It's a Different Game
Aluminum Bat Performance vs. Wood Bat Performance

Jeff Bagwell, Mike Bordick, Pie Traynor, Carlton Fisk, Mickey Cochrane, Tino Martinez, Eric Milton, Mark Mulder, Frank Thomas, and Mo Vaughn all have something in common. They are just a few who competed in the Cape Cod League, the premier amateur baseball program in the nation since 1885. In 2002 there were over 180 former Cape Cod League players on major league rosters, including American League Cy Young winner Barry Zito, AL Rookie of the Year Eric Hinske, and All-Stars Nomar Garciaparra, Todd Helton, and Lance Berkman.

In 1967, Thurman Munson led the league in hitting and was named the Cape's Most Valuable Player. A few other well-known MVPs over the years include Nat Showalter (better known as "Buck"), Steve Balboni, Ron Darling, Terry Steinbach, Greg Vaughn, Jason Varitek, and Darin Erstad. The award for the league's batting champion is named in Munson's honor, and among the recent winners are Bobby Kielty (1998), Lance Berkman (1996), Josh Paul (1995), and Lou Merloni (1992). The league established a Hall of Fame in 2000, and members include Mike Flanagan, Chuck Knoblauch, and Robin Ventura.

The nonprofit league was reorganized in 1963 with help from the NCAA and remains an NCAA-approved summer league, one of a growing number around the country, most of which use wooden bats. In 2003, the teams included the Bourne Braves, Brewster Whitecaps, Cotuit Kettleers, Falmouth Commodores, Harwich Mariners, Hyannis Mets, Orleans Cardinals, Wareham Gatemen, and Yarmouth-Dennis Red Sox.

The league is important to professional scouts because it doesn't allow the use of aluminum bats, though metal was used from 1975 to 1985. But the league returned to wood in 1986, and it has not looked back. The opportunity for collegiate players to prove they can make the switch from aluminum to wood makes the league more attractive and one reason it is much sought after by the best college prospects.

For the past six years I have been analyzing wood bat performance in the league compared to the high-tech aluminum bats used during the players' collegiate season. What follows is the results of that study. The parameters are: (1) only Division I hitters and pitchers were included; (2) hitters needed a minimum of 70 at-bats in the Cape Cod League to be included; (3) pitchers needed a minimum of 25 innings pitched. This means only regular players were considered.

In 2002, there were a total of 94 Division I hitters and 74 Division I pitchers who met the criteria. NCAA spring statistics (aluminum bat) of the players were compared to their summer Cape Cod League statistics (wood bats). Thus, the comparison is for the same players during the same year. While the level of players, both hitters and pitchers, is better in the Cape Cod League, the major variable is the bat.

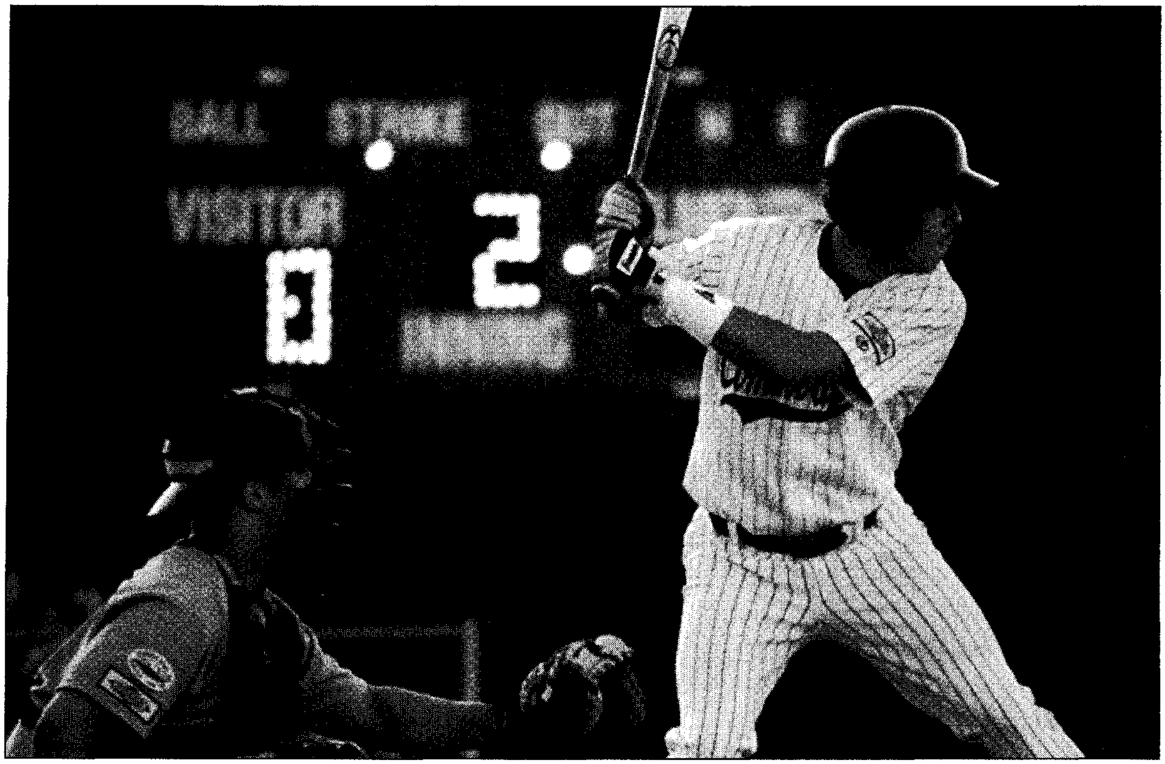
The difference in offensive performance in 2002 from the aluminum to the wood bat is dramatic. Using 94 Division I hitters, comparisons were made in the following offensive categories:

Table 1.

2002 HITTING STATS	ALUM.	WOOD	DIFF. (w/WOOD)
BA	.333	.231	-0.102
SLG	.518	.318	-0.200
HR/AB	1/29	1/75	-0.61%
R/AB	1/4.4	1/8.9	-0.51%
RBI/AB	1/4.9	1/10.3	-0.52%
K's%	15.8%	24.6%	+8.8%
BB	11%	9.4%	-1.6%

2002 PITCHING STATS	w/ALUM.	w/WOOD	DIFF. (w/WOOD)
ERA	4.28	2.48	-1.80
H/9	9.3	7.2	-2.1
K/9	7.3	8.4	+1.1
BB/9	3.1	3.0	-0.1
Opp. BA	.266	.220	-0.046

BILL THURSTON *has served as the head baseball coach at Amherst College for 39 years. He was the NCAA Baseball Rules Editor from 1985 to 2000, head coach for the Australian National team in 1984-85, and pitching coach for the USA team in 1986. He is a pitching consultant for the American Sports Medicine Institute in Birmingham, AL.*



Falmouth Commodores' shortstop Fernando Puebla readies to make contact versus the Hyannis Mets as Ryan Garko catches at Falmouth's Elmer "Guv" Fuller Field. Puebla is using a Barnstable bat, handmade in Cape Cod.

From this statistical analysis, it is unquestionably evident that during actual game use, aluminum bats dramatically outperform wood bats.

Furthermore, one has to question the reliability of the results of various lab tests comparing the performance of wood and aluminum. In the case of the Baum Hitting Machine Test used by the NCAA for the BESR standard, the test was compromised by comparing a heavier wood bat to a lighter aluminum bat. On a machine that has fixed-swing speed, both bats, regardless of weight, are swung at an identical speed; therefore, the more mass, the higher exit speed. Of course, a batter cannot swing a heavier unbalanced bat as fast or with as much control as with a lighter, better-balanced bat. Also, the Baum test standard was raised so that many of the 2000 aluminum bats could pass the test.

According to Professor James Sherwood (the test administrator), scaling the weight of the aluminum bat from 29.8 to 34 ounces could theoretically result in a ball exit speed of 113.25 mph, a difference of +16.37 mph with a comparable wood bat. Plus, metal bats have a total sweet spot much larger than wood (thus the ball is hit faster and more frequently). The sweet spot is the area on the bat that produces the maximum exit speed and causes the least amount of bat vibration.

I believe another factor with the Baum bat test that prevents the results from correlating to actual performance in the field is that the swing and pitch speeds are set so low that the resulting collision force of 132 mph does not trigger the trampoline effect in aluminum bats. The collegiate game is played at speeds of 160-180 mph.

Until the MOI (moment of inertia) of aluminum bats is mandated to match that of the various length of pro wood bats, the performance, and more important batted-ball exit speed, will never be like wood. It is obvious, even with the recent changes in aluminum bat standards by the NCAA, there continues to be major differences in bat performance on the playing field. The present MOI standard for aluminum bats is not comparable to the normal MOI of wood bats.

The following charts 2-7 compare the same 2002 players in their Division I games versus their performance in the Cape Cod League. This demonstrates the dramatic difference in actual games between aluminum and wood bat performance during the 2002 season only:

Table 2. 2002 batting average

BATTING AVG.	WITH ALUM. Nº	%	WITH WOOD Nº	%
.400 - .499	4	4%	0	–
.350 - .399	25	27%	0	–
.300 - .349	37	39%	6	6%
.250 - .299	23	24%	26	28%
.200 - .249	5	5%	33	35%
.150 - .199	0	–	24	25%
.100 - .149	0	–	5	5%

MAJOR DIFFERENCES	ALUM.	WOOD
% of batters BA .300+	70%	6%
% of batters BA -.200	0%	31%
BA Range high/low	.428/.212	.348/.143
Overall BA	.333	.231

Conclusion: BA difference of -.102 points using wood bats.

Table 3. 2002 slugging percentage

SLUGGING PCT.	WITH ALUM. Nº	%	WITH WOOD Nº	%
.700+	2	2%	0	–
.650 - .699	5	5%	0	–
.600 - .649	16	17%	0	–
.550 - .599	9	10%	0	–
.500 - .549	16	17%	1	1%
.450 - .499	16	17%	4	4%
.400 - .449	16	17%	11	12%
.350 - .399	9	10%	14	15%
.300 - .349	4	4%	23	24%
.250 - .299	1	1%	17	18%
.200 - .249	0	–	17	18%
.150 - .199	0	–	6	6%
.100 - .149	0	–	1	1%

MAJOR DIFFERENCES	ALUM.	WOOD
% of batters SLG .500+	51%	1%
% of batters SLG -.200	1%	44%
Overall SLG	.518	.318

Conclusion: SLG difference of -.200 points using wood bats.

Table 4. 2002 home runs per at-bat

HR/AB	WITH ALUM. Nº	%	WITH WOOD Nº	%
1/1 - 1/9	1	1%	0	–
1/10 - 1/19	21	22%	2	2%
1/20 - 1/29	17	18%	7	7%
1/30 - 1/39	12	13%	8	8%
1/40 - 1/49	7	7%	6	6%
1/50 - 1/59	5	5%	5	5%
1/60 - 1/69	8	8%	3	3%
1/70+	20	21%	26	27%
0/70+	3	3%	37	39%

MAJOR DIFFERENCES	ALUM.	WOOD
Overall HR Frequency	1/29 AB	1/75 AB
% with 1 or 0 HR per ≥70 AB	24%	67%
% with 1 HR per ≤29 AB	40%	10%

Conclusion: Home runs decreased 61% with wood bats.

Table 5. 2002 runs scored per at-bat

R/AB	WITH ALUM. Nº	%	WITH WOOD Nº	%
1/1 - 1/2.9	2	2%	0	–
1/3 - 1/4.9	58	62%	1	1%
1/5 - 1/6.9	26	28%	10	10%
1/7 - 1/8.9	8	8%	27	29%
1/9 - 1/10.9	0	2%	28	30%
1/11 - 1/12.9	0	–	11	12%
1/13 - 1/14.9	0	–	8	8%
1/15 - 1/16.9	0	–	4	4%
17+	0	–	5	5%

MAJOR DIFFERENCES	ALUM.	WOOD
% with 1 run per ≤5 ABs	64%	1%
Overall R/AB Frequency	1/4.4	1/8.9

Conclusion: R/AB difference of -.51% using wood bats.

Table 6. 2002 RBI per at-bat

RBI/AB	WITH ALUM. Nº	%	WITH WOOD Nº	%
1/1 - 1/2.9	3	3%	0	–
1/3 - 1/4.9	47	50%	1	1%
1/5 - 1/6.9	32	34%	7	7%
1/7 - 1/8.9	9	9%	20	21%
1/9 - 1/10.9	3	3%	15	16%
1/11 - 1/12.9	0	–	19	20%
1/13 - 1/14.9	0	–	13	14%
1/15 - 1/16.9	0	–	5	5%
1/17 - 1/18.9	0	–	3	3%
19+	0	–	11	12%

MAJOR DIFFERENCES	ALUM.	WOOD
% with 1 RBI per ≤5 ABs	53%	1%
% needing 9+ AB per RBI	3%	70%

Conclusion: Overall, hitters using aluminum bats drove in twice the number of runs than they did using wooden-bats.

Table 7. 2002 strikeouts per at-bat

K/AB	WITH ALUM. Nº	%	WITH WOOD Nº	%
41 - 45%	1	1%	4	4%
36 - 40%	0	–	5	5%
31 - 35%	1	1%	14	15%
26 - 30%	4	4%	22	23%
21 - 25%	16	17%	20	21%
16 - 20%	27	29%	19	20%
11 - 15%	28	30%	10	11%
6 - 10%	17	18%	0	–
0 - 5%	0	–	0	–

MAJOR DIFFERENCES	ALUM.	WOOD
Overall % of K	15.8%	24.6%
% of batters with ≥30% K	2%	30%
% of batters with < 20% K	77%	31%

Conclusion: Strikeouts were 8.8% more frequent with wood bats.

SIX-YEAR COMPARATIVE STUDY

Using the same criteria for Division I hitters and pitchers who played in the summer Cape Cod League for the past six seasons, the bat performance difference between aluminum and wood bats is consistent and dramatic. Same player, same year, the major variable, the bat.

Table 8. Comparison of batting average over six seasons

	ALUM.	WOOD	DIFF.
1997	.339	.232	-.107
1998	.329	.329	-.082
1999	.334	.248	-.086
2000	.325	.239	-.086
2001	.316	.232	-.084
2002	.333	.231	-.102
AVG.			-.091

Table 9. Comparison of home runs over six seasons

	ALUM.	WOOD	DIFF.
1997	1/25	1/74	-66%
1998	1/25	1/72	-.65%
1999	1/25	1/57	-.56%
2000	1/32	1/76	-.58%
2001	1/37	1/96	-.61%
2002	1/29	1/75	-.61%
AVG.			-.61%

Table 10. Comparison of slugging average over six seasons

	ALUM.	WOOD	DIFF.
1997	.551	.325	-.226
1998	.527	.350	-.177
1999	.542	.345	-.197
2000	.501	.330	-.171
2001	.470	.304	-.166
2002	.518	.318	-.200
AVG.			-.190

Table 11. Comparison of K% for hitters over six seasons

	ALUM.	WOOD	DIFF.
1997	17%	24%	+7%
1998	–	–	–
1999	–	–	–
2000	16%	22%	+6%
2001	15%	24.5%	+9.5%
2002	15.8%	24.5%	+8.8%
AVG.			+ 7.3%

Table 12. Comparison of runs scored per at-bat over three seasons

	ALUM.	WOOD	DIFF.
2000	1/4.5	1/7.4	-39%
2001	1/4.8	1/9.4	-49%
2002	1/4.4	1/8.9	-52%
AVG.			-49%

Table 13. Comparison of ERA over six seasons

	ALUM.	WOOD	DIFF.
1997	4.77	2.62	-2.15 or -45%
1998	5.01	3.50	-1.51 or -30%
1999	4.54	3.18	-1.36 or 31%
2000	4.11	3.15	-.096 or -24%
2001	4.34	2.25	-2.10 or -48%
2002	4.28	2.48	-1.80 or -42%
AVG.			-1.65 earned runs

Conclusion: Over a six-year period, pitchers gave up 1.65 fewer runs per game per pitcher when facing hitters using wood bats. That means on average there were 3.30 fewer earned runs each game.

PITCHING

Of the Division I pitchers, 74 pitched a minimum of 25 innings in the Cape Cod League during the 2002 season.

Table 14. 2002 earned run average

	WITH ALUM.		WITH WOOD	
K/AB	№	%	№	%
7.00+	3	4%	0	--
6.00 - 6.99	4	5%	0	--
5.00 - 5.99	14	19%	2	2%
4.00 - 4.99	23	31%	5	7%
3.00 - 3.99	17	23%	23	31%
2.00 - 2.99	10	13%	16	22%
1.00 - 1.99	3	4%	23	31%
0.00 - 0.99	0	--	5	7%

MAJOR DIFFERENCES	ALUM.	WOOD
% with ERA of 4.00+	59%	4%
% with ERA of 3.00 or less	17%	60%
Overall ERA	4.28	2.48

Conclusion: Each pitcher facing wood bats had a decrease of 1.80 earned runs per game, so two teams' pitchers had 3.60 fewer earned runs per game than those facing aluminum bats.

Table 15. Hits allowed per nine innings pitched in 2002

	VS. ALUM. (COLLEGE)	VS. WOOD (CAPE LEAGUE)
H/9 of 9 hits or less	34%	83%
H/9 of 7 hits or less	5%	37%
BA against these pitchers	.266	.220

WWW.SPORTSPIX.BIZ

Wareham at Hyannis in August 2002.

Also, hitters using aluminum bats hit twice as many home runs against these pitchers as when they used wood bats. Remember, the college offensive lineups are not nearly as talented as the hitters playing in the Cape Cod League.

RISK OF INJURY FROM BATTED BALLS

While this study focused on the different performance levels between aluminum and wood bats, another major problem that needs to be addressed is the increased risk of injuries from batted balls off the high-performance aluminum bats. It is well-documented from lab tests, field tests, and various studies that the ball is hit with greater velocity—and hit faster, more frequently—with an aluminum bat. There are many reasons for this:

1. Factors of increased batted-ball exit speed.
   a. Greater swing speed. A hitter can swing a lighter, better-balanced bat faster than a heavier or end-heavy bat.
   b. The trampoline and hoop effect of a thin, hollow tube versus a solid wood bat.
   c. The balance point (MOI) of an aluminum bat is closer to the handle, allowing greater head-of-the-bat speed and better bat control. Note: In a test completed at Amherst College in November 2002, six hitters had 90 recorded hits off a tee (stationary ball) with both wood bats (2) and aluminum (2) bats. The average increase in batted-ball exit speed was 5.9 mph when using aluminum bats. This is a major and significant increase because with pitched-ball speed and the trampoline effect added, the batted-ball exit speed off aluminum will dramatically increase over that of wood.

2. The ball comes off an aluminum bat faster and faster more frequently.
   a. The diameter of the bat is larger and stays larger, longer, down the length of the bat.
   b. Sherwood states that the sweet spot is four times larger than in a normal wood bat.
   c. The bat is better balanced. A hitter can control the swing better, start the swing later, and track the pitch longer (good contact more frequently leads

to higher batting averages).

d. Over a six-year period, hitters struck out 7.3% more often when using wood bats. With strikeouts, pitchers avoid the risk of being struck and injured by a batted ball.
e. The Cape Cod League study demonstrates how differently the game is played using wood. Over a six-year period:
  - Batting averages decreased by .091 points.
  - Slugging percentage decreased by .190 points.
  - Home runs decreased by 61%.
  - Earned run average decreased by 1.65 per pitcher, and 3.30 per game.
f. Players are now bigger and stronger than years ago and hit the ball harder and faster when they make good contact. Fielders do not have as much time to field the batted ball or to defend themselves.
g. If a player is struck by a batted ball hit with greater velocity, chances are there will be a more severe injury.

If experienced major league pitchers are being struck by line drives off wood bats, then we should not be using bats that outperform wood. Amateur pitchers are less experienced than pro pitchers and often are not aware of the danger. They don't anticipate and react properly.

Against aluminum bats, most pitchers cannot effectively pitch inside with the fastball; they have to pitch away more often. There are more balls hit up the middle (toward the pitcher) on outside pitches than on inside pitches. Again, this factor increases the risk of injury.

By 1998, the NCAA Umpire Improvement Program Committee was so concerned about the safety of their field (base) umpires that they instituted an umpire position change. They moved the field umpires back, farther away from home plate to give them more time to react to batted balls.

MAJOR QUESTIONS

Why don't various lab test results correlate to bat performance in the field, in games, or even during batting practice? Could all of these injuries we have experienced since the mid '90s through 2002 have occurred with balls hit off wood bats? Maybe, but we will never know. Chances are that they would have occurred less frequently with wood bats: The pitchers would have more time to react, since off wood the ball is not hit as fast or hit hard as frequently.

I believe that based on the statistical evidence of this study, and statistics developed over the past six years, my conclusions are convincing. The collegiate game played with the present high-performance aluminum bats is not remotely close to the traditional game played with wood bats. For the safety of the players, to bring the game back in balance, and to restore the integrity of the game, I hope that in the near future that true wood-performance standards are put in place at all levels of amateur baseball.

REVERSAL OF FORTUNE **Submitted by David MacDonald**

"Manager [Harry] Davis of the Cleveland Ball Team has barred crap shooting by his players. 'I don't object to a little game of poker, but I draw the line at craps.' The Cleveland manager claims that players don't play well when they owe each other money for gambling debts." —*Ottawa Journal*, March 12, 1912

(Davis's team, featuring Joe Jackson and Nap Lajoie, went 54-71 before he was axed. Joe Birmingham took over and the team went 21-7 the rest of the way.)

W. C. MADDEN

The Rise and Fall of Scoring at the College World Series

When the College World Series began in 1947, the wooden bat was the only weapon in the batter's arsenal. Scores were typical of the times with scoring for the two teams averaging about eight to eleven runs per game. By 1972 and 1973, that average had dipped to seven runs per game, a number that was soon history as the NCAA made some changes.

In 1974, the NCAA decided to allow hitters to use aluminum bats during the season and in the CWS. The ostensible reason was saving money: metal bats were much more expensive, but they outlasted their wooden counterparts ten times over. The metal monsters didn't break and they were only pulled when they developed flat spots. The NCAA also decided that year to implement the designated hitter rule like that used in the American League. Pitchers would no longer be required to bat.

The seven runs per game average became a memory, as did almost all the batting records. Within six years every record was shattered, not counting the first three years of the series because of the limited amount of games in those days.

The first record to go was the doubles mark of 34, topped by one in 1975. That record continued to be broken. Next to go was the home run record; the mark of 17 was broken with 21 in 1976. That record continued to rise as coaches turned to the long ball more for run production. The highest batting record of .266, (except for .314 in 1947 when only two teams played) set in 1971, was broken in 1977, when the average moved up to .284. That record continued to rise as well.

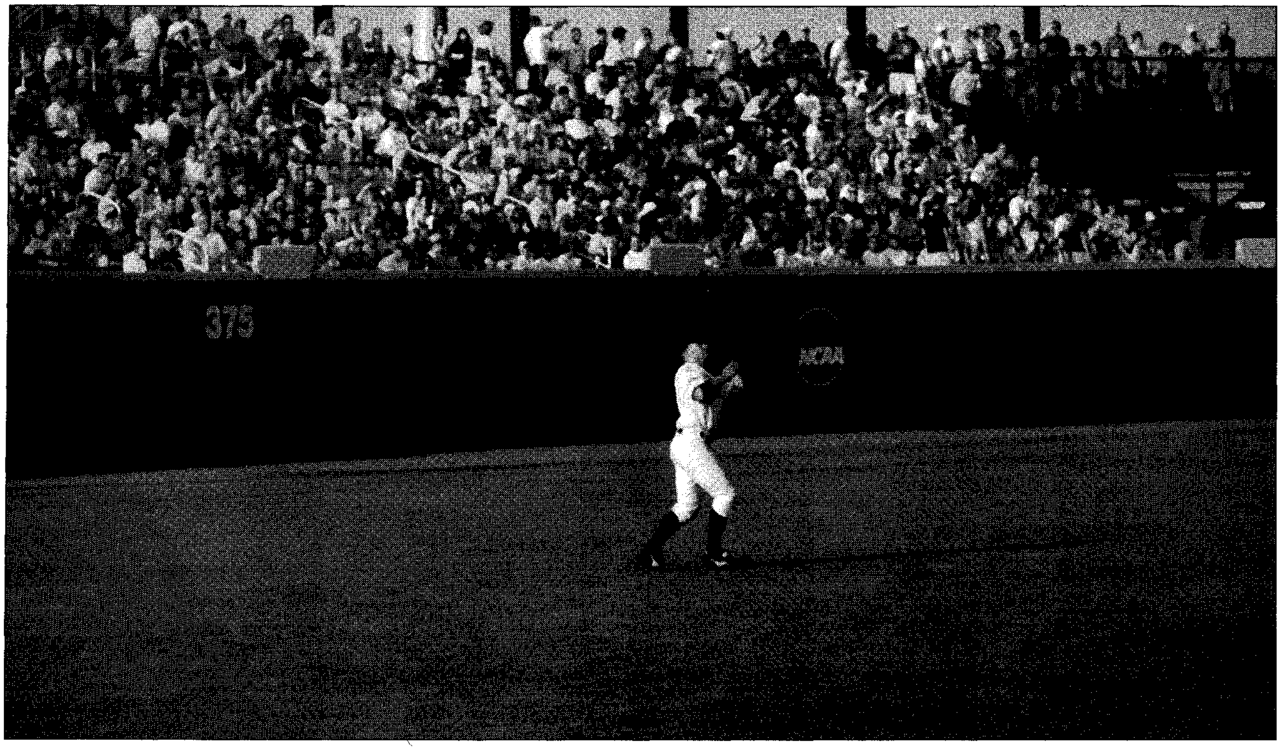
The slugging percentage was also easily surpassed that year as the benchmark of .365 was raised to a robust .419. That record also continued to rise through the years. The average number of runs record (11.5) was broken in 1979, when 11.6 runs were scored. That too continued to increase through the years, too. Corresponding with that was a rise in the RBI record as well (144 to 153). The hits record (272) went by the wayside in 1979, when 294 hits were made. The only record that did not fall and still stands today is the triples mark: the record is 22 three-base hits in 1962. Interestingly, the number of triples has decreased steadily through the years till it reached less than 10 per year after 1992.

While the hitting records generally climbed, pitching records, not unexpectedly, went in the other direction. Before 1974, the highest ERA was 4.32. That was broken in 1978 (4.78) and continued to rise. Shutouts took a nosedive and, mirroring the majors, the number of complete games rapidly decreased. Before 1973, an average of about 13 complete games were registered per year out of the 15 games held. But 1977 was the last year that complete games were in the double digits. Before 1974, the record number of shutouts was six; since then the high has been three. For one six-year stretch, 1996 to 2001, no shutouts were thrown. Nowadays, complete games and shutouts are a rarity, but on the flip side, the number of saves has risen.

The eclipsing of hitting records reached a peak in 1998, when the series was dubbed "gorilla ball" because of the prolific scoring. Bat manufacturers were creating more powerful bats, and colleges were producing prolific home run hitters. It read like a football score in the championship game when USC whipped Arizona State, 21-14. Wes Rachels knocked in a record seven runs for USC in the game. The 1998 Series shattered numerous records: batting average (.318), hits (327), home runs (62), RBI (212), total bases (558), and slugging average (.543). The average number of runs went up to 16 a game. ERA zoomed to 7.24. With no complete games or shutouts recorded, the save record was also broken with nine.

By 1998 many of the individual hitting records had also been broken. Bud Hollowell's record of four

W. C. MADDEN *of Fishers, Indiana is the author of the* College World Series Record Book *and a history of the College World Series, which will be published by McFarland later this year. He has also written several other baseball books, including* History of the Western League *and two books about the All-American Girls Professional Baseball League.*



Rosenblatt Stadium's power alley in left center used to be 360 feet before the walls were moved back before 2002 to 375.

home runs in a Series in 1963 was tied six times by 2000. J. D. Drew of Florida State and Edmund Muth of Stanford each hit three home runs in a game for a record. Drew also set the most total bases record (12), which was later tied. Even the doubles record went by the wayside in 2003, when Ryan Garko of Stanford hit four in one game. LSU set the record for most home runs (8) in a 12-10 win over USC in 1998.

The NCAA decided it was high time to make some changes to tone down the scoring barrage. They put limits on the bat manufacturers and moved the walls back in Omaha before the 2002 Series. Rosenblatt Stadium officials moved the fences back fifteen feet in the power alleys and three feet down the lines. The height of the walls varied between two feet and ten feet in center and eight feet elsewhere. A balance returned as averages declined.

By 2003, coaches were talking more about "small ball," instead of "long ball." Even the home run hitting contest was discontinued. The only offensive record eclipsed in 2003 was hit batsmen.

ALAN NATHAN

Bat Performance Standards in NCAA Baseball

Starting with the 2000 season, the NCAA requires that bats be subjected to performance testing. Bats are tested by impacting a ball moving at 70 mph with a bat rotating about a point 6" from the knob (toward the barrel) such that the speed of the bat at a point 6" from the barrel (toward the knob) is 66 mph.[1]

The post-impact ball speed is measured at different impact locations along the barrel. Bats are certified for use in officially sanctioned NCAA games if the ball exit speed does not exceed 97 mph at any of the impact locations, corresponding to a BESR of 0.728 (see corresponding article). The 97 mph upper limit was arrived at by testing a large sample of different wood bats, where it was found that the maximum exit speed was 96 mph. The maximum for aluminum was increased by 1 mph to allow for uncertainties in the measurement. This means that *for bats swung with comparable speed*, the ball exit speed is essentially the same for wood and aluminum bats.

However, this does not mean that wood and aluminum bats will necessarily perform identically in the field, because wood and aluminum bats are not swung at comparable speeds. Even for bats with the same weight, the *weight distribution* is generally very different for a wood and aluminum bat; a typical wood bat has more of its weight concentrated in the barrel and farther from the hands.

One way to characterize the weight distribution is the so-called moment of inertia (MOI), which is a measure of how far the weight is concentrated from the hands. A bat with a smaller MOI has the weight concentrated closer to the hands and will be easier to swing. Likewise, a bat with a larger MOI will have the weight farther from the hands and will be harder to swing. Typically, aluminum bats of a given length and weight have a smaller MOI than a wood bat with the same length and weight.

There is now a growing amount of scientific evidence which shows an inverse relationship between the MOI of a bat and the speed with which it can be swung.[1] The smaller the MOI, the faster the bat can be swung. Since an aluminum bat generally has a smaller MOI than a wood bat of comparable length, an aluminum bat can be swung faster and will therefore perform better in the field.

Recognizing the importance of MOI for bat performance, the NCAA supplemented their impact testing with restrictions on the weight and the MOI of bats. The weight restriction is the so-called "-3 rule," which means that a 34" bat must weigh at least 31 ounces. The MOI restriction is also based on length. For example, a 34" bat can have an MOI no smaller than 9700 oz-in;[2] for reference, a typical wood bat of that length has an MOI of about 11,000 oz-in.[2] Even with this restriction, aluminum bats can be expected to outperform wood bats in the field. As an example, consider a wood bat with an 11,000 MOI performing at 96 mph in the test and an aluminum bat with a 9700 MOI performing at 97 mph in the test. Using the Fleisig swing-speed data, one can estimate that with the higher swing speed, the aluminum bat will perform at about 101.5 mph in the field, or 5.5 mph faster than the wood bat. This gives rise to an additional 20-30 feet on a long fly ball.

ALAN NATHAN *is a Professor of Physics at the University of Illinois at Urbana-Champaign and the Associate Head of the Physics Department. He earned his Ph.D. at Princeton in 1975 and is a Fellow of the American Physical Society. His primary research is in the quark structure of protons and neutrons, but he moonlights doing work in the physics of baseball. He is a member of the scientific panel that advises the NCAA on issues regarding baseball bat performance.*

NOTES

1. The procedures are described in the September 1999 NCAA report (www.ncaa.org/releases/miscellaneous/1999/1999092901ms.htm).
2. See, for example, the paper of Fleisig, et al., published in the journal *Sports Engineering*, vol. 5, pp 1-14, 2002.

JAMES ASHTON-MILLER, MICHAEL CARROLL, KENNETH JOHNSON, AND ALAN NATHAN

Ball Exit Speed Ratio (BESR)

The NCAA requires that all nonwood bats be certified so as to limit their "liveliness." The certification process is accomplished by measuring the performance of a bat under controlled conditions and then assigning a number to it; this number is known as the BESR (Ball Exit Speed Ratio). To be certified, the BESR of the bat must fall at or below a predetermined value set by the NCAA. This paper discusses the concept of the BESR.

THE BALL-BAT COLLISION

Figure 1 shows a ball and a bat just before the collision and the ball just after the collision (the position of the ball after the collision has been moved downward for the sake of clarity). The speeds involved in the collision are:

V_{PITCH} = speed of the pitched ball just before it collides with the bat.

V_{BAT} = speed of the bat just before it collides with the ball. This is the bat speed at the point of impact.

$V_{BALL\ EXIT}$ = exit speed of the ball just after it leaves the bat.

WHAT IS THE BESR?

The BESR is a number, once known, that allows one to determine the ball exit speed $V_{BALL\ EXIT}$ when the bat speed V_{BAT} and the pitch speed V_{PITCH} are specified. The relationship between the BESR and these speeds is:

$$V_{BALL\ EXIT} = (BESR + \quad) V_{BAT} + (BESR - \quad) V_{PITCH} \qquad (1)$$

DR. KENNETH JOHNSON *is a Professor of Physics, Emeritus at Southern Illinois University.* **DR. JAMES ASHTON-MILLER** *is a Research Professor in the Departments of Mechanical Engineering and Biomedical Engineering at the University of Michigan. He is a past president of the American Society of Biomechanics.* **DR. MICHAEL CARROLL** *is former Dean of the George Brown School of Engineering at Rice University. These four professors comprise the NCAA Baseball Research Panel.*

As an example, suppose the BESR for a particular ball-bat collision is 0.65, and that the bat and pitch speeds are V_{BAT} = 70 mph and V_{PITCH} = 75 mph. The ball exit speed would be

$$V_{BALL\ EXIT} = (0.65 + \quad)(70mph) + (0.65 - \quad)(75mph) = 92mph$$

Conversely, if one measures the bat speed, the pitch speed, and the ball exit speed, then Equation 1 can be used to determine the BESR (see Equation 2 below).

Note from Equation 1 that greater values of the BESR give rise to greater ball exit speeds. Therefore, the BESR is a measure of the "liveliness" of the ball-bat collision and it includes, for example, any "trampoline" effect that the non-wood bat may display (due to its barrel being temporarily deformed by the ball during the collision).

WHERE DOES BESR GET ITS NAME?

When one algebraically solves Equation 1 for the BESR the result is

$$BESR = \frac{V_{BALL\ EXIT} + \quad (V_{PITCH} - V_{BAT}\ )}{(V_{PITCH} + V_{BAT}\ )} \qquad (2)$$

When the speeds of the pitched ball and bat are the same ($V_{PITCH} = V_{BAT}$), Equation 2 becomes

$$BESR = \frac{V_{BALL\ EXIT}}{V_{PITCH} + V_{BAT}}$$

We see in this case that the BESR is equal to the ratio of the ball exit speed $V_{BALL\ EXIT}$ to the relative speed ($V_{PITCH} + V_{BAT}$) of the pitched ball and bat before the collision. Hence, the name "Ball Exit Speed Ratio."

HOW BESR DEPENDS ON THE PROPERTIES OF THE BALL AND BAT

Figure 2 illustrates a ball just before colliding with the bat. The bat is assumed to be clamped in a hitting machine and is free to rotate in the plane of the paper about the pivot point.

The physics of the collision is described by applying the law of conservation of angular momentum to the ball-bat interaction. When this law is used, along with the definition of the coefficient of restitution (see below), we arrive at Equation 1, where the BESR is given in terms of the properties of the ball and bat [1,2,3]:

$$\mathrm{BESR} = \frac{e + \frac{1}{2}\left(1 - \frac{mr^2}{I_p}\right)}{1 + \frac{mr^2}{I_p}} \qquad (3)$$

where:

- e = coefficient of restitution of the ball-bat collision. The coefficient of restitution is defined as the ratio of the relative speed of the ball and bat after the collision to that before the collision. Suppose that, before the collision, the ball and bat are moving toward each other with a relative speed of 160 mph. Suppose, further, that after the collision the ball and bat are moving with a relative speed of 80 mph. Then the coefficient of restitution of the ball-bat collision is (80 mph)/(160 mph) = 0.5.
- m = mass of the ball.
- r = distance from the pivot point to where the ball hits the bat (see Figure 2).
- I_P = moment of inertia of the bat about the pivot point. This parameter depends on the mass of the bat as well as how the mass is distributed relative to the pivot point. The more the mass is concentrated away from the pivot point, the larger is the moment of inertia.

Note that BESR depends on the properties of the ball (m), the bat (I_P), and the ball-bat collision (e and r).

WHY USE THE BESR RATHER THAN SPECIFY A BALL EXIT SPEED?

In general, different bat testing laboratories use different types of hitting machines: (1) the pitched ball is moving and the bat is initially stationary, (2) the ball is stationary and the bat is initially moving, and (3) both the pitched ball and bat are initially moving. Even if each type of hitting machine is set up to have the same relative speed ($V_{PITCH} + V_{BAT}$) of the pitched ball and bat, the ball exit speeds will be different. However, all types of machines will give the same value for the BESR. This result, while not obvious, is a direct consequence of Equation 1.

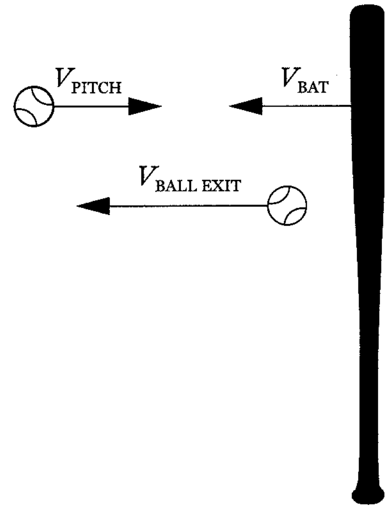


Figure 1. The ball-bat collision.

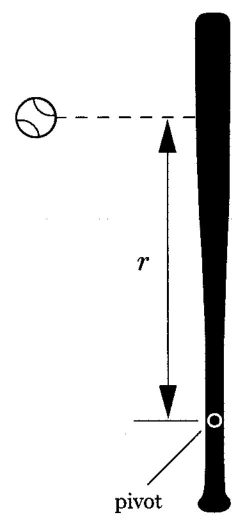


Figure 2. The bat pivot point and the distance r from the pivot point to where the ball collides with the bat.

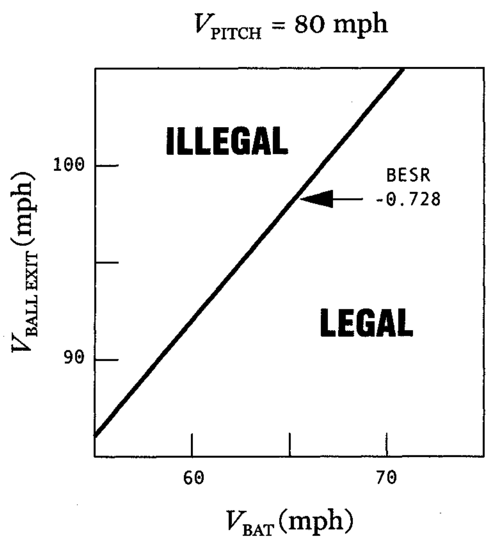
WHAT IS THE MAXIMUM ALLOWED VALUE FOR THE BESR?

When bats were first tested in 1999, an initial lot of baseballs was used. The tests were conducted by using a pitch speed of 70 mph and a bat speed (at a point 6 inches from the end of the barrel) of 66 mph. Under these conditions, the best major league wood bat yielded a BESR of 0.728, which the NCAA then set to be the maximum allowed value.

Figure 3 shows a plot of ball exit speed ($V_{BALL\ EXIT}$) versus bat speed (V_{BAT}) for the case when the pitch speed is V_{PITCH} = 80 mph. The straight line represents Equation 1 in which the BESR has been set to the legal limit of 0.728. Any bat that gives rise to a ball exit speed at or below this line is legal. Likewise, any bat that produces a ball exit speed above this line is illegal.

Subsequent tests on non-wood bats used different lots of new baseballs. Because the properties of balls differ from lot to lot, even when they are stored and used in a humidity-controlled room, the BESR is adjusted to account for these differences. Therefore, the maximum allowed value for the BESR changes slightly, depending on the particular lot of baseballs used in testing a given non-wood bat. However, in every case, the BESR of the non-wood bat is always compared with that of major league wood bats tested in the same machine with the same lot of baseballs under standardized ball-bat testing conditions.

Figure 3. Plot of ball exit speed vs. bat speed when the pitch speed is 80 mph



NOTES

1. Carroll, M. M. "Assessment and regulation of baseball bat performance," *Symposium on Trends in the Application of Mathematics to Mechanics*, edited by P. E. O'Donoghue and J. N. Flavin (Elsevier, Amsterdam, 2000), p. 17.
2. Nathan, A. M. "Dynamics of the baseball-bat collision," *Am. J. Phys.* 68, 979–990 (2000).
3. Nathan, A. M. "Characterizing the performance of baseball bats," *Am. J. Phys.* 71 (2), 134–143 (February 2003).

J. ERIC BICKEL, Ph.D.

Why It's So Hard to Hit .400

New Insights into an Old Statistic

Batting average measures the batter's ability to avoid striking out and his ability to "hit 'em where they ain't." However, the relative importance of these two skills is blurred by the traditional way in which we write AVG (H/AB). By rewriting AVG as a function of strikeouts per at-bat and hits per ball put in play, we gain greater insight into what has driven AVG over the last 100 years, why hitting .400 is so difficult, and whether or not batters can really place base hits.

To see the impact of strikeouts and putting the ball in play, let's write AVG in terms of these effects, as shown in Equations 1 and 2. BIP is the number of times the batter put the ball in play—not including strikeouts.

$$\mathrm{AVG} = \frac{\mathrm{H}}{\mathrm{AB}} = \frac{\mathrm{H}}{\mathrm{AB}} \times \frac{\mathrm{BIP}}{\mathrm{BIP}} = \frac{\mathrm{BIP}}{\mathrm{AB}} \times \frac{\mathrm{H}}{\mathrm{BIP}} \qquad (1)$$

From Equation 1 we see that AVG is equal to the fraction of at-bats where the ball was put in play (BIP/AB) multiplied by the fraction of balls put in play that went for hits (H/BIP). This is intuitive—first the batter must put the ball in play and then it has to "become" a hit.

Since there are only two possible ways to have an at-bat: (1) the ball is put in play or (2) the batter strikes out, BIP/AB = 1 - K/AB, where K/AB is the fraction of at-bats where the batter struck out. Let's call the term K/AB the batter's "strike out average" or KAVG and H/BIP his "in-play average" or IPAVG. We can then write Equation 1 as

$$\mathrm{AVG} = (1 - \mathrm{KAVG}) \times \mathrm{IPAVG} \qquad (2)$$

Several other authors have referred to IPAVG as hits per ball in play (HPBP) or batting average per batted ball.[1] I prefer IPAVG because it more closely ties with "AVG", where it plays a major role, and yet is not really a "batting average" because it rules out the possibility of a strikeout.[2]

To prove that Equation 2 works, we can calculate Ted Williams' 1941 batting average, which was .406. In 1941, Ted Williams had 456 at-bats, struck out 27 times, and tallied 185 hits. He put the ball in play 429 times (456−27), not including sacrifices. Therefore, his KAVG = K/AB = 27/456 = .059 and his IPAVG = H/BIP = 185/429 = .431. According to Equation 2, his AVG was (1 − .059) × .431 = .941 × .431 = .406.

Before moving on, we should note that Equation 2 makes it clear why batting average by count is misleading. If batting average is calculated for all counts with less than two strikes, then KAVG = 0 because no at-bat ends in a strikeout with less than two strikes. Therefore, AVG with less than two strikes is simply equal to IPAVG. Consider the following two situations: (1) Barry Bonds had an at-bat with less than two strikes and (2) Barry Bonds had an at-bat with two strikes. In which case is it more likely Bonds got a hit? When he had less than two strikes because we are sure he did not strike out. Comparing AVG across two-strike and non-two strike counts is meaningless! Yet, STATs Inc continues to publish AVG by count.[3] For more detail, please see Bickel and Stotz (2002).[4]

The same reasoning that we used to develop Equation 2 can be applied to SLG. In this case, SLG measures the batter's ability to avoid striking out and the number of bases he obtains when he puts the ball in play. This is shown in Equation 3.

$$\mathrm{SLG} = 1 - (\mathrm{K} - \mathrm{AVG}) \times \mathrm{IPSLG} \qquad (3)$$

In Equation 3 we have called Bases/BIP the "In-Play Slugging Average" or IPSLG.

Equations 2 and 3 suggest many interesting questions. What are the average KAVG, IPAVG, and IPSLG for major league hitters? Has this changed over time? Has the change in AVG over the last 100 years been driven by changes in KAVG, IPAVG, or both? What

J. ERIC BICKEL *is President of Competitive Edge Decision Systems, which provides pitch/hit-charting software to amateur and professional baseball teams.*

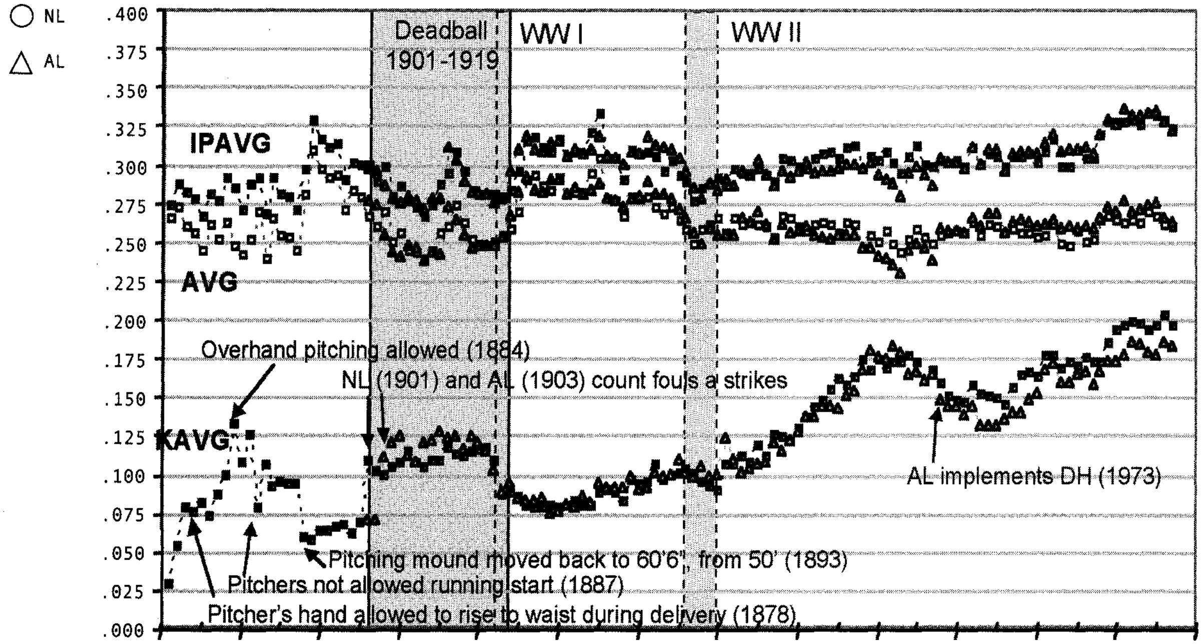


Figure 1. *KAVG, IPAVG and AVG for the AL and NL since 1876.*[5]

does it take to hit .400? Did Ted Williams hit .406 by avoiding strikeouts, or was he better at hitting them where they ain't? KAVG seems to be more under the control of the batter than IPAVG, in that IPAVG is driven by the structure of the game (e.g., the number and location of fielders, field dimensions, etc.) and, to some degree, just plain luck. Do some batters really have the ability to "place" the ball so that they get more hits when they put the ball in play? Can a major league hitter improve his AVG more by improving his KAVG or IPAVG? Do pitchers differ in the IPAVG and IPSLG that they allow?

In baseballprospectus.com, Voros McCracken reported that pitchers differ very little in their ability to prevent hits once the ball is put in play. In our words, they differ little in the IPAVG that they allow. Many other researchers have confirmed this effect.

KAVG AND IPAVG: A HISTORICAL PERSPECTIVE

Figure 1 displays KAVG, IPAVG, and AVG for the AL and NL since 1876. The lowest data set is KAVG. The circles represent the NL and the diamonds the AL. The middle data set is AVG for the NL and AL. And, finally, the top dataset is IPAVG for the NL and AL.

The Deadball Era (1901-1919) is clearly visible in the IPAVG data, with two dramatic declines—which would not be matched until World War II (WW II). It is interesting to note that strikeouts were relatively high during this era. The KAVGs recorded during this time, would not be seen again until the mid-1950s. It is generally agreed that Deadball Era ended in 1920 with the banishment of the spitball and other "freak" pitches.[6] However, KAVGs began to collapse as early as 1917, and IPAVGs began their dramatic increase in 1919 (mostly in the AL). These years coincide with America's involvement in WW I, but we are unsure as to the relative significance of WW I (e.g., loss of key players), the banishment of the spitball, and other factors.

IPAVG has generally been above .300 since the 1950s. NL and AL IPAVG data sets are similar because the rules governing balls put in play have been consistent between the two leagues. The AL gained a slight advantage over the NL with the introduction of the DH in 1973.

The trend of IPAVG has been generally upward since 1977 and is currently around .322 or almost 33%, with a dramatic increase starting in 1993. An IPAVG of about 33% means that the average MLB batter has about a one-third chance of getting a hit if he puts the ball in play. In a statistical sense, about two-thirds of the field is covered by the defense. Interestingly, 0.333 is very close to the IPAVG that Dean Stotz and I found in college baseball.[7] Clearly, IPAVG is driven to a very large degree by the structure of the game (e.g., the shape of the field, the number of fielders, the distance to the outfield wall).

KAVG has been more variable and swung wildly between 1876 and 1901, as the NL tinkered with the number of balls allowed before a walk, the distance from the pitcher's mound to home, and whether or not

Table 1. Impact of some rule changes since 1876

SEASON	RULE CHANGE	KAVG	IPAVG	AVG
1876	Initially a strikeout effectively requires 4 strikes Initially a base on balls effectively requires 9 balls			
1878	Pitcher's hand must stay below waist during delivery, instead of below hip	+49%	-2%	-4%
1880	Number of balls required for a walk decreased to 8	+7%	-4%	-4%
1881	Pitcher's mound moved back from 45' to 50' Number of balls required for a walk decreased to 7 Number of strikes required for a strikeout reduced to 3 on a called third strike	-11%	+5%	+6%
1883	Pitcher's hand must stay below shoulder during delivery, instead of below waist Foul balls caught on one bounce no longer count as outs	+14%	+6%	+4%
1884	Overhand pitching allowed Number of balls required for a walk reduced to 6	+34%	-2%	-6%
1886	Number of balls required for a walk increased to 7	+17%	+6%	+4%
1887	Pitchers no longer allowed a running start, may only take one step Number of balls required for a walk reduced to 5	-37%	+2%	+7%
1888	Number of strikes required for a strikeout reduced to 3	+36%	-8%	-11%
1889	Number of balls required for a walk reduced to 4	-13%	9%	11%
1892	Outfield fence must be at least 235' from home plate	--	-3%	-3%
1893	Only round bats may be used Pitcher's mound moved back from 50' to 60'6", however, mound could be raised	-37%	+10%	+14%
1894	Foul bunts count as strikes for the first time	-2%	+10%	+11%
1895	Foul tips count as strikes if caught within the 10' catcher's box	+10%	-4%	-4%
1901	NL counts foul balls as strikes for the first time	+58%	--	-4%
1903	Mound height limited to 15" AL counts foul balls as strikes for the first time	-3% +55%	+4% -3%	+4% -7%
1911	Introduction of cork-centered baseball	+1%	+7%	+7%
1920	Foreign substances may not be applied to ball. Ball may not be deliberately scuffed	-6%	+5%	+5%
1926	Pitchers may use rosin bags Outfield fence must be at least 250' from home plate	+4%	-3%	-4%
1969	Mound height limited to 10" Strike zone reduced (armpits to top of knees)	-3%	+4%	+5%
1973	AL introduces the designated hitter	-10%	+6%	+9%

foul balls counted as strikes. Figure 1 highlights some of the rule changes. Table 1 presents a more exhaustive list of rule changes and what happened to KAVG, IPAVG and AVG.[8]

Changes to the pitcher's delivery, reducing the number of strikes required for a strikeout and counting foul balls as strikes had dramatic effects on strikeouts. We also see that IPAVG increased by 7% after the introduction of the cork-centered baseball in 1911. However, the Deadball Era resumed after this increase.

It is interesting that KAVG has doubled during the last 50 years; batters are twice as likely to have their AB end in a strikeout now than they were in 1945. This increase in KAVG has tended to hold batting averages in check.

HITTING .400

What does it take to hit .400? Equation 2 shows that at a minimum the batter must get a hit 40% of the time he puts the ball in play (IPAVG = .400). In this case, he will hit .400 *if* he never strikes out. If he strikes out in 10% of his ABs (KAVG = .100), he would need to have an IPAVG of .444 ($.4 \div [1 - .9]$)—or he would have to get a hit 44.4% of the time he put the ball in play.

A KAVG of .100 is quite low by today's standards, being about half of the current 0.190 average. Furthermore, an IPAVG of .444 is astronomical by any historical measure. Table 2 presents the top 20 IPAVGs since 1913 (note: we have not included seasons prior to 1913 because individual strikeout data is unavailable in the AL until this time).[9]

Table 2 shows that IPAVG has been over .444 in only 11 player seasons since 1913 (about .12%). We also see that home run hitters have generated most of the top IPAVGs. This is understandable given that balls hit into the stands do not have the possibility of being caught. The record IPAVG of .478 belongs to Manny Ramirez. However, he "only" hit .351 because almost 27% of his ABs ended in strikeouts. Babe Ruth compiled five of the top 20 IPAVG seasons. Only three of the top 20 IPAVGs have resulted in AVGs over .400—all of them by Rogers Hornsby.

Table 3 runs down the top 20 batting averages since 1913. No batter has hit .400 or higher with an IPAVG below .420 (Ty Cobb, 1922). Furthermore, no .400 AVG has ever been accompanied by KAVG of greater than .080 (Rogers Hornsby, 1922). This is about one-third the current MLB average.

A COMPARISON OF WILLIAMS, BRETT AND GWYNN

Before leaving this historical look at IPAVGs and KAVGs, let's look at four more incredible player-seasons: Ted Williams 1941, Ted Williams 1957, George Brett 1980, and Tony Gwynn 1994.

In 1941, Ted Williams' KAVG was 42% lower and his IPAVG 45% greater than the league average (see

Table 2. Top 20 IPAVGs since 1913

PLAYER	LEAGUE	YEAR	AB	AVG	KAVG	IPAVG
Manny Ramirez	AL	2000	439	.351	.267	.478
Babe Ruth	AL	1923	522	.393	.178	.478
Babe Ruth	AL	1920	458	.376	.175	.455
Rogers Hornsby	NL	1924	536	.424	.060	.450
Jim Thome	AL	2001	526	.291	.352	.449
Jose Hernandez	NL	2002	525	.288	.358	.448
Babe Ruth	AL	1924	529	.378	.153	.446
Sammy Sosa	NL	2001	577	.328	.265	.446
Mo Vaughn	AL	1997	527	.315	.292	.445
Manny Ramirez	AL	1999	522	.333	.251	.445
Babe Ruth	AL	1921	540	.378	.150	.444
Sammy Sosa	NL	2000	604	.320	.278	.443
Larry Walker	NL	2001	497	.350	.207	.442
Mo Vaughn	AL	1998	609	.337	.236	.441
Babe Ruth	AL	1926	495	.372	.154	.439
Jeff Bagwell	NL	1994	400	.368	.163	.439
Andres Galarraga	NL	1993	470	.370	.155	.438
Rogers Hornsby	NL	1925	504	.403	.077	.437
Rogers Hornsby	NL	1922	623	.401	.080	.436
Lou Gehrig	AL	1927	584	.373	.144	.436

Table 3. Top 20 AVGs since 1913

PLAYER	LEAGUE	YEAR	AB	AVG	KAVG	IPAVG
Rogers Hornsby	NL	1924	536	.424	.060	.450
George Sisler	AL	1922	586	.420	.024	.430
George Sisler	AL	1920	631	.407	.030	.420
Ted Williams	AL	1941	456	.406	.059	.431
Rogers Hornsby	NL	1925	504	.403	.077	.437
Harry Heilmann	AL	1923	524	.403	.076	.436
Rogers Hornsby	NL	1922	623	.401	.080	.436
Bill Terry	NL	1930	633	.401	.052	.423
Ty Cobb	AL	1922	526	.401	.046	.420
Lefty O'Doul	NL	1929	638	.398	.030	.410
Harry Heilmann	AL	1927	505	.398	.032	.411
Rogers Hornsby	NL	1921	592	.397	.081	.432
Tony Gwynn	NL	1994	419	.394	.045	.413
Harry Heilmann	AL	1921	602	.394	.061	.419
Babe Ruth	AL	1923	522	.393	.178	.478
Harry Heilmann	AL	1925	573	.393	.047	.412
Babe Herman	NL	1930	614	.393	.091	.432
Ty Cobb	AL	1913	428	.390	.072	.421
Al Simmons	AL	1931	513	.390	.088	.427
George Brett	AL	1980	449	.390	.049	.410

Table 4. Comparison of Williams, Brett, and Gwynn

		INDIVIDUAL			LEAGUE			DIFFERENCE		
YEAR	PLAYER	AVG	KAVG	IPAVG	AVG	KAVG	IPAVG	AVG	KAVG	IPAVG
1941	Ted Williams	.406	.059	.431	.266	.102	.297	52%	-42%	45%
1957	Ted Williams	.388	.102	.432	.255	.138	.296	52%	-26%	46%
1980	George Brett	.390	.049	.410	.269	.133	.310	45%	-63%	32%
1994	Tony Gwynn	.394	.045	.413	.270	.184	.327	46%	-75%	26%

Table 4). In 1957 his IPAVG was even higher than 1941, but he struck out almost twice as often as 1941. George Brett and Tony Gwynn had incredibly low KAVGs—less than .050, or about 70% lower than the league average (much better than Ted Williams). However, their IPAVGs were too low to break the .400 barrier.

AN EXPLOSION IN OFFENSE

IPAVG has increased dramatically since 1993. As we know, this increase has been driven by a major increase in home runs. Figure 2 makes this abundantly clear.

Several interesting trends appear in Figure 2. First, the fraction of balls put in play that have become singles (1BIP) has remained quite stable since 1946. Next, the likelihood of a triple (3BIP) has decreased substantially since 1876. This must be driven in large part to changes in field dimensions. And, most strikingly, the likelihood that a ball put in play becomes a home run (HRIP) increased dramatically since the middle of the century and has exploded since 1993.

Many of these trends become apparent by looking at averages over several decades, as shown in Table 5. Table 6 presents the change in the average to 2002.

The chance of a single, given the ball has been put in play, has only increased by 1% from the 1876-1886 average (.217 vs .219). This should not be surprising given that the factors that determine whether a ball put in play becomes a single (field dimensions, location of fielders, speed of batted ball) have not changed appreciably, if at all, since 1876. On the other hand, the likelihood of doubles, triples, home runs, and all extra base hits (Extra H) has changed dramatically over the years. Likewise, IPSLG has seen significant increases.

Table 5. Hit type IPAVG and IPSLG by decade (NL and AL)

DECADE	1B	2B	3B	HR	EXTRA H	IPSLG
1876-1885	0.217	0.043	0.014	0.004	0.061	0.363
1886 1895	0.221	0.042	0.018	0.009	0.070	0.397
1896-1905	0.233	0.040	0.017	0.006	0.063	0.387
1906-1915	0.224	0.040	0.016	0.005	0.061	0.371
1916-1925	0.228	0.047	0.016	0.009	0.072	0.405
1926-1935	0.223	0.056	0.014	0.015	0.085	0.438
1936-1945	0.218	0.051	0.011	0.016	0.078	0.419
1946-1955	0.213	0.047	0.010	0.024	0.082	0.434
1956-1965	0.214	0.047	0.009	0.031	0.087	0.459
1966-1975	0.218	0.046	0.008	0.026	0.080	0.439
1976-1985	0.219	0.051	0.008	0.026	0.086	0.451
1986-1995	0.219	0.057	0.007	0.031	0.095	0.477
1996-2002	0.219	0.064	0.007	0.039	0.111	0.525

Table 6. Change in IPAVG and IPSLG since 1876 (NL and AL)

YEARS	1B	2B	3B	HR	EXTRA H	IPSLG
1876-2002	1%	50%	-52%	824%	80%	45%
1886-2002	-1%	53%	-62%	320%	59%	32%
1896-2002	-6%	60%	-59%	570%	76%	36%
1906-2002	-2%	61%	-57%	678%	81%	41%
1916-2002	-4%	38%	-56%	323%	54%	30%
1926-2002	-2%	16%	-51%	156%	30%	20%
1936-2002	1%	27%	-39%	139%	41%	26%
1946-2002	3%	36%	-33%	63%	35%	21%
1956-2002	2%	37%	-24%	27%	27%	14%
1966-2002	0%	41%	-13%	49%	38%	20%
1976-2002	0%	26%	-19%	51%	29%	16%
1986-2002	0%	14%	-6%	27%	17%	10%

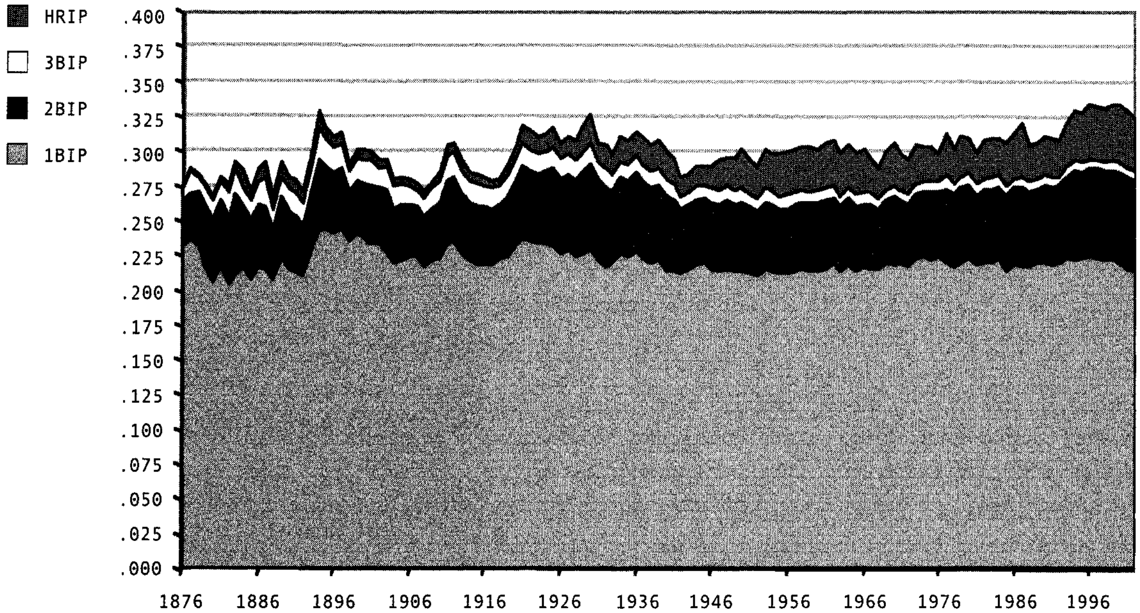


Figure 2. *Hit type KAVG, and IPAVG by decade (NL and AL)*

PLACING BASE HITS

Do batters have the ability to place base hits? This is a question worthy of more space than we can provide here. However, a quick look suggests that, yes, the ability to place base hits is a skill. Table 7 presents the highest and lowest IPAVGs (in total and by hit type) and KAVG for the 2002 season (only players eligible for a batting title were considered).

Jose Hernandez led the majors with a .448 IPAVG (1.33 times the league average), while Neifi Perez's IPAVG was only .261 (0.77 times the league average). Even if we exclude home runs, and thereby do not include power hitters, we find a significant difference between player performance—Jose Hernandez's .377 versus Jay Gibbons' .219. Not surprisingly, the difference between the best and the worst is even greater for other hit types.

We also see a significant difference in terms of KAVG. Jose Hernandez led the majors with a .358 KAVG in 2002. That's right, the same Jose Hernandez that led the league in IPAVG! 35.8% of Jose's at-bats resulted in strikeouts, but if he put the ball in play he got a hit 44.8% of the time. That is the sixth highest IPAVG since 1913 (see Table 2). Jason Kendall had the lowest KAVG with only 5.3% (.053) of his ABs ending in strikeouts.

Table 7. Highest and lowest IPAVGs and KAVG in 2002

STAT	PLAYER	INDIVIDUAL	RELATIVE TO MLB
IPAVG	Jose Hernandez	0.448	1.33
	Neifi Perez	0.261	0.77
IPAVG w/o HRs	Jose Hernandez	0.377	1.30
	Jay Gibbons	0.219	0.75
1BIP	Luis Castillo	0.302	1.40
	Jay Gibbons	0.149	0.69
2BIP	Bobby Abreu	0.110	1.61
	Luis Castillo	0.034	0.50
3BIP	Brad Wilkerson	0.023	3.50
	(20 Players)	0.000	0.00
HRIP	Jim Thome	0.152	3.23
	Fernando Vina	0.002	0.04
KAVG	Jose Hernandez	0.358	2.05
	Jason Kendall	0.053	0.30

As you can see, batters appear to differ to a significant degree in their ability to get a hit once they put the ball in play and their ability to avoid strikeouts. However, the spread among players is greater for KAVG than it is for IPAVG. That is, players differ to a greater degree in their ability to avoid striking out than in their ability to get a hit once they put the ball in play.

Further research should focus on whether or not a player's performance is consistent over time.

CONCLUSION

Batting average measures the batter's ability to avoid striking out and his ability to "hit 'em where they ain't." By rewriting AVG as a function of these two effects, we gain greater insight into what has driven changes in AVG over the last hundred years, why hitting .400 is so difficult, and whether batters can really place hits.

ACKNOWLEDGMENTS

The author thanks Jim Charlton and three anonymous reviewers for many helpful suggestions.

NOTES

1. Dean Stotz and I referred to this as IPA in our 2002 BRJ article (Bickel, J. Eric, and Dean Stotz, "Batting Average by Count and Pitch Type: Fact and Fallacy," *The Baseball Research Journal*, No. 31, 2003, pp. 29-34). In that same issue, Dick Cramer referred to this concept as "batting average per batted ball" (Cramer, Dick, "Preventing Base Hits: Evidence That Fielders Are More Important Than Pitchers," *The Baseball Research Journal*, No. 31, 2003, pp. 88-92). See also, Birnbaum, Phil, "Factors Affecting Pitcher Ball-in-Play Average," *By the Numbers*, May 2001, pp. 8-11.
2. John Burnson developed equation 1 independently in *Ron Shandler's 2003 Baseball Forecaster* (pp. 7-8). Stanford Baseball has been calculating IPAVG since 1998, and this stat has been publicly available since 2000.
3. In addition, Jim Albert and Jay Bennett analyze AVG by count in their book *Curve Ball* (see pp. 87-107). They note, "Players who have small pitch count effects are relatively unlikely to strike out." This is expected because AVG for batters that do not strike out a lot is almost equal to IPAVG, which is not as dependent on the count as KAVG.
4. Bickel, J. Eric and Dean Stotz, *op cit.*
5. This data was obtained from the "The Lahman Baseball Database" (www.baseball1.com). Individual player data on strikeouts is unavailable from 1897 to 1909 in the NL and 1901 to 1912 in the AL. However, league total strikeout data was obtained from *Total Baseball.*
6. Steinberg, Steve L.,. "The Spitball and the End of the Deadball Era," *The National Pastime*, No. 23, 2003, pp. 7-17.
7. Bickel, J. Eric, and Dean Stotz, *op cit.*
8. These rule changes were obtained from David Nemec's outstanding book *The Rules of Baseball* (1994, Lyons and Burford Publishers) and *Total Baseball.*
9. Only player seasons for which the batter had enough plate appearances to be considered for a batting championship were included.

ROBERT K. ADAIR

Cameras and Computers, or Umpires?

According to baseball rules: "A STRIKE is a legal pitch when so called by the umpire, which is not struck at, if any part of the ball pass though any part of the strike zone.

"The STRIKE ZONE is that area over home plate the upper limit of which is a horizontal line at the midpoint between the top of the shoulders and the top of the uniform pants and the lover level is a line at the hollow beneath the kneecap. The strike zone should be determined from the batter's stance as the batter is prepared to swing at a pitched ball."

Simply pictured, any pitched ball that would touch an imaginary vertical extension of the five-sided home plate extending above a horizontal plane at the knees and terminating below a horizontal plane at the letters, as defined by the rule book, is a strike.

For more than a century an umpire (human by statute if not in the view of some fans) has called balls and strikes. With modern technology, fast video cameras and fast computers analyzing the video images, that determination can be made instrumentally. Such systems have been used in televised game broadcasts for some time. In 2002, Major League Baseball arranged for the operation of the QuesTec™ system in the Anaheim, Arizona, Boston, Cleveland, Houston, New York (Shea), Milwaukee, and Tampa Bay ballparks to evaluate umpires in a program labeled the Umpire Information System (UIS).

All such systems work in approximately the same manner—tracking the flight of basballs as they near the plate. This reporter has been given access to detailed technical data on one such system as well as rather less information on QuesTec™, which is closely held. Also, my extensive experience in the computer analysis of photographs of tracks of elementary particles moving through liquid hydrogen has been surprisingly useful in giving me insights into baseball-track reconstruction problems.

The sketch of Figure 1 suggests the positions of the cameras in a typical UIS system. The primary video cameras, *C-L* and *C-R*, are mounted high in the grandstand overlooking first and third base. Auxiliary cameras (labeled *c* in the diagram) are placed low, often near the dugouts, and are used to set the high and low limits of the strike zone at the letters and knees of each batter. Constrained by the ball park architecture, the cameras are placed somewhat differently in each park.

Using the limited information I have about QuesTec™, and more extensive information about a similar system, I can reconstruct the UIS measuring process in a manner that cannot be far from the mark. Each of the two primary cameras takes a "picture" of the ball every 1/30th of a second as it nears the plate. The resultant images reflect the position of the ball at intervals of the order of four feet over about the last 30 feet of the ball's flight until it nears the plate. Blocked by the batter, the ball is not actually tracked over the plate but the ball position at the plate is calculated extrapolating from the measurements of its approach trajectory.

The image planes of the cameras used for baseball tracking are typically made up of three sets (for brightness, red, and blue) of CCDs (charge-coupled devices) with 758 (vertical) x 494 (horizontal) pixels. The image of the 2.9" diameter ball is then roughly 8 pixels vertically and perhaps 10 horizontally as the ball image is elongated about an inch by the motion of the ball during the 1/1000th of a second "exposure." There is no physical shutter; the CCDs are read electronically. The information from the image planes of the two cameras is then processed by the computer to construct the position of the ball as it moves toward home plate. The position of the center of the ball at each picture time can be reconstructed from the images captured by the two cameras to an accuracy of about " horizontally

ROBERT ADAIR *is Sterling Professor Emeritus of Physics at Yale University and a member of the National Academy of Sciences. His research has largely been connected with the properties of the elementary particles and forces of the universe. At the request of his friend Bart Giamatti, he served as official Physicist to the National League 1987-1989.*

and $\frac{3}{16}$" vertically. In this reconstruction, the computer corrects for some classes of distortions, such as barrel or pin-cushion lens distortions.

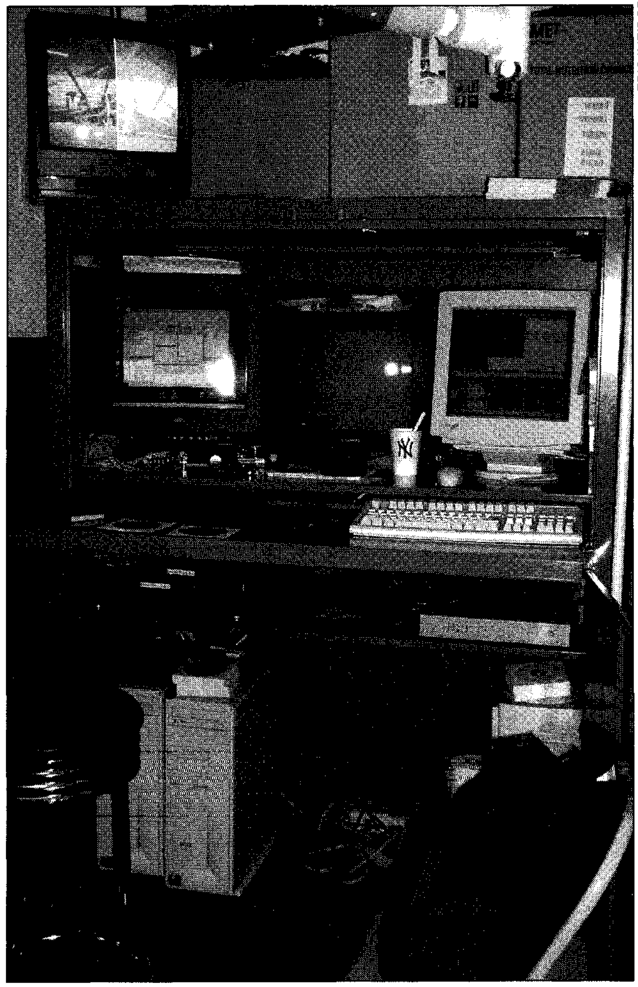
QuesTec™ has made some measurements of the accuracy of their system that—according to their reports—appear to be quite well-done, if somewhat limited. After a preliminary set of experimental measurements at Tampa, they tested their system at Fenway, where they measured the differences between the calculated and measured impact points of 44 pitches that hit a backboard at home plate. Measuring only the random variations, they found a "root-mean-square" (average) left-right error of 0.40 inches and an up-down error of 0.47 inches.

This assessment of error did not include some systematic errors such as errors in the calibration of the system during a game, which can be significant. The extensive calibration set up for the Fenway test was quite special and not used later in real game situations. Nor was the data sample sufficiently large to evaluate the likelihood of rare large errors. If the "random" errors were made up of very many independent small errors, the probability of large errors is well-defined; only about one pitch in 80 would have a left-right error from random fluctuations greater than an inch, and one in 25 an up-down error greater than an inch. However, in most real measurements, the probability of large deviations—the "tails" of the distribution—are greater than expected from a single normal distribution. Nevertheless, I would expect that the probability of a random error greater than two inches would be negligible.

However, most complex systems make occasional "mistakes" as well as errors, and mistakes are not easily tractable. My own experience of a half century of experimental physics tells me that while it is no inconsiderable effort to construct a system such as QuesTec™ which operates correctly 98% of the time, it is often the very devil to get to 100%.

During the year 2002, UIS was used to check the ball-strike calls of plate umpires.

The system requires an operator who uses two computers and several monitors. Before the game, he calibrates the system—that is, locates home plate—through a defined procedure using one of the monitors. He may repeat the calibration during the game. As each pitch is made, the operator records the disposition of the pitch



ERIC IMEIDOPF

The Umpire Information System at Yankee Stadium.

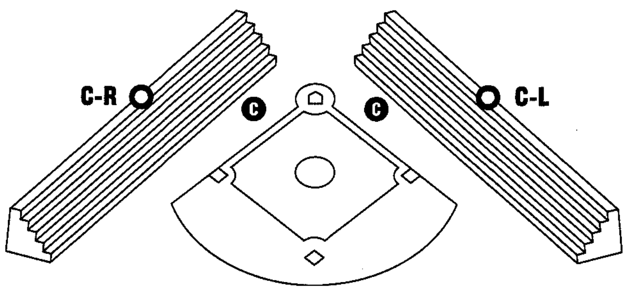


Figure 1. Typical UIS camera locations.

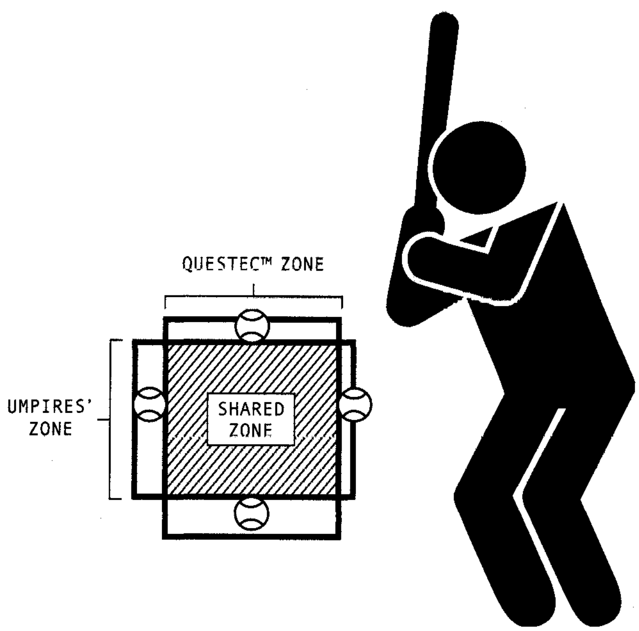


Figure 2. Comparison of the different strike zones of umpires and the Questec™ system.

on the scoring computer, pressing buttons that signify "swung at," or "called by the umpire." Warm-up throws, etc., are automatically recorded by the system but elicit no input label. There is also a button that signifies "bad track," presumably pressed when the system fails transparently. The path of each pitch, as determined by the tracking cameras, is recorded by the system and displayed on a monitor. A second monitor shows the game from the center-field camera.

After the game the operator views the pictures taken, using the ground-level cameras, of the batters at the plate in one of the monitors and sets the high and low strike levels at the knees and letters of each player as the player stands waiting for a pitch.

With the upper and lower limits of the strike zone defined for each batter, the operator sets up the computing program that calls balls and strikes for each pitch and reviews each pitch. At this time he identifies "bad tracks," where the computed track is clearly not in accord with other information. Then a CD is produced for the plate umpire that shows the computed ball tracks and the system and umpire ball-strike calls for each pitch with a letter *C* (correct) if the umpire and UIS give the same call, *A* (acceptable) if the calls disagree but a change of two inches in the computed trajectory would bring the calls in agreement, and *N* (not acceptable) if the calls disagree by greater than two inches. The operator manually excludes the information on the "bad track" pitches that he identifies, which are usually, if not always, labeled *N* by the system.

This system was used to evaluate umpires in about 600 games in 2002. A total of 83,891 pitches were recorded where both the umpire and UIS called balls and strikes. Man and machine agreed (*C*) on 71,164. They disagreed (*A*) on 4,970 where the trajectory difference was less than two inches, a difference that was considered within the uncertainties in the system, and hence the umpires' calls were acceptable. However, they strongly disagreed (*N*) on the call—by more than two inches in the trajectory—in 7,757 pitches. Thus man and machine differed strongly on about 9% of the called pitches, for an average of about 14 pitches per game.

There were differences in the scores of the 79 different umpires. I made a statistical analysis of those scores and found that the differences between umpires were almost wholly due to chance fluctuation in the scoring process and not from whatever differences there might be in the umpires' ball-strike judgments. Although the umpires believe that they had been told in the Major League Baseball Umpire Manual that the 2002 results from the UIS system were to be used only for "training to improve performance" and "no umpire will be judged . . . for the 2002 season," they believe that their scores according to UIS were considered in the appointment of umpires for the post-season series.

The errant calls are classified in the table where U-s, Q-b means that the umpire called the pitch a strike and the computer called it a ball. Conversely, U-b, Q-s stands for pitches that the umpire called balls and the computer strikes.

OUTSIDE CORNER		HIGH	
U-s, Q-b	3.336	U-s, Q-b	301
U-b, Q-s	122	U-b, Q-s	943

INSIDE CORNER		LOW	
U-s, Q-b	622	U-s, Q-b	18
U-b, Q-s	208	U-b, Q-s	2007

It is evident that the differences were not random, and hence that most of the N (not acceptable) umpire calls were not simply due to erratic judgments. As the most extreme discrepancy, the computer called 2,007 low pitches strikes that the umpires called balls, while the umpires called only 18 low pitches strikes that the computer judged were balls. Clearly the computer strike zone extended far below the umpires' zone.

With the aid of a plausible model, I have estimated the difference between the strike zones as shown in the cartoon of Figure 2. Independent of which is "right" or even if there is any precise meaning to "right," UIS and the umpires call different strike zones. The computer zone is much narrower than the umpire zone but longer in the vertical direction. That difference between zones accounts for about half of the 9% of pitches where man and machine differ. If the umpires would adjust by calling balls much tighter on the outside corner, a little tighter on the inside corner, allowing a slightly higher strike, and calling almost any low ball that isn't in the dirt a strike, their UIS error score would decrease by about 4.5%.

Moreover, a substantial fraction of those errors would probably be from the QuesTec™ system—operator plus machine. While the rate seems to vary widely game-to-game and park-to-park, the operators seem to have thrown out an average of five pitches per game (or about 3% of the pitches) as "bad tracks," QuesTec™ system mistakes. How many of the 4.5% residual errors were then actually less dramatic QuesTec™ mistakes rather than umpire errors? Perhaps almost all! Indeed, the umpires contend that the UIS system miscalls occasional pitches by more than a foot—and sometimes much more.

Though the umpires could adjust to the UIS strike zone, such an umpire adjustment would change the game of baseball. The pitch on the outside corner now called a strike, that many pitchers (e.g., Greg Maddux) live by, would be gone. Conversely those pitchers that like to throw high and hard together with a splitter low and slow (e.g., Hideo Nomo) might do very well.

Baseball, in the form we know, is a game with an important tradition that goes back more than a century. My grandfather Ted Wiegman, who played sandlot baseball at third base in Fort Wayne, Indiana, of the 1890s, if alive would have no trouble identifying the modern game today with the game he loved as a youth. If somehow umpires (all umpires!) have in recent years drifted away from the traditional game and dramatically changed the strike zone, surely it would be desirable to bring it back to the letter of the rule book, and the UIS could assist in doing that. Indeed, with only a little more development we can envision balls and strikes called only by machine with the umpires relegated to judging base-running plays and such.

However, if the present umpires' strike zone is really the traditional zone, called that way long ago by umpires like Jocko Conlon, Bill Klem, and Cal Hubbard, we should be wary about changing the zone and the game. I have made some computer simulations that suggest to me that the umpires' zone is largely the natural human interpretation of the rule book zone and, most likely, that zone has always been as it is called now. In the future, highly developed machines could surely call balls and strikes more accurately than human umpires (and pitching machines can throw harder than Randy Johnson!), but for the time being let's keep humans judging and playing this human game.

TERRY BAHILL AND DAVID G. BALDWIN

The Vertical Illusions of Batters

For many decades, batters have maintained that a high, hard fastball can accelerate and rise suddenly as it nears the strike zone. They claim the ball appears to "jump" a foot or more with an explosive burst of speed. For example, in describing Dwight Gooden, Tony Gwynn stated, "He rears back and throws you that high-rising fastball whenever he needs a big pitch" (Gwynn and Rosenthal). The rising fastball is often called "smoke," "cheese," or "express."

According to principles of physics, a rising fastball (in the strike zone) is impossible unless delivered with a low side-armed or underhand release (Karnavas, Bahill and Regan; Bahill and Karnavas). Regardless of the delivery, all pitches must decelerate and all are in a free fall during flight. Yet most batters claim the explosive riser can be thrown by pitchers with an overhand or three-quarter delivery. This widespread belief has persisted in spite of the recognition by some coaches that such a pitch cannot occur (House; Thrift and Shapiro; Mike Scioscia quoted by Will).

A related phenomenon is the curve or sinker that appears to "drop off the table." This is described as a sharp, downward break and sudden deceleration as the pitch nears the strike zone. Like the rising fastball, the breaking pitch contradicts the laws of physics—the trajectory of every spinning pitch is smooth with no sudden changes. As these phenomena must be perceptual illusions, what is causing them?

WHAT THE BATTER DOES DURING THE PITCH

The timeline of the batter's activities during the pitch has been described by Watts and Bahill (2000) and Adair (2002). During the early phase of the ball's flight, the batter gathers sensory information about the pitcher's release point and the trajectory of the ball (Baker, Mercer, and Bittinger,; Schmidt and Ellis). Concurrently, the batter compares this information with mental models of previous pitches he or she has experienced (Bahill and Karnavas; Gray). A model with characteristics that appear to be consistent with the current pitch will be used in predicting the location and time of contact.

After the observational phase, the batter computes where and when the ball will make contact with the bat. In addition, the swing/take decision is made in this second phase. Once this decision processing is complete, the batter begins the swing of the bat (and decision making occurs only with respect to checking the swing). Upon completion of the pitch, the batter gathers final information and compares it with the predictions.

In this scenario, the batter uses information gathered in the first part of the pitch to estimate time until contact and the height of the ball at contact. The batter must determine the pitch speed in order to estimate the height of the ball (the ball is in a free fall throughout its flight—the longer the pitch is in the air, the farther it falls). Underestimation of pitch speed results in anticipation that the potential contact point is lower than it actually is and can produce the illusion of the rising fastball. An analogous explanation of the breaking pitch is based on speed overestimation (Karnavas, Bahill, and Regan, 1990; Bahill and Karnavas, 1993).

EYE-TRACKING STRATEGIES

The angle of the ball's trajectory with the batter's line of sight increases during the flight of the ball. Consequently, even on pitches of moderate speed, the image velocity increases until the batter can no longer maintain the image on the fovea of the eye. Thus, the batter cannot use smooth-pursuit eye movement to track the pitch the entire distance to the contact point.

To compensate, batters use one of two strategies

DAVE BALDWIN *pitched for the Senators, Brewers, and White Sox during the 1960s and '70s, and earned a Ph.D. in genetics and an M.S. in systems engineering. He has been a scientist, engineer, artist, and poet.* **TERRY BAHILL** *received his Ph.D. from the University of California, Berkeley, in 1975, and has been a Professor of Systems Engineering at the University of Arizona since 1984. He is the co-author of* Keep Your Eye on the Ball: Curve Balls, Knuckleballs and Fallacies of Baseball.

in tracking the pitch (Bahill and LaRitz). *The optimal learning strategy* allows the batter to see the ball hit the bat. The batter tracks the ball over the early part of its trajectory with smooth-pursuit eye movement, then makes a saccade (leap) to a predicted point of bat-ball collision. The batter continues to follow the ball with peripheral vision, letting the ball catch up to the foveae. The batter then resumes smooth-pursuit tracking with the image of the ball on the foveae. It is called an optimal learning strategy because the batter predicts where the contact point will be, and sees the ball's position when it contacts (or fails to contact) the bat. The batter uses this feedback to make better predictions when the pitcher throws a similar pitch.

The optimal hitting strategy does not allow the batter to see the ball hit the bat. With this strategy, the batter tracks the ball with smooth-pursuit eye movements and falls behind late in the pitch. It is called the optimal hitting strategy because the batter keeps the eyes on the ball longer. This should allow late adjustments to the swing/take decision. We have no evidence that batters voluntarily switch between these two strategies.

Either strategy allows the batter to track the ball with smooth pursuit long enough to start the swing. Then, the momentum of the bat does not allow the batter to alter the timing or height of the swing—the only change the batter can make during the swing is to check it by pulling the bat toward the rear shoulder. Therefore, neither tracking strategy allows an adjustment for greater accuracy during the swing.

With respect to monitoring the actual height of the pitch when it reaches the potential contact point, the optimal hitting strategy might be inferior because the batter cannot see the ball at this point but must infer its position by where the ball is caught—several feet behind the contact point. Since experienced catchers catch the ball with glove moving toward the center of the strike zone (to influence the umpire), the batter could be misled in judging the height of the pitch.

BATTERS' PREDICTIONS AND THE RISING FASTBALL

Although retinal image information provides an accurate cue for the time until contact, it provides poor cues for absolute distance to the ball and for its line-of-sight speed (Bahill, Karnavas, and Regan, 1990). Classical stereoscopic depth perception is of little help in this regard. Although this system provides a precise indication of relative depth (i.e., the difference between the x-axis distances of two objects imaged near the fovea), it provides little indication of distance. In tracking the pitched ball, the batter has one object, the ball, imaged on the fovea. Therefore, the batter cannot measure the distance to the ball or the pitch speed; the batter can only estimate them.

Bahill and Karnavas present the following psychophysical explanation for the rising fastball. The batter can only approximate pitch speed and the time since release of the pitch. The batter uses these data in conjunction with his or her experience (mental models of past pitches) to estimate the distance to the ball. The batter then uses this estimate and the ball's retinal image velocity to estimate the vertical velocity. From the vertical velocity and the time until contact, the batter can estimate how far the ball will fall in the last one-third of its flight, thereby predicting the height of the ball at the potential contact point.

Figure 1a and Table 1 illustrate the results of simulation studies of 95 and 90 miles per hour (mph) fastballs (Karnavas, Bahill, and Regan). These simulations include the effects of gravity and aerodynamic forces of lift and drag. In these studies, both pitches were launched one degree upward with 1500 revolutions per minute (rpm) of backspin. As shown in Figure 1d, the distance from the front of the pitcher's rubber to the plate's back vertex is 60.5 feet. The pitcher released the ball about five feet in front of the rubber. Therefore, the simulated release point was 55.5 ft. away from the vertex. We assume the batter's head was aligned with the front of the plate and the bat hit the ball about 1.5 ft. forward of the head. The plate measures 17 inches from back vertex to front edge. Thus, the bat-ball collision point was assumed 3 ft. in front of the vertex (represented by bottom row values of Tables 1 and 2). The pitcher's release point was assumed six feet high.

Now consider an example of a visual judgment error. Suppose the pitcher throws a series of 90 mph pitches, followed by a 95 mph fastball. Assume the batter uses a 90 mph mental model to interpret retinal image information about the 95 mph pitch. Suppose the batter tries to estimate the ball's vertical speed 200 msec after the ball left the pitcher's hand. The actual pitch (a 95 mph fastball) would be 28.6 ft. from the vertex of the plate (Table 1). By subtracting 1.5 ft. (the

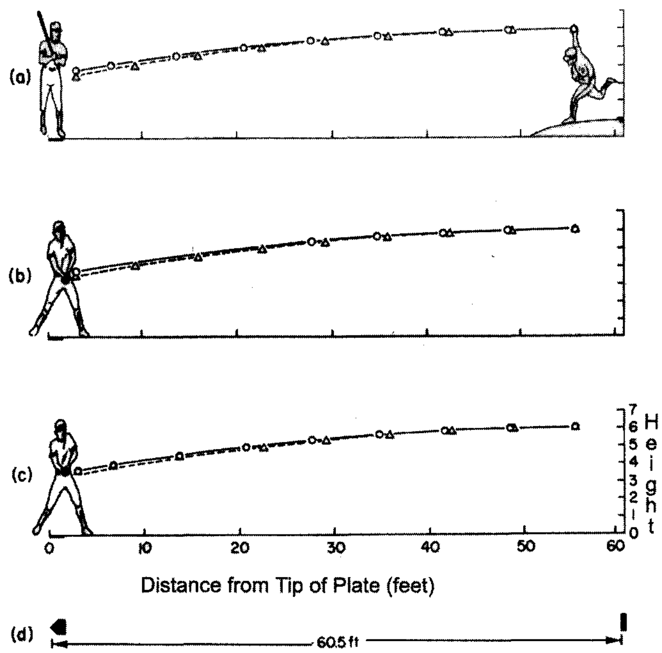


Figure 1. *(a) Computer simulation of the trajectory of a 95 mph fastball (solid line and circles) and a 90 mph fastball (dashed line and triangles). The slower pitch takes longer to get to the plate and therefore drops more. (b) Computer simulation of the trajectory of a 95 mph fastball (solid line and circles) and the batter's mental model of this trajectory (dashed line and triangles) when the batter underestimated the speed of the pitch by 5 mph. (c) The same simulation as Figure 1b, except that when the ball was 20 feet in front of the plate, the "batter" realized his mental model was wrong and corrected it, thus putting his mental model triangles on the 95 mph trajectory. (d) Physical dimensions for adult baseball. [From Karnavas, Bahill and Regan, 1990.]*

distance forward from vertex to the batter's eye alignment) from 28.6, we get 27.1 ft. as the distance from the ball to the batter's eyes. At this distance, its vertical velocity of 6.4 ft/sec (derived from gravitational effects) would produce a retinal velocity of 13.3 deg/sec. The actual height of the ball at the potential impact point is 3.56 ft. (Table 1).

Table 1. Trajectories of fastballs

TIME SINCE	95 MPH FASTBALL			90 MPH FASTBALL		
RELEASE (MSEC)	DIST. (ft.)	HGT. (ft.)	MPH	DIST. (ft.)	HGT. (ft.)	MPH
0	55.5	6.00	95.0	55.5	6.00	90.0
50	48.6	5.86	93.3	49.0	5.86	88.5
100	41.8	5.67	91.7	42.5	5.68	87.0
150	35.2	5.43	90.2	36.2	5.44	85.6
200	28.6	5.15	88.6	30.0	5.16	84.3
250	22.2	4.83	87.2	23.8	4.84	82.9
300	15.8	4.46	85.7	17.8	4.47	81.6
350	9.6	4.05	84.4	11.9	4.05	80.4
400	3.4	3.59	83.1	6.0	3.59	79.2
450	3.0	3.56	83.0	–	–	–
426	–	–	–	3.0	3.33	78.6

However, if the batter thinks the pitch is a 90 mph fastball, this model would translate to a pitch 28.5 feet away (30.0 – 1.5) at 200 milliseconds (msec) after release. At this distance, a retinal image velocity of 13.3 deg/sec would indicate the vertical velocity was about 6.7 ft/sec. The batter would think the ball was falling farther than it really was. The batter would predict the height at impact to be 3.33 ft. and would tend to swing under the ball. Therefore, if the batter made a saccadic jump to the predicted point of contact, this point would be below the ball when the ball caught up with the eye, and the ball would seem to jump upward—in this example by three inches. This error of visual judgment could be avoided if the batter had an accurate visual cue to the ball's absolute distance or its speed, but the batter has no direct optical sense for these two important parameters.

BATTERS' PREDICTIONS AND THE BREAKING PITCH

Spinning baseballs follow smooth parabolic trajectories. The 90 mph fastball of Table 1 and Figure 1 falls more than 2.5 ft. in its flight to the plate. A plot of this vertical distance as a function of time would be parabolic. In the first, second, third, and fourth 100 msec periods, the ball falls 3.8, 6.2, 8.3, and 10.3 inches, respectively. The ball drops progressively more in each period, but it follows a smooth parabolic trajectory (in agreement with Adair, 2002). The drop can be enhanced with the addition of a vertical Magnus force due to topspin on the ball. Table 2 shows the results of simulations of 80 and 75 mph drop curves (defined as pitches with pure topspin). Both were launched upward at an angle of 2.5 degrees with 1900 rpm. We used a formula from Watts and Bahill (2000) to calculate the downward force due to spin.

Table 2. Trajectories of drop curves

TIME SINCE RELEASE (MSEC)	80 MPH		75 MPH	
	DIST. (ft.)	HGT. (ft.)	DIST. (ft.)	HGT. (ft.)
0	55.5	6.00	55.5	6.00
50	49.7	6.20	50.0	6.18
100	44.0	6.28	44.7	6.25
150	38.3	6.24	39.4	6.20
200	32.7	6.09	34.1	6.05
250	27.2	5.83	29.0	5.78
300	21.9	5.47	23.9	5.41
350	16.5	4.99	18.8	4.93
400	11.3	4.41	13.9	4.35
450	6.11	3.73	9.0	3.67
480	3.0	3.27	–	–
500	–	–	4.2	2.89
513	–	–	3.0	2.68

Consider the 75 mph pitch. The ball falls 2.4 inches between 100 and 200 milliseconds, and 7.7, 12.7, and 17.5 inches in the following 100 millisecond periods. Once again, the ball drops more in each period, but it still follows a smooth parabolic trajectory. This is a gradual curve rather than a sharp break. To undergo a sharp downward break a 75 mph curve would have to drop 2.4 inches in the early 100 millisecond period, but more than 17.5 in the last 100 milliseconds.

To explain the breaking pitch, we will suppose the pitcher threw the 75 mph drop curve of Table 2. It would drop 25 inches in the last 150 milliseconds before contact. However, if the batter overestimated the pitch speed and thought it was 80 mph, the batter would expect it to fall 21 inches in the last 150 milliseconds. Thus, if the batter took his eye off the ball 150 milliseconds before the projected time of contact, and saw it again when it arrived at the potential contact point, he would think that it broke downward 4 inches. Therefore, we suggest that the apparent break of some pitches might result from the overestimation of pitch speed in the batter's mental model—the opposite of the explanation for the rising fastball.

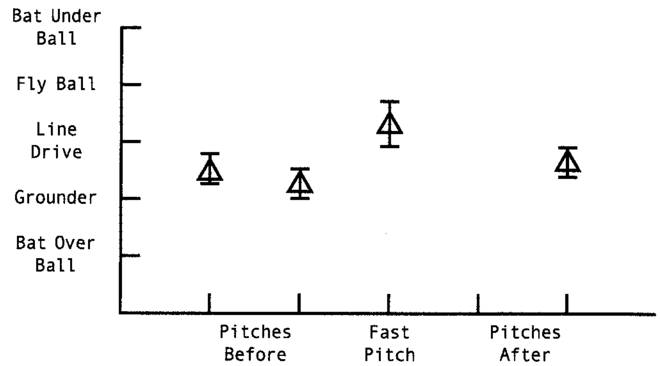


Figure 2. Averaged data from seven batters, showing that when an unusually fast pitch was thrown, most batters swung under the ball. The triangles are the mean values, and the vertical bars are the 95% confidence intervals.

EXPERIMENTAL ASSESSMENT OF THE MODEL

To assess the model, Bahill and Karnavas ran experiments using a pitching machine. They threw 450 pitches to seven subjects: 3 adults and 4 boys aged 9, 11, 11 and 13. Pitching speed was set at 50 mph, with an occasional 55 mph pitch. The number of 50 mph pitches between these fast pitches was 3, 4, 5, or 6, chosen randomly. An observer (who did not know the pattern of pitches) recorded the relationship of bat and ball when the ball crossed the plate. The outcomes of the fast pitches and the two pitches before and two after were averaged, as shown in Figure 2. These results, statistically significant, show that on fast pitches batters swung below the ball, indicating they underestimated the speed of the pitch. This vertical error in judgment associated with error in speed estimation supports the illusion model.

PITCHERS' TACTICS

The pitcher's tactics are to select a pitch different from the batter's predictive model and to provide adequate vertical movement of the pitch (since vertical movement is more effective than horizontal). The most important factor in pitch selection is the change of speeds. Warren Spahn is quoted by Will as saying, "Hitting is timing. Pitching is upsetting timing." As a manager, Ted Williams exhorted his pitchers to never let a batter see consecutive pitches of the same speed. Pitching coach Johnny Sain expressed the same maxim, adding that the pitcher should work rapidly to allow the batter to retain a mental image of the previous pitch. In following this advice, pitchers amplify speed differences, thereby increasing the likelihood the batter will misjudge the vertical movement of the ball (Bahill and Baldwin). These tactics contribute to vertical illusions of batters.

CONCLUSIONS

Vertical illusions are caused by misjudgment of the ball's distance and speed coupled with accurate but misunderstood feedback about the prediction error. That is, the batter predicts the height of the potential contact point inaccurately, sees how far off the prediction was, and then misinterprets this error to be a phenomenon of the ball's flight. The likelihood the batter will undergo an illusion is increased by pitching tactics designed to confound the batter's judgment of pitch speed and pitch height at contact. The illusion is not directly related to the effectiveness of the pitch, however; commitment to the swing takes place before the batter experiences an illusion.

REFERENCES

Adair, R. K., 2002, *The Physics of Baseball*, 3rd ed., New York: HarperCollins.

Bahill, A. T. and Baldwin, D. G., 2003, "The Rising Fastball and Other Perceptual Illusions of Batters," *Biomedical Engineering Principles in Sports,* George Hung (ed.), New York: Kluwer Academic/Plenum Publishers.

Bahill, A. T. and Karnavas, W. J., 1993, "The Perceptual Illusion of Baseball's Rising Fastball and Breaking Curveball," *Journal of Experimental Psychology, Human Perception and Performance,* 19(1): 3-14.

Bahill, A. T. and LaRitz, T., 1984, "Why Can't Batters Keep Their Eyes On the Ball?," *American Scientist,* 72: 249-253.

Baker, D., Mercer, J., and Bittinger, M.L., 1993, *You Can Teach Hitting,* Carmel, IN: Bittinger Books.

Gray, R., 2003, "Markov at the Bat: A Model of Cognitive Processing in Baseball Batters," *Psychological Science,* 13(6): 542-547.

Gwynn, T. and Rosenthal, J., 1992, *Tony Gwynn's Total Baseball Player,* New York: St. Martin's Press.

House, T., 1994, *The Pitching Edge,* Champaign, IL: Human Kinetics.

Karnavas, W. J., Bahill, A. T., and Regan, D. M., 1990, "A Model for the Rising Fastball and Breaking Curve Ball," *Proceedings of the 1990 IEEE International Conference on System, Man, and Cybernetics,* 925-928.

Schmidt, M. and Ellis, R., 1994, *The Mike Schmidt Study: Hitting Theory, Skills and Technique,* Atlanta: McGriff and Bell.

Thrift, S. and Shapiro, B., 1990, *The Game According to Syd,* New York: Simon and Schuster.

Watts, R. G., and Bahill, A. T., 2000, *Keep Your Eye on the Ball: Curveballs, Knuckleballs and Fallacies of Baseball,* New York: W.H. Freeman and Co.

Will, G. F., 1990, *Men at Work*, New York: Macmillan.

JOE D'ANIELLO

DiMaggio's Hitting Streak

High "Hit Average" the Key

If it weren't for eBay, I never would have written this article. While browsing around the mega auction site I discovered a 1995 *BRJ* for sale. I entered a last-minute bid (yes, I'm one of *those* guys) and won the auction with a $5 bid. One article in the 1995 *BRJ* was titled "Streaks" by Neal Moran. Moran discussed streaks of all kinds and referred to a 1994 *BRJ* article by Charles Blahous titled "The DiMaggio Streak: How Big a Deal Was It?" Blahous estimated that DiMaggio had a .013% chance of hitting in 56 games or about a 1 in 750 chance. Last year Michael Freiman ("56 Game Hitting Streaks Revisited," 2002 *BRJ*) resurrected the topic and estimated that DiMaggio's streak odds were 1 in 9,545. Moran, on the other hand, felt that it would be "more accurate [to run] a series of simulated 1941 seasons based on DiMaggio's overall batting statistics, rerun the simulation about a zillion times and see how often a fifty-six game hitting streak [came] up." As a computer programmer, this challenge appealed to me. But I took the project a step further than what Moran suggested. Instead of using DiMaggio's entire 1941 season, I wanted to put the hitting streak under a microscope. In other words, using the furious pace that produced DiMaggio's historic streak, what kind of odds was the Hall of Famer facing—or, what where the odds of DiMaggio doing what he did when he did it?

CUP OF (MR.) COFFEE

Having done past simulations (see 2000 *BRJ*, "The Ten Thousand Careers of Nolan Ryan"), one thing I learned was that if the sample is large enough, the results are almost predictable. That is, if the odds of something happening are 1 in 1,000, it may happen three times in 1,000 or it may happen not at all, but it won't happen 95 times. Yet there was something in the amateur mathematician in me that wanted me to at least understand why DiMaggio beat such long odds. In order to do so, I had to start simple. A hitter who hits well enough to sustain a long hitting streak would be someone who hits around .375, or three hits in eight at-bats, on average. My first task was to figure out what the odds were of a player coming up from the minors for a late September cup of coffee, playing two games, getting three hits in eight at-bats and having a two-game hitting streak. At first I thought the odds were 50-50. That is, his hit totals in the two games could be 2-1, 1-2, 3-0 or 0-3. That approach is incorrect because the distribution of 2-1, 1-2, 3-0 and 0-3 isn't equal—there are more 2-1 or 1-2 possibilities than there are 0-3 or 3-0. It turns out there are 56 ways to get three hits in eight at-bats with four at-bats per game. Table one shows the 56 permutations:

Table 1. 56 permutations (0=out, 1=hit,); a space has been left between each "game"

0000 0111	0000 1101	0001 0011	0001 0110
0001 1010	0010 0011	0010 0110	0010 1010
0100 0011	0100 0110	0100 1010	1000 0011
1000 0110	1000 1010	0011 0001	0011 0100
0101 0001	0101 0100	0110 0001	0110 0100
1001 0001	1001 0100	1010 0001	1010 0100
1100 0001	1100 0100	0111 0000	1101 0000
0000 1011	0000 1110	0001 0101	0001 1001
0001 1100	0010 0101	0010 1001	0010 1100
0100 0101	0100 1001	0100 1100	1000 0101
1000 1001	1000 1100	0011 0010	0011 1000
0101 0010	0101 1000	0110 0010	0110 1000
1001 0010	1001 1000	1010 0010	1010 1000
1100 0010	1100 1000	1011 0000	1110 0000

For example, in the left uppermost cell, the hitter batted four times in the first game and failed to get a hit (0000). In the second game he went three-for-four, getting hits in his last three at-bats (0111). Of those 56 combinations only eight fail to secure a two-game hitting streak, the first four and the last four. The number or permutations, *p*, is determined by the following equation:

```
p = pa! / (h! × ~h!)
```

JOE D'ANIELLO *lives in Niskayuna, New York, with his wife and two children. He met Joe DiMaggio at a baseball card show in 1986, and Joltin' Joe graciously posed for a picture with the author's father.*

Where *pa* is plate appearances, *h* is hits and *~h* is the number of non-hit plate appearances (*pa* – *h*). In other words $p = (1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8) \div ((1 \times 2 \times 3) \times (1 \times 2 \times 3 \times 4 \times 5))$ or in real numbers $p = 40320 \div (6 \times 120) = 56$.

That means that Joe Cupacoffee has a 48/56 (85.7%) chance of having a two-game hitting streak, much better than the 50% I originally thought. But as Mr. Cupacoffee's games increase linearly, *p* increases exponentially. My computer could handle only up to an eight-game hitting streak, which took two hours to run and had 225 million permutations. I estimated that a 12-game hitting streak (40 billion permutations) would take my computer a couple of days and a 14-game hitting streak (7 trillion) about seven years.

Table 2. Odds of .375 hitter having a two-, four-, six-, or eight-game hitting streak, assuming the hitter gets four at-bats per game

G	AB	H	STREAKS	PERMUTATIONS	STREAK %
2	8	3	48	56	85.7
4	16	6	4,480	8,008	55.9
6	24	9	466,944	1,307,504	35.7
8	32	12	51,429,376	225,792,840	22.8

Obviously, the more games played, the more difficult it is to maintain a hitting streak. Basically, each additional two games played had roughly two-thirds the chance of its predecessor of being successful (i.e., 55.9 ÷ 85.7 = .65). Blahaus noted in his article that each game plate appearance above and beyond four added little to the odds of getting a hit in that game, but reducing the number of plate appearances significantly impacted the odds. If Mr. Cupacoffee's at-bat/game split is changed from 4+4 to 3+5, his odds of reaching a two-game hitting streak slip from 85.7% to 80.4% (45/56). With a 2+6 split, his odds of success see a steeper slide to 71.4% (40/56).

OH! THOSE BASES ON BALLS

It has been argued that DiMaggio's temperament factored into the hitting streak equation. That may be so, but his batting habits had more of an impact. The man knew he was the big gun in the Yankee lineup, and he wanted to swing the bat, figuring he had a better chance for a hit than the hitters behind him in the lineup. This logic may go against current wisdom that hails high on-base percentages, but it is essential for maintaining a long hitting streak. During his streak DiMaggio walked only 21 times and was hit by a pitch twice. That's just 23 plate appearances that DiMaggio "wasted" in his batting streak. When a hitting streak is in progress there are only two results: a hit or a non-hit. Outs, errors, sacrifice flies, walks, hit by pitcher... all of them do nothing to forward a hitting streak. True, if the hitter has *all* his plate appearances in a game result in walks, hit by pitches, sacrifices (not sacrifice flies) or reaching first on catcher's interference, the game will be excluded from the hitting streak and the streak will continue. DiMaggio, however, didn't have the luxury of such a ruling, and in general, a walk is as much an anathema to a hitting streak as a no-decision is to a pitcher trying to win twenty games. Throughout the streak DiMaggio batted .408 (91 for 223), but his hit average (Hits ÷ Plate Appearances) was .370. How great is a hit average of .370? Had DiMaggio been able to keep up that pace all season, his hit average would have been the sixth highest ever. The top ten list below should offer no surprises:

Table 3. The top ten hit averages of all time

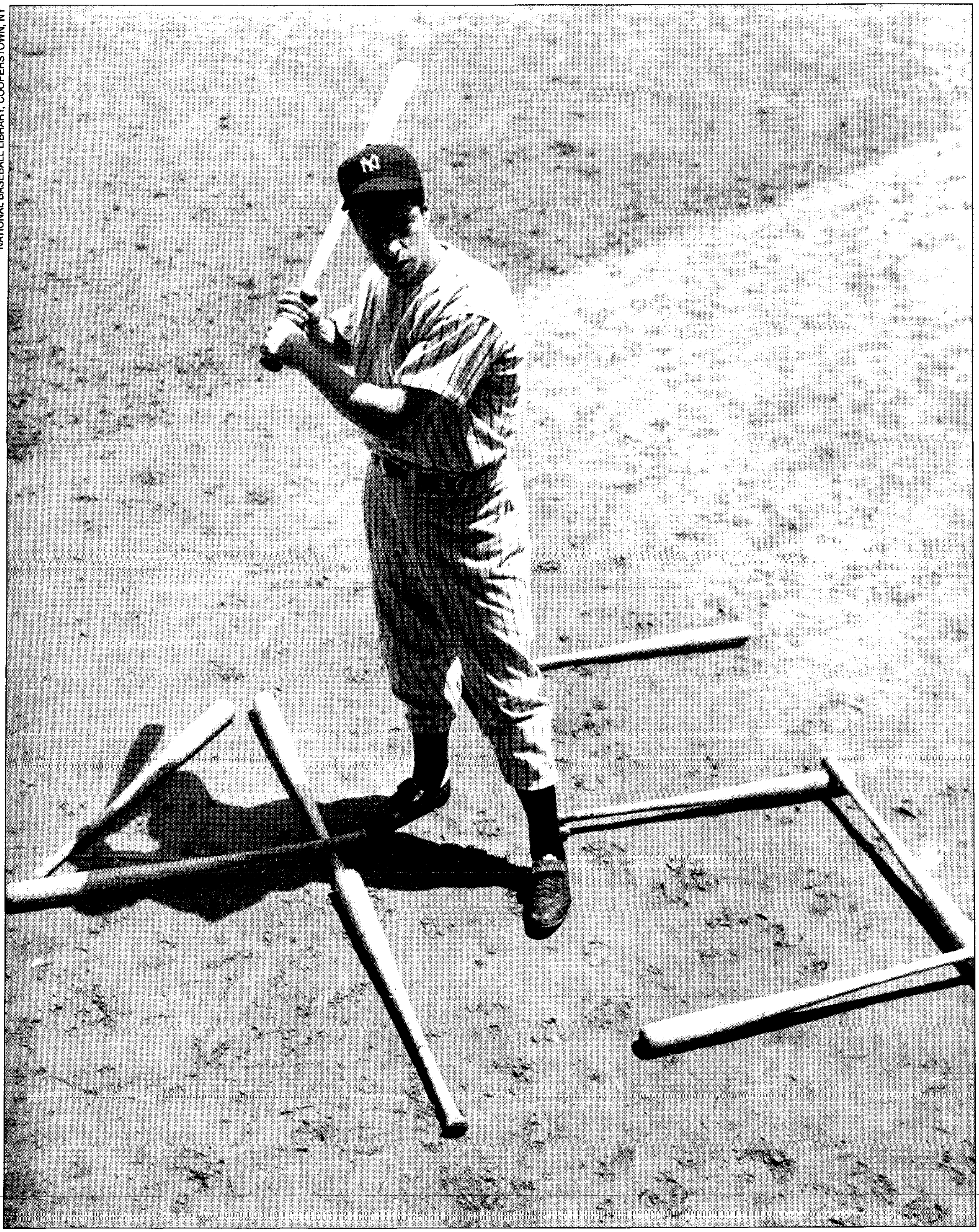
PLAYER	YEAR	HIT AVG.	BATTING AVG.
Nap Lajoie	1901	.399	.422
Ty Cobb	1911	.386	.420
George Sisler	1922	.386	.420
George Sisler	1920	.378	.407
Ty Cobb	1912	.376	.410
Bill Terry	1930	.368	.401
Joe Jackson	1911	.367	.408
Al Simmons	1925	.367	.384
Al Simmons	1927	.363	.392
Rogers Hornsby	1922	.363	.401

For the 1941 season, DiMaggio's hit average of .311 was his fourth best behind 1939 (.340, 64th best in history), 1940 (.313) and 1937 (.312). The top mark of the new millennium is Nomar Garciaparra's .333 (98th best) in 2000. The only player to top .350 since 1930 is Tony Gwynn with .352 (26th best) during the strike-shortened 1994 season.

THE SIMULATION

Based on plate appearances in each game, I simulated DiMaggio's hitting streak exactly as it occurred in 1941. DiMaggio had three games with just three plate appearances and in game 49, a five-inning rain-shortened game, had just two plate appearances. Joltin'

NATIONAL BASEBALL LIBRARY, COOPERSTOWN, NY



Joe's 91 hits during the streak were randomly sprinkled throughout his 246 plate appearances. After one million simulations DiMaggio had 155,536 occurrences where he failed to get a hit in the first game (1 in 6 chance). He had a 54% chance of having a four-game hitting streak, and this matches up well with Joe Cupacoffee's chances as listed in Table 2. DiMaggio had a 1 in 126 chance of reaching the halfway mark (28 games). As for going all the way, it happened 15 times in my one million simulations, giving DiMaggio a 1 in 66,667 chance of success. If you think a sample of one million isn't large enough, I ran the program another one million times. In the second million DiMaggio had sixteen 56-game hitting streaks. If DiMaggio were able to keep up his .370 hit average over an entire season—and that's asking a lot—he would have a 1 in 673 chance of getting a 56-game hitting streak in 1941. The number 673 is derived by dividing his odds (66,667) by the 99 opportunities for getting a 56-game hitting streak in a 154-game season.

WHAT ABOUT TED WILLIAMS?

Interestingly, Ted Williams began a 23-game hitting streak the same day DiMaggio started his record streak. Williams not only batted .406 in 1941, but he out-hit DiMaggio .412 to .408 during the 56-game streak. Nevertheless, Williams' penchant for walking made it virtually impossible for him to sustain a long hitting streak. Williams walked 147 times and was hit by three pitches in 1941. When factored to the same amount of plate appearances that DiMaggio had during his hitting streak (246), Williams would have sacrificed more than twice as many plate appearances over 56 games costing him 15 hits. The Splinter's hit average of .305—the best of his career—was far below DiMaggio's .370. I modified the program to simulate Ted Williams's chances of getting a 56-game hitting streak based on his 1941 statistics (factored down to 56 games), and it wasn't even close. In one million simulations, DiMaggio had 883 hitting streaks of at least 40 games in duration; Williams had nine, with a 43-game streak being the longest. Given that DiMaggio's 1941 hit-average of .311 wasn't much higher than Williams' only heightens the magnitude of his accomplishment.

THE FUTURE

There are some records that will never be broken, like Owen Wilson's 36 triples or Walter Johnson's 113 shutouts. But most of those records are due to a game that has changed. It's unlikely that any pitcher starting his career today will have 113 complete games, let alone that many shutouts. But every day there are hundreds of hitting streaks in progress. In one month a hitter can go from no hitting streak at all to past the halfway mark of the Clipper's streak, yet in 62 years nobody has seriously challenged this magnificent accomplishment and only Pete Rose has managed to get even 75% toward the goal.

But with odds of 1 out of 66,667 for a .408 hitter—hit average .370—it might be wiser to put your money on 37 triples.

SOURCES

Blauhous, Charles. "The DiMaggio Streak: How Big a Deal Was It?," *Baseball Research Journal*, No. 23. Society for American Baseball Research, 1994.

Freiman, Michael. "56-Game Hitting Streaks Revisited," *Baseball Research Journal*, No. 31, pp. 11-15. Society for American Baseball Research, 2002.

Moran, Neal. "Streaks: Statistics vs. Serendipity," *Baseball Research Journal*, No. 24, pp. 79-80. Society for American Baseball Research, 1995.

Neft, David S., Richard M. Cohen, and Michael L. Neft. *The Sports Encyclopedia: Baseball.* New York: Griffin Trade Paper, 2000.

Seidel, Michael. *Streak: Joe DiMaggio and the Summer of '41.* Lincoln, NE: University of Nebraska Press, 2002.

BOB BROWN AND PETER GOODRICH

Calculating the Odds
DiMaggio's 56-Game Hitting Streak

A certain ballplayer, with a lifetime batting average of .280, is about to begin a 10-game road trip. What are the chances that, sometime during the 10-game span, he will "hit safely" in seven straight games? You probably haven't given this question too much thought over the years. Now, here's one you have thought about (perhaps even lost a little sleep over): A lifetime .343 hitter, averaging 4.5 plate appearances a game (Joe DiMaggio's career statistics through 1940), is about to begin a season in which he will play 139 games. What are the chances that, sometime during the season, he will "hit safely" in 56 straight games?

Hitting safely in 56 straight games is a lot like tossing 56 "heads" in a row. Since there's only *one* way it can happen, the probability is easy to compute (in the case of the latter, 1/2 raised to the 56th power). A far more interesting question is this: If a fair coin is tossed 139 times, what are the chances that, sometime during the experiment, "heads" will turn up 56 times in a row? How many ways can this happen? We really couldn't tell you, but here's one:

$$\text{Scenario A: } H_1T_2T_3 \ldots T_{29}\underbrace{H_{30}H_{31} \ldots H_{85}}_{\text{56 consecutive heads}}T_{86} \ldots H_{139}$$

Here's another:

$$\text{Scenario B: } T_1H_2H_3 \ldots T_{29}\underbrace{H_{30}H_{31} \ldots H_{85}}_{\text{56 consecutive heads}}T_{86} \ldots T_{139}$$

Notice that in both scenarios, the "run" of 56 heads begins at the same point, flip #30, but the two are *not* alike. For one thing, scenario A begins with a head, while scenario B begins with a tail. They end differently, too. There are so many different scenarios, in fact, we couldn't begin to write them all down. And so it is with the "streak." No one yet has been able to figure out how many different ways it could happen. And until someone does, we'll never know how improbable it really was. (In case you're wondering, DiMaggio's 1941 season could have turned out in any one of 2^{139} totally different ways. That's a 7 followed by 41 zeroes!)

Imagine our excitement, then, when not too long ago we ran across an article by Charles Blahous entitled "The DiMaggio Streak: How Big a Deal *Was* It?" (*BRJ*, 1994, pp. 41-43), where, in the space of a few paragraphs, the question seemed to have been answered once and for all.

Using a combination of probability theory and intuitive reasoning, he estimated the chances to be around .00134 (or, once every 746 years). Right away we knew something must be wrong. As proof, take a look at a normal distribution table and check out how much area falls under the curve to the *right* of $\mu + 3\sigma$, for all intents and purposes, the *largest* number in the distribution. It's .00135. So, what Mr. Blahous is saying, essentially, is that DiMaggio's streak has roughly the *same* chance of happening as an American-born male has of reaching a height of 6'4" (tall, to be sure, but obviously, not one-of-a-kind tall). If it were really this likely, then it almost certainly should have happened, at the very least, a couple of more times, perhaps as many as four or five times total, in the past 100 years. But it hasn't. Know why? Because the streak is about as likely as seeing someone 7'4" tall on the mound (as opposed to, say, Roger Clemens, who, though *very* tall at 6'4", is fully one foot shorter). Any bets, now, on how long before it happens again?

As a result of all this, we decided to take a look at how Mr. Blahous arrived at this figure. Let's review it together.

PETER GOODRICH, *a 1957 NY Giants batboy, and* **BOB BROWN**, *a life-long Giants fan, who grew up just blocks from the Polo Grounds, now teach courses in statistics and strategic management at Providence College.*

HOW BIG A DEAL WAS THE STREAK?

In figuring the odds, he turned his attention to the following:

1. When DiMaggio stepped up to the plate, what were his chances of getting a hit? Of not getting one?

   Mr. Blahous used Joe's 1941 batting average of .357 (though he probably should have used his lifetime mark of .325—or better yet, his ratio of hits to plate appearances, $H \div (AB+W) = 193 \div (541+76) = .313$. After all, a walk is really no different than a flyout, a groundout, or a strikeout. It's a wasted opportunity). His chances of not getting a hit are then $1-.357 = .643$.

2. What constituted a "typical" game for DiMaggio in 1941?

   Since Joe played in 139 games and had 541 "official" at-bats, a typical game for him that year would have been a game in which he got up to bat 3.89 times officially.

3. In a typical game, what were his chances of not getting any hits? Of getting at least one?

   He handled this one by raising .643 to the 3.89 power (.179). He then subtracted this figure from 1 to get Joe's chances of hitting safely in any game he played (.821).

4. What were his chances of hitting safely in 56 straight games?

   $.821^{56}$, or about .000016

5. How many opportunities did he have during the season to initiate a 56-game hitting streak?

   The streak could have started on Opening Day, or the next day, or even as late as the 84th game. He then reasoned that since there were 84 ways to fit the streak into the season, the final probability must be .000016 times 84, or .00134.

Now, all this seems reasonable (after all, .00134 is fairly small), but if Mr. Blahous had made just a few more calculations, he would have realized something was terribly wrong. To show this, we've taken the liberty of writing down a *general formula* for the procedure described above using the symbols *n, p, L*, and *G* (something Mr. Blahous should have done). The formula is:

(CHANCE OF HITTING SAFELY IN L STRAIGHT GAMES) × **(Nº OF WAYS TO FIT AN L GAME STREAK INTO A SEASON G GAMES LONG)**

$$P(L,G) = (1-(1-p)^n)^L \times (G - L + 1);\ L = 1, 2, \ldots, G$$

where

n = official at bats per game
p = batting average
L = length of hitting streak in games
G = number of games in season

and

P(L,G) = chances of hitting safely in L straight games sometime during a season G games long

Substituting n=3.89, p=.357, and G=139 into P(L,G) yields

$$P(L,139) = (.821)^L \times (140 - L);\ L = 1, 2, \ldots, 139$$

To prove that the formula *doesn't* work, it is sufficient to find just *one* value of L such that P(L,139) > 1 (since no probability can exceed 1).

It turns out there are 23 such values (L=1, 2, . . . , 23). Let's try L=20:

$$P(20,139) = (.821)^{20} \times (120) = 2.32$$

That's 232%! Well, you know what they say. Nothing difficult is ever easy. But don't take it so hard, Mr. Blahous. After all, you're in good company. Nobody else has been able to solve the problem, either.

WHAT MAKES THE PROBLEM SO DIFFICULT?

The truth is, and there's absolutely no way of getting around it, the only way to figure the chances that a ballplayer will hit safely in *L* straight games sometime during a season *G* games long is to (1) write down, literally, every possible way the hitting streak could come about (a formidable task even for relatively small values of G), (2) work out the chances that each of them will actually happen (easy, but time consuming), and (3) add them all up. As an example of what's involved here, imagine that a season consists of merely G=10 games and that, at some point during the season, a certain player manages to hit safely in L=7 of them in a row. What are the odds? Well, on any given day, the player either gets a hit (or perhaps, more than one)—that is, he "succeeds"—or he does not—he "fails." Now, keep in mind that in a 10-game season, any one of 11 possible hitting streak lengths can happen. Even in a season this short, there are 1,024 totally different ways the season could turn out (2 raised to the 10th power), only 12 of which will result in a 7-game hitting streak. Imagine trying to write down all 1024 of them on paper. Unless you had a lot of time on your hands, you probably couldn't do it. Table 1 shows the 12 possible ways the player could have put his seven-game streak together.

Assuming this player has some constant chance, say p(f), of going *hitless* in any one of these 10 games and, thus, some chance 1−p(f) = p(s) of coming away with *at least one safety*, then, by using what is known as the "multiplication rule for independent events," we get the probabilities shown in the far right column.

Table 1. Calculation of P(L=7, G=10)

	GAME										
SCENARIO	1	2	3	4	5	6	7	8	9	10	PROBABILITY
1*	s	s	s	s	s	s	s	f	f	f	$p(f)^3 p(s)^7 = .0022$
2	s	s	s	s	s	s	s	f	f	s	$p(f)^2 p(s)^8 = .0043$
3	s	s	s	s	s	s	s	f	s	f	$p(f)^2 p(s)^8 = .0043$
4	s	s	s	s	s	s	s	f	s	s	$p(f)^1 p(s)^9 = .0087$
5*	f	s	s	s	s	s	s	s	f	f	$p(f)^3 p(s)^7 = .0022$
6	f	s	s	s	s	s	s	s	f	s	$p(f)^2 p(s)^8 = .0043$
7*	f	f	s	s	s	s	s	s	s	f	$p(f)^3 p(s)^7 = .0022$
8	s	f	s	s	s	s	s	s	s	f	$p(f)^2 p(s)^8 = .0043$
9*	f	f	f	s	s	s	s	s	s	s	$p(f)^3 p(s)^7 = .0022$
10	f	s	f	s	s	s	s	s	s	s	$p(f)^2 p(s)^8 = .0043$
11	s	f	f	s	s	s	s	s	s	s	$p(f)^2 p(s)^8 = .0043$
12	s	s	f	s	s	s	s	s	s	s	$p(f)^1 p(s)^9 = .0087$
											.052

To illustrate, let's return for a moment to our hypothetical .280 hitter. Suppose his lifetime statistics are:

G	H	AB	W	PA	H/PA	PA/G
1000	960	3428	572	4000	.24	4

Each time he comes to bat (walks included), he has a 24% chance (960/4000) of getting a hit (or, equivalently, a 76% chance of not getting one). His chances, then, of going hitless in a typical 4 at-bat game? About 1 chance in 3 (.76 raised to the 4th). This leaves him with 2 chances in 3 of getting at least one hit (thus keeping the streak alive).

Making these two substitutions (that is, p(f) = 1/3 and p(s) = 2/3), we see that a .280 hitter has a bit more than 1 chance in 20 (*.052*) of hitting safely in 7 straight games sometime during a 10-game season. The 10 other hitting streak lengths, L=0, 1, 2, . . . , 6, 8, 9, 10, would account for the remaining 95%, or so, of the probability in the distribution. Not exactly easy, is it? And there's no other way to do it.

The four scenarios in Table 1 that are marked with asterisks (nos. 1, 5, 7, and 9) are, according to Mr. Blahous, the only four ways a seven-game hitting streak can come about in a 10-game season. What he fails to realize, of course (and this is where he runs into trouble), is that what happens on either side of the streak, both before it ever begins as well as after it eventually ends, is every bit as important in figuring the overall probability as the length of the streak itself. In short, every game matters. *Every single one of them.*

But ignore them he does. Not to mention scenarios 2, 3, 4, 6, 8, 10, 11, and 12. How would Mr. Blahous have figured the chances of such a streak? Using his (*incorrect*) methodology, he likely would have raised .6757 to the 7th power, then multiplied by 4:

$$P(L=7,\ G=10) = (1-(1-.280)^{3.428})^7 \times (10-7+1)$$

$$= (.6757)^7 \times 4$$

$$= .257$$

resulting in an answer about five times larger than the true probability.

Think about it logically for a minute. No math. Just plain, ordinary common sense. Could a .280 hitter have this much of a chance, better than one chance in four, of hitting safely in seven straight games sometime during a season *only* 10 games long? With the available supply of probability limited to 100% (a supply that must be sufficient to cover all the outcomes possible, 11 in this case), *how could so much of it be concentrated at one single point*, a point that one would not ordinarily associate with having much of a chance to happen in the first place. Well, it couldn't. To begin with, he'll probably only get hits in 7 games total, and it's fairly unlikely that all 7 will come in a row. If he gets hits in 8 or 9 games, he's got a better chance of hitting in 7 straight. Of course, if he gets hits in 6 games or less (or all 10), he's got no chance whatsoever of accomplishing the feat. Where does this leave us? Hitting safely in seven straight games, particularly in a season this short, sounds awfully hard to do. Don't you think? But if you're still not convinced, there's one more thing we can do.

A SIMPLE EXPERIMENT: 1,000 SEASONS WITH A DIE

How about an experiment to settle the issue once and for all. Get a six-sided die. Toss it. If it turns up 1, 2, 3, or 4 (which it will 2/3 of the time), write down an *s* (for success). If it turns up 5 or 6, write down an *f* (for failure). Repeat 9 more times. Each 10-toss sequence (the equivalent of a 10-game season) will produce an L-value. Carry out the experiment 100 times. Keep track of how often L=0, 1, 2, . . . , 10. Here's what we got:

STREAK LENGTH (L)	0	1	2	3	4	5	6	7	8	9	10
OCCURRENCES	0	2	17	27	20	15	11	2	3	2	1

It took us perhaps half an hour to generate these results. Desperate for more data, though, we turned to our students for help. Forty-five volunteered to carry out the same procedure 20 times each (without knowing what the output would be used for). Here's what they got:

STREAK LENGTH (L)	0	1	2	3	4	5	6	7	8	9	10
OCCURRENCES	0	20	140	221	188	131	89	52	27	17	15

Combining all the output yields the distribution shown below:

STREAK LENGTH (L)	0	1	2	3	4	5	6	7	8	9	10
% OF TIME	0	.022	.157	.248	.208	.146	.10	.054	.03	.019	.016

Notice that:

1. In 54 of the 1,000 seasons, his longest hitting streak was 7 games, within 2 of what we expected (and no where near the 250 or so Mr. Blahous would have predicted).
2. He hit in every game 16 times, just one short of what we expected (found by raising 2/3 to the 10th power).
3. He never had a hitless 10-game season (the chance of which, .000017, is found by raising 1/3 to the 10th).
4. Most of the time (in 248 seasons), the best he could manage was a 3-game hitting streak.

The most important thing we learned, however, was that the *more* we simulated, the *closer* the actual and theoretical results tended to be.

For example, after 100 seasons our .280 hitter had just two 7-game hitting streaks (about half of what we anticipated). He ended up, though with 54 such streaks in 1,000 seasons (a virtual bull's-eye). This gave us an idea. Since we have no way of figuring out mathematically just how likely a 56-game hitting streak is, why not simulate a whole bunch of 139-game seasons, and

NATIONAL BASEBALL LIBRARY, COOPERSTOWN, NY

keep track of how often it happens (just like we did above). That's exactly what we decided to do.

Two problems immediately arose. First, we couldn't possibly carry out the simulation by hand (a 10-game season is one thing; a 139-game season is quite another). Clearly, we would need the assistance of a computer. Second, since the probability of a hitting streak this *long* is so *small*, we're obviously going to have to carry out the experiment a fairly large number of times in order to avoid making the kind of mistake that could easily ruin the investigation, namely, concluding that it can't *possibly* happen simply because we never *saw* it happen. We know it's going to. Eventually. On some kind of *regular* basis, too. But how long before *this* happens? 5,000 runs? 10,000? 50,000? To be on the safe side (that is, in order to be absolutely certain we wouldn't over- or underestimate the chances of such a streak), we decided to run the experiment a *million* times (10 trials of 100,000 runs each). Yes, that's an awful lot, but we wanted to be sure.

THE SIMULATION

In order to run the simulation, we only need to know one thing: Joe's chances of hitting safely in any *one* of those 139 games. Rather than use his career statistics or his 1941 batting data alone, we decided to base our estimate on his performance covering the first five years of his career, 1936-1940.

Table 2. Joe DiMaggio's batting statistics, 1936-40

GAMES	HITS	AT-BATS	WALKS	PLATE APPEARANCES
686	969	2827	260	3087

Though he averaged .343 over this period, his chances of getting a hit each time he came to the plate were only .314 (969/3087). Dividing 3,087 by 686, we see that he averaged exactly 4.5 plate appearances per game. Now, if we raise .686 (Joe's chances of going "hitless" each time he came to the plate) to the 4.5 power, we get his chances of going "hitless" in a typical game. This yields .183. Thus, we estimate his chances of getting *at least one hit* in any one of those 139 games at around .817.

The simulation is straightforward and easy to carry out. It works like this: For each game, a three-digit number is drawn from a random numbers table (computer generated). If it falls anywhere between 000 and 816, this means DiMaggio "hit safely" in that day's game (he gets an *s*). If it falls between 817 and 999, he went "hitless" that day and gets an *f*.

On the first run (season #1), DiMaggio's longest hitting streak was 17 games (games 65 through 81). After 13 runs (the actual length of Joe's career), nothing longer than 23 straight. After 1,000 runs, still no streaks of any real consequence. After 5,000 runs, the longest streak we observed was 36 games (season #3,932). We were beginning to think *it* would never happen. Then on run 7,693, *it* finally did. A 59-game marvel covering games 29 through 87. If we had stopped simulating at this point, we would have given Joe one chance in 7,693 of hitting in 56 games or more consecutively. We kept going. After 10,000 runs, he had done it twice. After 50,000 runs, 15 times. After 100,000 runs, 23 times total, or once every 4,348 seasons. All the results are summarized below.

Table 3. Summary of 100,000 simulated seasons

STREAK LENGTH (L)	OCCURRENCES 1ST 50,000 RUNS	OCCURRENCES 2ND 50,000 RUNS
10 or less	1,393	1,413
11-15	16,061	15,944
16-20	18,083	18,179
21-30	12,492	12,465
31-40	1,724	1,754
41-55	232	237
56+	15	8
TOTALS	50,000	50,000

He hit in 56 straight just six times (Mr. Blahous would have predicted 134 such streaks). His shortest streak was seven games (achieved 25 times); his longest, a mind-boggling 70. About two out of every three streaks fell between 11 and 20 games in length—better than 93% between 11 and 30. The entire batch averaged out to 18.70 with a standard deviation (sd) of 4.77, meaning that what happened in 1941 is the equivalent, more or less, of reaching into the adult American male population and pulling out someone 3" taller than "Shaq" (not someone merely 6'4", which would have put him in the extreme right-hand tail of a normal distribution, but someone fully *8 sd's taller than the average man*). It shouldn't have happened. And probably won't ever again. No matter how many times you repeat the experiment.

Thrilled with our results, but wondering whether 23 "successes" in 100,000 runs was typical or not, we decided to repeat the experiment. Not surprisingly, we didn't get 23 the second time around. We got only 21. Since there was some variation, we thought, what the hell, let's keep going and see what happens. So we tacked on 800,000 more runs (bringing the grand-total to one million 139-game seasons). Don't laugh. It took a long time. Here are the results:

Table 4. Summary of 1,000,000 simulated seasons

TRIALS ($1X10^5$ RUNS EACH)	56+ GAME STREAKS	56 GAME STREAKS
1	23	6
2	21	7
3	21	4
4	22	6
5	24	5
6	19	5
7	24	6
8	23	5
9	21	4
10	24	6
TOTALS	222	54

Notice how *close together* these numbers are. Solid evidence, we think, that our final estimate of Joe's chances of hitting safely in 56 or more games consecutively sometime during the 1941 season,

$$P(L\geq 56,\ G=139) = 222/1 \times 10^6 = .000222$$

is correct. Not to mention our estimate of his chances of hitting in 56 straight,

$$P(L=56,\ G=139) = 54/1 \times 10^6 = .000054$$

SUMMARY

Mr. Blahous sums up by saying, "What all this means is that DiMaggio, given his .357 batting average and the at-bat chances that he received, did something that he shouldn't have been expected to do unless he hit that way for 1,038 years." Well, we're not sure where this figure comes from (it's actually closer to once every 18,519 years), but we are sure of this: what DiMaggio did in 1941 has no chance, realistic or otherwise, of ever happening again. No one's seriously approached the record in 60 years (not even Rose). Want a comparison? In 1884, "Old Hoss" Radbourn (playing for Providence, who else?) started 73 games, completed all 73, and won 60 of them. In his career, he started 503 games and failed to go the distance only 14 times. Must have been a real "horse," huh? Weighed 168, 15 pounds less than Gene "Stick" Michael. But the point we want to leave you with is this: there isn't one person on the face of the earth who thinks these records will ever be broken. It's the same thing with DiMaggio's 56-game hitting streak. It's here to stay.

HERM KRABBENHOFT

Ted Williams' On-Base Performances in Consecutive Games

Does Teddy Ballgame Hold That Important Major League Record Too?

On-Base Performance is very important—it is the absolute prerequisite for the most critical aspect of playing winning baseball—scoring runs. In order for a team to score a run, at least one of its players must get on base.

The conventional metric for On Base Performance is On Base Average (OBA), oftentimes called On Base Percentage (OBP)—which is the total number of times a player gets on base *safely* divided by his total plate appearances. "Officially," there are three ways in which a batter can get on base safely: (1) by getting a base hit; (2) by drawing a walk; (3) by being hit by a pitch. Not included in the "official" means of getting on base safely are plays involving catcher's interference, dropped third strikes, a fielder's choice, or fielding errors.

From a historical perspective, On Base Average (or On Base Percentage) came into being in the middle 1950s, thanks to two men—Allan Roth (a baseball statistician with the Brooklyn Dodgers) and Branch Rickey (the general manager of the Pittsburgh Pirates). In an August 2, 1954, article in *Life* magazine, Rickey stated, "The ability to get on base, or on base average, is vital." He deduced his position from a statistical analysis carried out by Roth which revealed a strong correlation between on-base average and runs scored—"OBA went hand in glove with runs scored." Rickey concluded his article with the following statement: "Baseball people—and that includes myself—are slow to change and accept new ideas. But they will accept this new interpretation of baseball statistics eventually. They are bound to."

It took 30 years for Rickey's stance to reach the mainstream—it was not until 1984 that *The Sporting News Baseball Guide* (as well as *Street & Smith's Baseball Yearbook* and the annual editions of Neft and Cohen's *Sports Encyclopedia: Baseball*) began including OBA in their statistical tabulations.

HERM KRABBENHOFT *is a retired organic chemist. He is the only person to ever have his research published in* Baseball Digest, Baseball Weekly, The Sporting News, *and* Baseball America *as well as* Die Makromoleculare Chemie *and also be granted a United States patent, #5,021,521 (Branched Thermoplastics)—in a single season (1991).*

Indeed, getting on base safely is now recognized as an essential, if not the preeminent, batting skill.

Switching gears, for a moment, it is pointed out that performance streaks are an interesting, if not also important, part of the diamond game. For example, tremendous excitement was generated by Cal Ripken in his quest for the record for *most consecutive games played, career*, which now stands at 2,632. Other interesting examples of major league records for consecutive performance streaks include (1) *most consecutive victories for a pitcher, season*—19 by Tim Keefe and Rube Marquard (in 1888 and 1912, respectively); (2) *most consecutive errorless games at shortstop, season*—100 by Rey Ordonez (in 1999); (3) *most consecutive seasons leading one's league in home runs*—seven by Ralph Kiner (National League, 1946-1952).

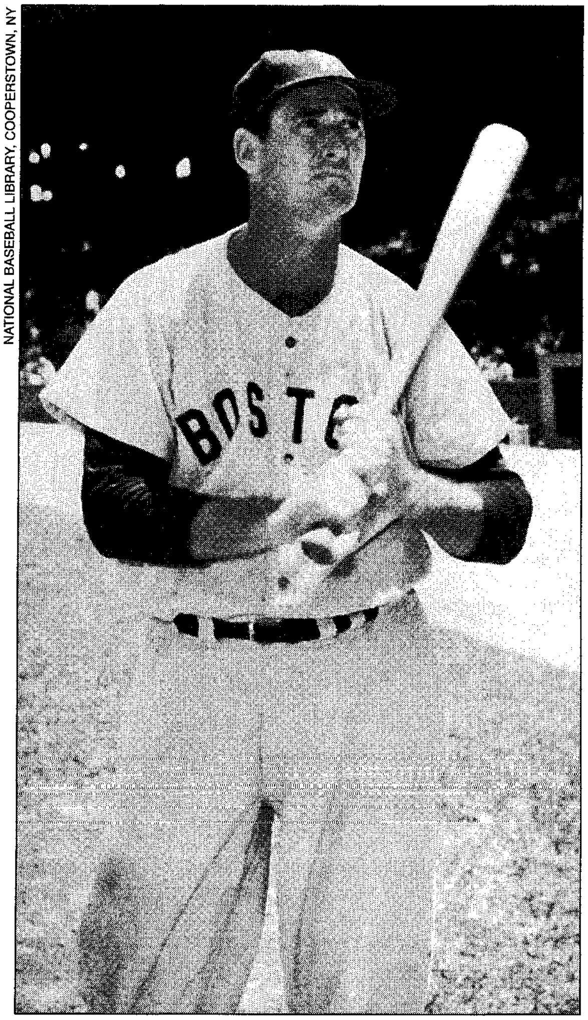
With that introduction, let's ask the following question: "Who holds the major league record for *most consecutive games reaching base safely, season*?"

Resorting to the various baseball record books and encyclopedias to find the answer is fruitless—none of them provides the answer for this important record. Of course, the major league record *for most consecutive games getting on base via a base hit, season,* is widely known—56 by Joe DiMaggio (in 1941). And the ML record *for most consecutive games getting a base on balls, season,* is also recorded—22 by Roy Cullenbine (in 1947). But there's no mention of *most consecutive games reaching base safely, season.*

The results of my research, directed toward providing the answer to the above question, are presented here.

RESULTS AND DISCUSSION

The starting point for my research endeavor was Joe DiMaggio's record 56-consecutive-games hitting streak. His consecutive games on base safely (CGOBS) streak had to be at least 56 games. My initial query



NATIONAL BASEBALL LIBRARY, COOPERSTOWN, NY

was, "What did The Yankee Clipper do in the game(s) immediately before he began his 56-game hitting streak and immediately after his 56th game. Checking out the official 1941 day-by-day records for DiMaggio revealed some very interesting findings.

Looking at the front end of his 56-game hitting streak, DiMaggio began it on May 15. In the game on May 14, while Joltin' Joe went 0-for-3, he did get on base via a walk; therefore, he had reached base safely. However, in the May 13 contest, Joe went 0-for-4 with no walks and he was not hit by a pitch. So, now DiMaggio had a CGOBS streak of at least 57 games.

Turning then to the back end of DiMaggio's 56-game hitting streak, it was found that while he went 0-for-3 in the game on July 17 (the game his hitting streak ended), he did get on base with a walk. Moreover, Joe then embarked on 16-game hitting streak from July 18 through August 2. Finally, in the game on August 3, he went hitless in four at-bats and had no walks and was not hit by a pitch.

Therefore, from May 14 through August 2, Joe DiMaggio achieved a CGOBS streak of 74 games—a truly phenomenal total!

But is this the major league record for *most consecutive games on base safely, season*?

Perhaps Joe D's 74 CGOBS streak is the record; perhaps not—there could be other players who have exceeded that number. Thus, DiMaggio's 74 CGOBS streak served as the benchmark for my research.

In considering players who might have assembled a CGOBS streak longer than DiMaggio's 74, Ted Williams appeared to be a particularly good candidate. That's because The Splendid Splinter accomplished an extraordinary on-base performance record, as indicated by the following:

- He had the highest single season On Base Average in ML history—.553 (in 1941)—until Barry Bonds surpassed it with a .582 mark in 2002.[1,2,3]
- Ted fashioned the highest career OBA in ML history (.482).
- The Thumper holds the mark for most seasons leading the league in OBA (12).
- *The Sporting News Complete Baseball Record Book* lists Williams with the major league record for *most consecutive plate appearances on base safely*—16 (in 1957).

So did Williams ever surpass DiMaggio's benchmark 74 CGOBS streak?

To find out, I carefully examined the official day-by-day records for Ted Williams for each season in his major league career. The pertinent results:

1939 Williams began his ML career with a CGOBS streak of 15 games (April 20 through May 14, first game). Later on in the season (May 26 through June 11), he compiled a CGOBS streak of 17 games, which would be his longest of the season.

1940 Ted improved his CGOBS streak performance somewhat in his sophomore season, putting together a 29-game skein of reaching base safely (April 28 to June 4). That streak came to an end in a 14-inning contest at Fenway Park at the hands of St. Louis Browns pitchers Jack Kramer and Bob Harris, who combined to inflict an 0-for-7 on Ted. Williams made the final out of the game with the tying run on base. That 0-for-7 would be the worst single-game 0-for in Ted's career.

1941 Williams blossomed in 1941, achieving two major CGOBS streaks. Curiously, Ted embarked on his first long streak on the same date (May 15) that Joe D began his 56-game hitting streak. However, Ted's string was snapped on June 29 in Philadelphia after 44 games; Athletics hurler Jack Knott handcuffed Williams in four plate appearances.

Ted's second long streak started on July 12 (only a few days before DiMaggio's 56-game hitting streak would end). This streak reached 69 games and was still alive after the last game of the season on September 28.[4] Of course, that begs the question—"How would Teddy Ballgame begin the 1942 season?" He was only five CGOBS behind Joltin' Joe's benchmark of 74 CGOBS.

1942 Well, Ted Williams did indeed manage to get on base safely in each of his first five games of the 1942 season. But that was it. Nonetheless, those five games gave Teddy Ballgame a two-season CGOBS streak of 74 games. Over the course of two seasons he had "equaled" DiMaggio's phenomenal single-season mark of 74. (Arguably, Ted's two-season 74 CGOBS streak could be asterisked to distinguish it from Joe's single-season achievement.) Incidentally, the pitcher who prevented Williams from "eclipsing" the Yankee Clipper was Marv Breuer, Joe D's teammate. The right hander shut Williams down in four plate appearances at Yankee Stadium on April 19. Later in the season, Ted accomplished two rather long CGOBS streaks. He had a 35-game streak that ended on July 14 as Johnny Niggeling of the St. Louis Browns imposed an 0-for-4 on him. However, Ted started another streak the very next day; it lasted for 33 games (until August 19). So, Williams reached base safely in 68 out of 69 games. After the 1942 season, because of World War II, Williams was in the military service for the next three years.

1946 After his three-year stint in the military, Teddy Ballgame came back and began the 1946 season with a 41 CGOBS streak. That string was terminated on June 3 in Boston by White Sox pitchers Thornton Lee (a southpaw) and Earl Caldwell (a righty). Undaunted, Ted started another long streak in the very next game. This streak lasted 34 games. So, if it hadn't been for Lee and Caldwell, Williams would have passed Joltin' Joe's single-season mark of 74 CGOBS.

1947 While he had seven double-digit CGOBS streaks, Ted's longest CGOBS streak was a not-spectacular 25 gamer.

1948 Ted began the new campaign with a 65 consecutive games on-base safely streak. That streak—which encompassed all of April, May, and June, came to an end on July 5 at Fenway Park. Interestingly, the pitchers responsible for curtailing Ted's bid to overtake DiMaggio were Joe's teammates—Vic Raschi and Karl Drews. Then, in the very next game, Ted embarked on a 17-game CGOBS streak. He reached base safely in 82 out of 83 games. Had it not been for Raschi and Drews, Williams would have shattered Joe D's mark of 74.

1949 "Wait till next year!" is frequently exclaimed in Boston. And for Ted Williams, 1949 was "next year," at least in terms of his CGOBS performance. Starting on July 1 and lasting until September 28, Teddy Ballgame amassed an 84 consecutive games on-base safely streak. He surpassed DiMaggio's mark by 10 games, thereby establishing a new single-season CGOBS streak benchmark. The pitcher who snapped Ted's streak was Ray Scarborough of the Washington Senators. Scarborough struck out Williams twice and got him to fly out to short center in his three plate appearances. Interestingly, Ted was in the on deck circle when Johnny Pesky was retired for the third out in the top of the ninth.[5] Interestingly, Ted was in the on-deck circle when Pesky was retired for the third out in the top of the ninth.

1950 The 1950 campaign was a difficult one for Ted Williams. In the first half of the season, he put together CGOBS streaks of 28 and 36 games. Had it not been for an 0-for-5 performance on May 28, the two streaks would have been continuous—64 games. Then, in the All-Star game, Ted fractured his elbow and was out of the lineup until September.

1951 In what might be considered an off-year for Teddy Ballgame, the longest CGOBS streak that he could assemble was a 48-gamer (May 30 through July 16). Then, for the next two seasons (1952 and 1953), Ted spent nearly all of his time in the Marines fighting in the Korean War.

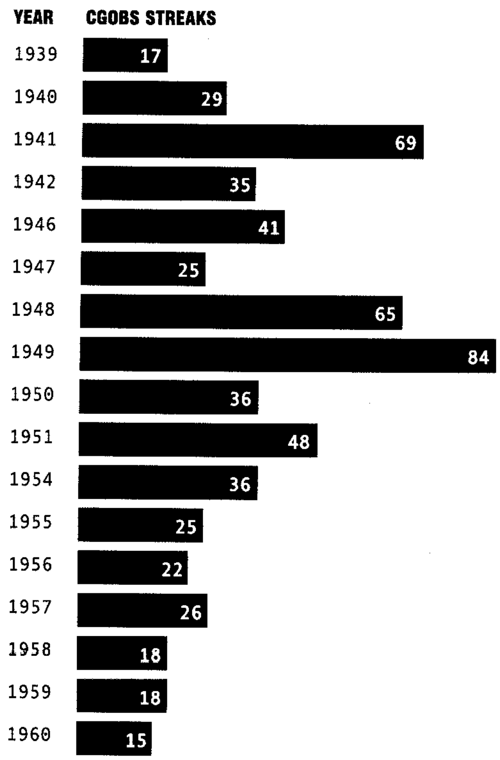
1954–1960 For several reasons, Ted wasn't a full-time, full-season player during the last seven years of his career. Consequently, he never neared the 84 CGOBS streak he achieved in 1949. His longest CGOBS streaks in each of his final seven ML seasons were: 36 in 1954, 25 in 1955, 22 in 1956, 26 in 1957, 18 in 1958, 18 in 1959, and 15 in 1960.

In summary, Chart 1 presents the longest CGOBS streak that Ted Williams achieved in each of his "full" major league seasons from 1939 through 1960. The 84 CGOBS streak in 1949 turned out to be the very best for Teddy Ballgame.

And that leads to the next question: how does The Splendid Splinter's 84 CGOBS streak stack up to the CGOBS performances of other players—especially those with a knack for getting on base frequently?

Ideally, to properly answer this question, one has to examine the consecutive games on-base safely per-

Figure 1. The longest single-season CGOBS streaks that Williams achieved in each season of his major league career (1939-1960)



formances of *every* player who's played in the major leagues since 1876. That, unfortunately, is not realistically possible.[6] However, for this project, I focused on four groups of players from the NL and AL:

(1) The top three to five players in On Base Average for each season from 1891 through 1977.[7]

(2) The top 25 players in (hits plus walks) per game for each season from 1946 through 1974.

(3) All players with consecutive games hitting safely streaks of at least 20 games.
I scrupulously examined the official day-by-day records for each of these players and determined their longest consecutive games on-base safely streaks. All together, I examined the day-by records of nearly 2,000 player seasons from 1891 through 1977.[7]

(4) For the seasons from 1969 and 1974 through 2002 (plus the 1967-1968 American League), thanks to the fantastic efforts and cooperation of Retrosheet—especially Dave Smith—I have the absolute longest consecutive games on-base safely streaks for every major league player (who appeared in at least 30 games in a given season). All together, another 12,000 additional player seasons were searched.

So, how does Teddy Ballgame's 84 CGOBS streak stack up? Table 1 provides the answer, listing all the players I have found who assembled a CGOBS streak of at least 50 games (including one player who achieved the feat during the 2003 season).[8,9] It presents the 40 times a CGOBS streak of at least 50 games in a single season has been achieved. The list is composed of 19 players from the American League—including Ted Williams, who appears three times and has the longest CGOBS streak. Joe DiMaggio and Ty Cobb are the only other AL players who reached the 50 CGOBS streak plateau more than once. The National League is represented by 17 players—including Duke Snider and Barry Bonds, who each achieved a CGOBS streak of 58 games—the longest one found in the senior circuit.[10] The only NL player to achieve a CGOBS streak of at least 50 games more than once was Bill Joyce.

As can be seen, Ted Williams' 84 consecutive games on-base safely streak does indeed stack up as the longest streak out of the very best on-base player seasons examined so far.

Is it the major league record? Very probably!

What is its significance? The following discussion may provide the answer.

In a recent article on Ted Williams,[11] in reference to Joe DiMaggio winning the 1941 American League Most Valuable Player Award, Williams is quoted as stating the following:

"I didn't feel robbed or cheated that year. I believe there isn't a record in the books that will be tougher to break than Joe's 56-game hitting streak. It may be the greatest batting achievement of all."

In terms of approachability, however, Ted's 84 CGOBS streak seems to be more difficult than Joe's 56-game hitting streak. Since DiMaggio achieved that streak in 1941, the closest any major league player has come to it was the 44-game hitting streak by Pete Rose in 1978. Forty-four is 78.6% of the way to 56. Since Williams achieved his 84-game streak in 1949, the closest any major league player has come to it (in a single season) were the 58 CGOBS streaks by Duke Snider in 1954 and Barry Bonds in 2003. Fifty-eight is 69.0% of the way to 84. Furthermore, the closest that any ML player has come to Teddy Ballgame's 84 CGOBS streak over two seasons was the 62 CGOBS streak by Mark McGwire at the end of the 1995 season and the beginning of the 1996 campaign.[9] Sixty-two is 73.8% of the way to 84.

So, with the above approachability considerations in mind, and paraphrasing Ted Williams, it can be argued that Teddy Ballgame's 84 consecutive games on-base safely streak "may be the greatest batting achievement of all." [12,13]

When will it be included in the record books?

Table 1. Single-season CGOBS streaks of 50 or more games

YEAR	TEAM	LG	YEAR	CGOBS
Ted Williams	BOS	AL	1949	84
Joe DiMaggio	NY	AL	1941	74
Ted Williams	BOS	AL	1941	69
Ted Williams	BOS	AL	1948	65
Johnny Tobin	STL	AL	1922	58
Duke Snider	BKN	NL	1954	58
Barry Bonds	SF	NL	2003	58
Cupid Childs	CLE	NL	1892	57
George Kell	DET	AL	1950	57
Wade Boggs	BOS	AL	1985	57
Ed Delahanty	PHI	NL	1896	56
Bill Joyce	WAS/NY	NL	1896	56
Arky Vaughan	PIT	NL	1936	56
Ryan Klesko	SD	NL	2002	56
Billy Hamilton	BOS	AL	1896	55
Ty Cobb	DET	AL	1915	55
Stan Musial	STL	NL	1943	55
Jim Thome	CLE	AL	2002	55
Bill Joyce	WAS	NL	1894	54
Ray Blades	STL	NL	1925	54
Luke Appling	CHI	AL	1936	53
Derek Jeter	NY	AL	1999	53
Shawn Green	LA	NL	2000	53
Denny Lyons	PHI	AA	1887	52
Ty Cobb	DET	AL	1914	52
Tris Speaker	CLE	AL	1920	52
Lou Gehrig	NY	AL	1934	52
Mel Almada	STL	AL	1938	52
Jimmy Wynn	HOU	NL	1969	52
Greg Gross	HOU	NL	1975	52
Tony Phillips	DET	AL	1993	52
Frank Thomas	CHI	NL	1996	52
Gary Sheffield	ATL	NL	2002	52
Joe Kelley	BAL	NL	1896	51
Babe Ruth	NY	AL	1923	51
Ken Williams	STL	AL	1923	51
Joe DiMaggio	NY	AL	1937	51
George Brett	KC	AL	1980	51
Vince Coleman	STL	NL	1987	50
Lou Whitaker	DET	AL	1991	50

REFERENCES AND NOTES

1. The .553 OBA value shown here is two points higher than the .551 value originally shown in the various baseball record books and encyclopedias. That's because I discovered an error in the official 1941 day-by-day records for Ted Williams: in the first game of a doubleheader on September 24, the official day-by-day records show that Williams was hitless in three at-bats and had no walks, and was not hit by a pitch. However, in checking the game accounts and box scores in various newspapers I determined unequivocally that Williams was walked twice in that game. Thus, he actually walked a total of 147 times (not 145) in 1941, and his corrected OBA is therefore .553.[2] It is gratifying to note that *The*

Sporting News Complete Baseball Record Book (2003 edition) does include the walk-related corrections: (1) On page 22, for "Highest on-base percentage, season (100 or more games)," Ted Williams is listed as the AL record holder with .553 in 1941. (2) On page 160, for the yearly leaders in "Bases On Balls, American League," Ted Williams is listed for 1941 with 147. (3) On page 178, for players with "1000 Bases On Balls," Ted Williams is listed with 2,021. Also, in *The Book of Baseball Records* (2003 ed.) by Seymour Siwoff of the Elias Sports Bureau, on page 386, for the "annual batting leaders in walks," Ted Williams is listed with 147. Similarly, the 2003 edition of *The Sports Encyclopedia: Baseball* lists Ted Williams as the AL leader in walks for the 1941 season with the corrected total of 147. Likewise, for the 1941 Boston Red Sox "roster," Ted Williams is listed with 147 walks (in boldface). In the section giving players' lifetime statistics, Ted Williams is listed with the corrected total of 2,021 career walks.

2. For a complete account of the two additional walks, see: (1) "Good Eye Gave Ted Williams More Walks," H. Krabbenhoft, unpublished report (November, 2002); (2) "An Error Discovered in Ted Williams's 1941 Walk Total," L. Spatz, SABR Baseball Records Committee newsletter, page 1 (June 2002); (3) "Baseball Records," *The SABR Bulletin*, page 4 [Volume 32 (July/August 2002). See also the following pertinent accounts: (4) "Ted's 1941 On-Base Percentage Mark Increases," B. Nowlin, *On Deck*, page 3 (Summer 2002); (5) "Two More Walks for Williams," R. Neyer, ESPN.com article (May 29, 2002).
3. Williams' .553 OBA in 1941 remains the record in the AL.
4. However, according to Ted's official day-by-day records, this CGOBS streak had been terminated after 64 games. In the first game of that September 24th doubleheader, the official records state that Williams went hitless in three at-bats and had *no* walks, and was *not* hit by a pitch. But, as indicated earlier,[1,2] Ted did indeed walk twice in that game, thereby extending his CGOBS streak to 65 games. He continued his streak for the last four games of the season, giving him a "living" 69 CGOBS streak.
5. The streak-ending game is famous—infamous—for another reason. According to Dave Halberstam in *Summer of '49*, the game is known as the "Scarborough Game." Scarborough defeated the Red Sox as the Nats rallied for two runs in the bottom of the ninth inning for a 2-1 victory which severely impaired Boston's pennant drive. Halberstam had this to say about Scarborough and Williams: "Scarborough was poison to Ted Williams. Scarborough could decoy Williams better than any other pitcher in the league. It was not just a matter of his pitch selection, it was his motion as well. He would show fastball, and then at the last second, go to his curve. Years later, Williams paid Scarborough the ultimate accolade by stating that he probably chased more balls out of the strike zone with Ray Scarborough than with any other pitcher in the American League."
6. Apparently, *no* day-by-day records (official or unofficial) have ever been compiled for the 1876-1890 NL seasons.
7. For the 1891-1919 period, Pete Palmer provided day-by-day HBP data, which are *not* in the official records. Pete also helped by going through his own day-by-day records to determine the longest CGOBS streaks for 24 players—something I couldn't do since day-by-day records weren't available at the Hall of Fame Library.
8. Table 1 also includes (through the courtesy of Bill Deane) Denny Lyons, who in 1887 achieved a CGOBS streak of 52 games with Philadelphia of the American Association; see Bill's SABR-L posting (May 22, 2001).
9. It should also be noted that Tom Ruane, utilizing the Retrosheet data base, previously reported the CGOBS streaks achieved by Boggs, Phillips, Brett, Whitaker, and Coleman in a SABR-L posting (June 11, 1999).
10. For discussions on Barry Bonds equaling Duke Snider's NL CGOBS streak mark, see: (1) Krabbenhoft, Herman. "Will Bonds Break NL CGOBS Streak Record?" in SABR-L posts on September 20-22, 2003; (2) Krabbenhoft, Herman. "Longest NL CGOBS streak—Snider & Bonds" in SABR-L post on September 23, 2003; (3) J. Roberts, "Bonds On," *The Giants Journal*, <members.aol.com/TGJDIR2/bondson.htm>, October 1, 2003.
11. "A Splendid Life: From Hitting Machine to War Hero, Ted Williams Left His Mark," B. Koenig, *USA Today Baseball Weekly*, July 11, 2002, v. 12, p. 3.
12. Some of the material discussed in this report was presented at the SABR32 Convention—"Teddy Ballgame's On Base Performances in Consecutive Games: Does Ted Williams Hold That Important Record Too?," H. Krabbenhoft, page 39. See also the following articles derived from the research presented here: (1) "Nobody Could Walk in His Footsteps," B. Nowlin, *Boston Globe*, page C-7 (July 8, 2002); (2) "The Greatest Streak Ever," B. Nowlin, *Diehard*, page 20 (August 2002); "Sheffield's Streak Short of Ted's," D. Jenkins, *Chattanooga Times Free Press*, page C-5 (August 4, 2002).
13. On hitting streaks, see: (1) C. Blahous, "The DiMaggio Streak: How Big a Deal Was It?" *The Baseball Research Journal* [1994, pages 41-43]; (2) M. Freiman, "56-Game Hitting Streaks Revisited," *The Baseball Research Journal* [2002, pp. 11-15].

ACKNOWLEDGMENTS

It is with great pleasure that I thank the following individuals for their contributions to my research efforts. Dave Smith wrote the computer program to extract the longest CGOBS streaks from the Retrosheet data base for all players from the AL in 1967 and 1968 and in the ML in 1969 and 1974-2002. Pete Palmer provided guidance on the official day-by-day records and (with collaboration from John Schwartz and Alex Haas) the pre-1920 HBP information. Bill Deane provided insightful comments and suggestions and the information on the Denny Lyons 1887 streak. Tom Ruane, utilizing the Retrosheet data base, provided the CGOBS streaks ≥45 games for the 1980-1998 period. Brian Rash alerted me to Barry Bonds' approach to the NL CGOBS streak record in 2003. While each of these individuals is an outstanding baseball researcher, their greatest contributions are being baseball research enablers. Finally, I also wish to express my gratitude to the Hall of Fame library staff of Rachael Kepner, Russell Wolinsky, Claudette Burke, and Tim Wiles for their cooperation and help in my examination of the official day-by-day records.

HERB CREHAN

Boston Red Sox Spring Training History

From 1901 to 2003

When the 2003 Boston Red Sox reported to Fort Myers, Florida, for spring training, state-of-the-art facilities, a battery of instructors, and a full staff of physical-training specialists awaited them. The minor league facility at City of Palms Park, where the team trains before the exhibition season begins, includes eight batting tunnels and sixteen pitcher's mounds. Every effort is made to ensure that the players have access to the best training facilities in major league baseball.

When Johnny Pesky reported to Sarasota, Florida, in 1942 for his first major league spring training session, conditions were comparatively spartan. Pesky was in Fort Myers as a special assignment instructor for the Red Sox in 2003. He has been associated with the club for more than 50 years and has over 60 training camps as a basis of comparison.

"The best way to describe the clubhouse in Sarasota in 1942 was an old barn with some lockers in it. Our manager, Joe Cronin, was a playing manager, so he had to spend time getting himself in shape. And we only had one diamond with a lumpy infield surface and a terrible hitting background.

"But I'll tell you one thing. We were awfully happy to be in Florida at a major league spring training camp."

This is the Red Sox 11th spring training camp in Fort Myers. Over the years the team has trained in 19 locations in 11 different states.

THE EARLY YEARS

Professional baseball teams have been heading to warmer climates for preseason training for over 125 years. The earliest teams were located in the North, and the trek south dates back to the beginning of baseball. The New York Mutuals trained in New Orleans, Louisiana, prior to their 1869 season.

The early emphasis of spring training was on getting the players back in shape. The 1886 Chicago White Sox trained in Hot Springs, Arkansas, where they took 20-mile hikes daily. The 2004 Red Sox have no long distance hikes planned for this year's camp.

The Boston Red Sox, or "Boston Americans" as they were known at the time, were formed in 1901 as charter members of the new American League. The team was established to compete with the Boston Nationals, who had been Boston's NL entry since 1876.

The new team assembled for the first time at historic South Station in March 1901, boarding a train headed to Charlottesville, Virginia, for spring training. The first recorded score for the Boston Americans was a 13-0 victory over the University of Virginia.

During the next five seasons the team selected the state of Georgia for its preseason training. After completing their successful first season, the Americans picked Augusta, Georgia, for their 1902 preseason headquarters. The following season the team shifted their training camp to Macon, Georgia. The 1903 Boston Americans won the first World Series ever played, and they elected to remain in Macon for spring training through 1906.

In 1907, the Americans made a major switch, moving preseason training to Little Rock, Arkansas, where they remained for two preseasons. Prior to 1908 training camp, the team was rechristened as the "Red Sox."

Spring training was a low-budget operation in the early years. Following the 1908 camp, the team elected to pay their expenses by leaving behind a spare outfielder to play for the Little Rock minor league team that season. Fortunately, the team retained an option for the player's future services, because that spare outfielder was Tris Speaker. Speaker went on to star for the Red Sox from 1909 to 1915 and later earned Hall of Fame selection based on his outstanding 22-year major league career.

HERB CREHAN *is the author of* Lightning in a Bottle: The Sox of '67 *(Branden Publishing, 1992) and a resident of Natick, MA. He writes extensively on baseball and its history for newspapers and for periodicals throughout New England. He is the Managing Director of the actuarial consulting firm Crehan & Associates of Natick, MA, and he is a member of the Society for American Baseball Research. Portions of this article originally appeared in* Red Sox Magazine *(2003, pp. 8-21).*

Table 1. Red Sox spring training sites

1901	Charlottesville, VA
1902	Augusta, GA
1903-1906	Macon, GA
1907-1908	Little Rock, AR
1909-1910	Hot Springs, AR
1911	Redondo Beach, CA
1912-1918	Hot Springs, AR
1919	Tampa, FL
1920-23	Hot Springs, AR
1924	San Antonio, TX
1925-1927	New Orleans, LA
1928-29	Bradenton, FL
1930-1931	Pensacola, FL
1932	Savannah, GA
1933-1942	Sarasota, FL
1943*	Medford, MA
1944*	Baltimore, MD
1945*	Pleasantville, NJ
1946-1958	Sarasota, FL
1959-1965	Scottsdale, AZ
1966-1992	Winter Haven, FL
1993-2003	Fort Myers, FL

* War Years

147° FARENHEIT

In 1909, preseason training was shifted to Hot Springs, Arkansas. Hot Springs is the site of underground thermal springs with temperatures that remain at a constant 147 degrees. The oldest park in the national park system, Hot Springs was the spring training headquarters for a number of major league teams.

The team remained in Hot Springs for two years and then shifted to Redondo Beach, California, for spring training in 1911. Redondo Beach was near the winter home of team owner John I. Taylor. This was the only spring that the Red Sox trained in California.

The Red Sox returned to Hot Springs, Arkansas, prior to their next season and the team won its' second World Series in October 1912. The team remained in Hot Springs for seven years, and they won four World Championships during that span. Apparently the powers of the springs were quite formidable.

Following their 1918 world championship, the team pitched camp in Florida for the first time. They were enticed to Tampa by John McGraw of the New York Giants, who recognized that emerging star Babe Ruth would draw fans to the exhibition games. The Babe did not disappoint. He hit one home run well over 500 feet, and a plaque near the spot where the ball landed recognizes the historic drive today.

In 1919 the Boston Red Sox finished a disappointing sixth in the American League. The team elected to return to Hot Springs, hoping that the elixir of the thermal springs would return them to their former glory. Unfortunately, the magic had vanished and so had The Babe. The team finished no better than fifth place following their next four spring training camps in Hot Springs.

Thinking that a total change of scene might change their fortunes, the team spent their only spring training in Texas prior to the 1924 season. A seventh-place finish was the best the team could muster after training near the Alamo in San Antonio. The team headed to New Orleans, Louisiana, in 1925.

These were the not-so-roaring '20s for the Boston Red Sox. In a March 12, 1925, dispatch to *The Sporting News*, correspondent Burt Whitman noted the team was "tickled to death with their infield situation." But his more prescient observation was, "One of the easiest things for the baseball man to do is to get optimistic in the spring of the year." The 1925 Red Sox finished ninth, 49.5 games behind the American League pennant-winning Washington Senators.

Despite two more preseasons in "The Big Easy," things just got tougher for the team in 1926 and 1927. They finished dead last both years. In 1928 the Red Sox returned to Florida, to Bradenton on the Gulf Coast.

By 1928 Florida had become the location of choice for major league spring training. Ten of the sixteen big-league teams trained in Florida in 1928, and the Grapefruit League was in full bloom.

Al Lang, a transplanted northerner and baseball die-hard, is acknowledged as the driving force in luring teams to the Sunshine State. After several false starts he convinced the St. Louis Browns to train in St. Petersburg in 1914. Recognizing the value of the Florida byline on the sportswriters' reports to their northern papers, other Florida cities aggressively recruited major league teams. The Florida land boom of the 1920s did the rest.

Two seasons in Bradenton produced two more last place finishes, and the Red Sox moved to Pensacola, Florida, for spring training in 1930 and 1931. The team finished eighth and sixth respectively, and moved their spring training headquarters to Savannah, Georgia, prior to the 1932 season. The regular season was probably the low point for the franchise. The team finished in last place, 64 games behind the New York Yankees. The total attendance at Fenway Park in 1932 was 268,715 fans.

SARASOTA, FLORIDA

Finances had become so difficult for Red Sox owner Bob Quinn that he had to borrow against his life insurance policy to fund the team's spring training expenses at their new location in Sarasota, Florida. But just prior to the 1933 season a knight in shining armor arrived to rescue the franchise. In February of 1933, Thomas Austin Yawkey agreed to buy the Boston Red Sox from Quinn and the Yawkey Era began.

Sarasota is located on the Gulf Coast, about halfway between Tampa and Fort Myers. John Ringling had selected the city as the winter headquarters for his Ringling Brothers, Barnum and Bailey Circus in 1927. In 1933, Sarasota was a sleepy little village of about 2,500 citizens.

Red Sox Hall of Fame second baseman Bobby Doerr attended his first spring training camp with the team in 1937, and he remembers Sarasota fondly. "It was a great place to train. I remember in the early years that once you left the small downtown area, it was just open land. Miles and miles of palmetto grass with nothing built on it.

"Old Payne Park wasn't much to speak of. It wouldn't begin to compare to today's parks. But there I was, an 18-year-old, 3,000 miles from home, and I'm playing ball with guys like Jimmie Foxx and Lefty Grove. What a thrill."

In 1938 Doerr was assigned the task of escorting young Ted Williams on the long train ride from California to Florida. Their trip took eight days, as torrential rains swept the southern region of the United States. "Ted was so excited. I remember two older women asking a conductor if he could get Ted to quiet down. He was using his pillow as a bat and trying to get major leaguers Babe Herman and Max West to give him hitting tips.

"I remember walking into the Red Sox clubhouse and introducing Ted to manager Joe Cronin. Ted's greeting to Cronin was 'Hi sport!' I figure Ted earned his ticket to our Minneapolis farm club right then."

Johnny Pesky remembers the long road trips to play exhibition games. "This was before the highways were built and before the fancy buses they have today. It was also long before they built the big bridge over Tampa Bay. I remember we used to take a bus over to Bradenton. Then we would get on a ferry to go over to Tampa to play the Cincinnati Reds. After the game we would have to repeat the whole process. It was a long trip, but we didn't mind it at all."

During World War II, clubs were prohibited from training south of the Potomac River, since the movement of military personnel was the nation's top travel priority. In 1943, the Red Sox trained within ten miles of Fenway Park at Tufts University in Medford, Massachusetts. In 1944, the team headed farther south to Baltimore, Maryland, and in 1945 the team trained in Pleasantville, New Jersey.

With the war successfully concluded, the team returned to Sarasota in 1946. That spring training is one of Johnny Pesky's favorite memories. "The war was over and you got to see guys you hadn't seen for two or three years. It was like a reunion. And we had a great team that year."

Doerr still remembers the camaraderie that the ballplayers had with the circus performers. "It was the darnedest thing. You would be walking down the street, and you would see people of all sizes and shapes. A lot of them were baseball fans, and we would see them at our games. Later, when I was married, I would bring my wife and son to spring training and we would go over to watch the circus performers.

"I remember when the circus train would leave to head north, we would all line up to wave good-bye. It was quite a sight. And I remember in 1950 when they filmed the movie *The Greatest Show on Earth*. Most of the film was shot in Sarasota, and we got to watch stars like Charlton Heston, Jimmy Stewart, and Barbara Hutton at work. It was a lot of fun watching them shoot the scenes."

BARNSTORMING

Both Pesky and Doerr have strong memories of barnstorming their way north at the conclusion of spring training. "We used to link up on a train with the Reds, and we would stop along the way to play exhibition games," Doerr recalls. "I remember playing in towns like Jacksonville, Florida, and Durham, North Carolina. We would change on the train before and after the games."

Pesky especially remembers his first barnstorming trip. "I was still fighting for a major league job, and I had a great game in Lexington, Kentucky. Manager Joe Cronin came up to me and said, 'John, you've made the club.' I'll always remember that."

Bobby Doerr recalls a barnstorming excursion that ended up in Louisville, Kentucky. Louisville was the home both of the Red Sox triple-A farm club for many years and Hillerich & Bradsby, the leading manufacturer of major league baseball bats at the time.

"Ted and I walked over to their factory one morning. I remember we got there a half hour before they opened and we sat on their front steps waiting for them to get there. We toured the factory, and of course Ted asked them a million questions. At one point he slipped a guy a $20 bill, which was a lot of money in those days, and told him to be sure they put extra piney wood in his bats. Ted would go to any length to make sure he was the best hitter in baseball."

For many years the Red Sox preseason concluded with the City Series against their crosstown rivals, the Boston Braves. The series began in 1925 and was always played just before opening day. The Braves and Red Sox met in the first Sunday major league game ever played in Boston on April 14, 1925. Attendance over the years ranged from a few thousand fans to the 33,279 who crowded Fenway Park on April 14, 1946. The teams played their last preseason game in Boston on April 12, 1953, at Fenway Park, honoring a commitment made prior to the Braves' March decision to move to Milwaukee.

Weather was always a question mark for the City Series, and that is what Bobby Doerr remembers best. "It seemed like it was freezing every time we played the Braves. We enjoyed playing them, but we had just spent six weeks in Florida and barnstorming in warm weather. Boy, was it ever cold for those games."

By the late 1950s the facilities at Payne Field in Sarasota had started to deteriorate. After several years of negotiating with the city of Sarasota for an upgrade to the ballpark, the team made the decision to relocate their spring training headquarters. After a total of 23 years in Sarasota, the Red Sox selected Scottsdale, Arizona, as their 1959 spring training site.

SCOTTSDALE, ARIZONA

Arizona was a newcomer to the spring training sweepstakes in comparison to Florida. The Cleveland Indians and New York Giants had moved their spring training camps to Arizona in 1947. This move wasn't the result of a sophisticated study of the advantages of the Arizona climate. Rather, Indians owner Bill Veeck owned a ranch in Tucson, and he convinced Giants owner Horace Stoneham to bring his team along so the Indians would have someone to play.

Scottsdale is located just east of Phoenix in Arizona's Valley of the Sun. When the town was incorporated as a city in 1951, its population of 2,000 occupied just one square mile. At the time of the Red Sox move, the city was not well-known. Today its population of more than 200,000, spreads over an area of 185 square miles.

When the Red Sox arrived in Scottsdale in 1959, they became the fourth member of the Cactus League. The Indians were still in Tucson, and while the Giants had moved from New York to San Francisco in 1958, they continued their spring training headquarters in Phoenix, Arizona. The other Cactus League member was the Chicago Cubs. The Cubs had moved to Mesa, Arizona, in 1952 after spending 24 years training on Catalina Island off the coast of southern California.

Arizona offers major league baseball teams consistently good weather. In March the average high temperature is 75 degrees. And sunshine is almost guaranteed. Unlike Florida, which is subject to extended rainy periods, the average precipitation level for Arizona for the month of March is three-quarters of an inch.

The players loved training in Scottsdale. Fall River, Massachusetts, native Russ Gibson was in the Red Sox organization for 13 years and caught for the major league team from 1967 to 1969. "I remember my first major league spring training camp like it was yesterday. I was a newlywed, and my wife and I got there about a week before camp started. I ran into 'The Monster' (Red Sox reliever Dick Radatz) the first day there, and he helped find us a place at the complex he stayed in.

"What a beautiful part of the country. You knew that every day would be as nice as the day before, or even nicer. And what beautiful scenery there is, with the mountains in the background. I had been at the Red Sox minor league camps at Deland and Ocala, Florida, so being in the big league camp in Scottsdale was really special.

"And the good weather made it easy to get in shape. The only problem was the dry heat. You would work up a sweat and five minutes later it would evaporate. But it was a great location for the hitters. The ball really carried in that air. It really gave the hitters a lot of confidence."

If the batters loved hitting in the Arizona air, the pitchers hated it. Former Red Sox pitcher Billy Monbouquette trained in Scottsdale for seven of his eight years with the team. Monbo, who is a member of the Red Sox Hall of Fame, came to dread facing Giant sluggers like Willie Mays and Willie McCovey.

"One time Willie Mays hit a rocket off me. I mean, it cleared the outfield fence at the 350-foot mark, it soared over the parking lot behind the fence, and it landed beside a swimming pool, which was at least 500 feet from home plate. When the ball left the bat, Yaz (Hall of Famer Carl Yastrzemski) didn't even move. Just stood there with his head down and his hands on his knees.

"I was waiting for him on the top step of the dugout when he came in, and I said, 'Don't you ever show me up like that again! You can at least make some effort to get back under it.' He looked me in the eye and said, 'Bill, I'm not going to play back at the swimming pool.'"

Another Red Sox pitcher of that era, Gene Conley, had a unique spring training challenge. Conley, who pitched for the Red Sox from 1961 to 1963, was also a key member of the Boston Celtics. The perennial world champion Celtics would finish their playoff run just about the time his baseball club was breaking camp. Conley would usually report to spring training in time to say good-bye to his baseball teammates. Conley solved his problem by recruiting retired ballplayers to play with him. "I would look around and see all these guys in their 1930s uniforms," Conley recalled. "It was like having my own Field of Dreams."

Johnny Pesky, who had played his entire big-league career training in Florida, managed the Red Sox in 1964 and 1965. "I loved it out in Scottsdale. Good weather every day, lots to do when you weren't at the ballpark. But you had to be careful about overrating the hitters or getting down on your pitchers. The ball really flew in that light air."

Scottsdale provided great weather and gorgeous scenery, but it wasn't ideal for the Boston Red Sox. The major league camp was separated from the minor league camp in Florida by 2,000 miles. Having only three teams to match their talent against was another drawback. And finally, it was a long way for their loyal fan base to travel. After seven seasons in Arizona, the team decided it was time to head back to Florida.

Table 2. Grapefruit League vs. Cactus League

YEAR	FLORIDA-BASED TEAM	ARIZONA-BASED TEAMS*
1960	12	4
1970	18	6
1980	18	8
1995	19	9
2003	18	12

* The Anaheim Angels train part of the year in Arizona and part of the year in Palm Springs, California.

WINTER HAVEN, FLORIDA

Prior to the 1966 season the Boston Red Sox relocated their spring training headquarters to Winter Haven, Florida. Winter Haven is a central Florida community of about 20,000, located between Orlando and Tampa. Winter Haven had some baseball history: it had been the spring training headquarters for the Philadelphia Phillies from 1928 to 1937.

Before the arrival of the Red Sox, Winter Haven was best known as the home of Cypress Gardens, a theme park featuring lush gardens, animals, and a well-known water-skiing show. Established in 1936, it is recognized as Florida's oldest tourist attraction.

The first spring training camp was relatively uneventful, and the team went on to finish in ninth place for the second straight season. But in 1967, things began to change. The biggest single change was the arrival of rookie manager Dick Williams. Williams ran his first camp like a Marine Corps drill sergeant. When his pitchers weren't throwing or running, he had them playing volleyball on the sidelines.

Russ Gibson still marvels at how organized Dick Williams was. "The pitchers and catchers arrived before the hitters, and as a catcher you spent hours and hours down in your crouch while the pitchers stretched their arms out. You would wear yourself out before the full camp even got under way.

"Dick was aware of this, and realized it was only hurting the catchers. He went out and hired a couple of local guys who could catch, and that kept us from breaking down. That guy [Williams], he thought of everything."

The 1967 season produced the "Impossible Dream Team" and the return of baseball as a New England passion. Spring training in 1968 was Gibson's favorite preseason. "Our 1967 season was so great that we couldn't wait to get back to Winter Haven. There was something really special about going to camp as the American League champs. We thought it would go on forever."

Spring training in 1968 was also special because the Winter Haven facilities had been expanded to provide space for all of the team's minor league players. The Dodgers had pioneered the concept of an organization-wide spring training camp, acquiring an old naval airbase in Vero Beach, Florida, in 1949. With the addition of four playing fields to the Chain-O-Lakes facility, every player in the Red Sox system was at the same location with the same professional instructors.

A sure harbinger of spring for New Englanders is the news report that "the Red Sox equipment truck has left Fenway Park and is en route to their spring training headquarters in Florida." For 24 years the man who made that happen was the late Jack Rodgers, who served as the team's traveling secretary from 1969 until his retirement in 1992.

"It was a little bit like running a small community for a couple of months," Rogers recalled. "I would go down to Winter Haven in mid-January to get things organized. We had to make sure that the facility was all set and to get the local staff organized. Everything had to be in place before the pitchers and catchers reported.

"We helped the players to find housing. The team headquarters was in the Holiday Inn, about a mile from the park. The hotel would be filled with fans from New England. We looked forward to seeing the same faces year after year. In those days we averaged about 3,500 fans at our exhibition games. If we drew over 4,000 it was a good crowd."

Rogers had to deal with myriad problems ranging from visa issues for players arriving from Latin America to overdue video rentals. Winter Haven became a home away from home for Rogers and his wife. "Winter Haven was a small town, and you knew that anybody you ran into was there for baseball. We made a lot of friends there. We still go back to visit."

What was the toughest part of the job for Rogers? "I was the last person to see a player who had been released. They had to see me to get their last check." Always a kind man, Rogers added, "A lot of times it was the best thing that could have happened to the player."

Jack Rogers retired from the Red Sox on December 31, 1992. In mid-January 1993 he was in Fort Myers, Florida, at the Red Sox new spring training facility. "They brought me back as a consultant to help out. It was a short retirement," Rogers chuckled.

FAMILY-FRIENDLY

Winter Haven holds some special memories for former pitcher Bob Stanley. Stanley, who tops the list for lifetime pitching appearances for the Sox, still remembers his first spring training camp. "I was an 18-year-old kid, away from home for the first time. I remember what it was like to compete for a job. We had all been local stars and now we were fighting for a spot.

"My first three years I stayed at the Howard Johnson's Motel with all the other minor leaguers. Then I got to move next door to the Holiday Inn with the major leaguers. That was a big deal.

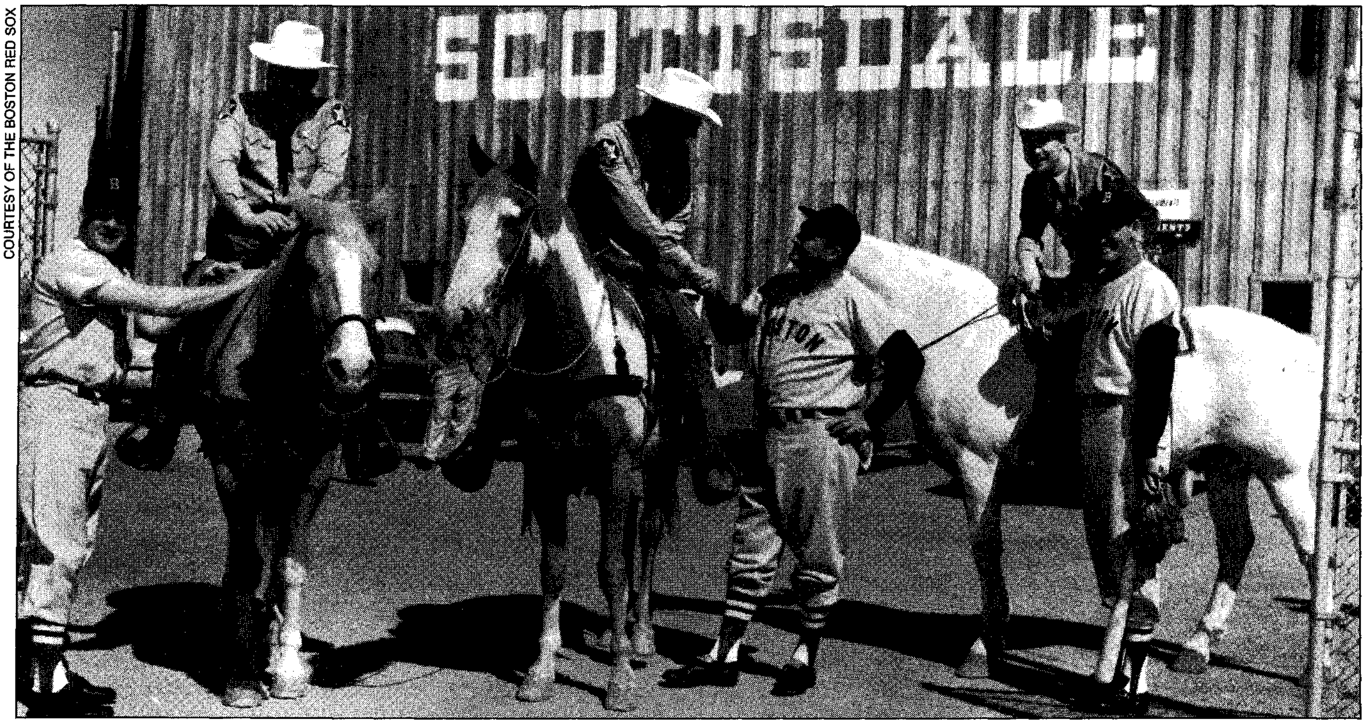
"Over the years as our family grew, it became a real family adventure. We always stayed in the same two adjoining rooms at the Holiday Inn, right beside the kiddies' pool. My wife could sit outside the rooms and enjoy the sun while our kids napped. As the kids got older, we arranged for a tutor so the kids could keep up with their class work and we could be together as a family.

"Jerry Remy used to stay at the Holiday Inn with his family too. Every night we would take our kids, and any other kids who happened to be around, and play ball by the pool. We used to have an Easter egg hunt out there every year. One year there was one egg that nobody could ever find. The next year we were playing by the pool, and I went into the bushes to find the ball and there was the missing egg. We had some great times."

When the Red Sox returned for spring training in 1989, it marked the team's 24th season in Winter Haven. "The Hayve," as the players affectionately called it, had surpassed Sarasota, Florida, as the team's longest running spring training site. But the city of Winter Haven was having trouble maintaining the facility to the Red Sox's standards.

Jim Healey, who served as the Red Sox point man in the selection of a new spring training location, remembers it well. Healey, who worked in the club's front office from 1975 to 2002, understood the city's plight. "The officials wanted to meet their commitment, but like all municipalities they were strapped for cash. Finally, it became an issue of safety. Some of the fields

COURTESY OF THE BOSTON RED SOX



Cowboy up? Jackie Jensen, manager Pinkie Higgins, and Ted Williams meet with the sheriff and two deputies.

were in such tough shape, we were afraid that a player would get hurt.

"We realized that it was time to find a new spring training location. We actually looked at 15 different alternatives in Florida. We came fairly close to working something out with Naples, Florida. When that didn't work out, they suggested that I talk to Mayor Wilbur Smith in Fort Myers."

Table 3. Red Sox Spring Training Sites Ranked by Number of Years

LOCATION	Nº YEARS
Winter Haven, FL	27
Sarasota, FL	23
Hot Springs AR	13
Fort Myers, FL	11
Scottsdale, AZ	7
Macon, GA	4
New Orleans, LA	3
Bradenton, FL	2
Little Rock, AR	2
Pensacola, FL	2
Augusta, GA	1
Baltimore, MD*	1
Charlottesville, VA	1
Medford, MA*	1
Pleasantville, NJ*	1
Redondo Beach, CA	1
San Antonio, TX	1
Savannah, GA	1
Tampa, FL	1

* War Years

FORT MYERS, FLORIDA

Fort Myers is a city of 50,000, located on the Caloosatchee River, about 65 miles south of Sarasota. Nearby Fort Myers Beach, Sanibel and Captiva Islands provide direct access to the Gulf of Mexico. The city first came to national prominence when inventor Thomas Edison built his winter home and a laboratory there in 1885. Edison was the driving force in importing and planting the trees that give the community its "City of Palms" nickname.

"The city was anxious to revitalize its downtown, and they saw our ballpark as a catalyst," recalls Jim Healey. "They were great to work with. The one stumbling block was the lack of land to build the major and minor league facilities together. We were able to locate the minor league facilities about two miles down the street, and that has worked out fine.

"The city voted six to one in favor of the bond issue for the facility, and we broke ground in March of 1992. I actually served as a 'member' of the Fort Myers city council for one day in order to put the project together. We worked with the HOK architectural firm from Kansas City, and created the ideal spring training ballpark. We started construction in May 1992, and we were finished in time for spring training in 1993.

"I remember the first game at City of Palms Park very well. We were scheduled to play Boston College, and they were stuck at Logan Airport in a snowstorm. I spent most of the day on the telephone with Massport. Their plane finally got clearance and they landed in Fort Myers at 5:30 P.M. for a 7:00 P.M. game. They came directly to the ballpark and played very well, so it all worked out."

City of Palms Park has a capacity of about 7,800, including standing room, and it has been a hit with fans from the beginning. In their first season the Red Sox drew 96,421 fans to 15 home games for an average of almost 6,500 per contest. By 2003, average home attendance had grown to over 7,500 per game.

"The Minnesota Twins had moved to Fort Myers in 1991, and initially they were very opposed to our move to the city," Healey remembers. "But we were convinced that our presence would create a rivalry that would help both clubs. And that's the way it has turned out. We have been selling out, even turning fans away, and the Twins drew over 100,000 fans last year."

Spring training gives the players a chance to get in shape, and it allows the front office to fine-tune the roster. And it is also a special time for the fans. George Berardi of Woburn has been to almost every Red Sox spring training camp since 1967.

"Bill Crowley, the late Red Sox public relations director, started the BoSox Club in 1966, and I was one of the early members. The Sox had finished ninth the year before, but it was still exciting to be in Winter Haven in 1967. We stayed at the Holiday Inn with most of the players, and in the evening you could sit around the pool and hear guys like Sam Mele and Frank Malzone talk 'inside baseball.'"

Berardi, whose father Joseph took him to his first Red Sox game in 1936, enjoys watching the young players move through the Red Sox system. "You can watch them at the beginning of camp and then see them get their assignments to the different levels. It's fun watching them develop. My wife Ann and I have made a lot of friends at spring training over the years. We've gotten to know players like Mike Andrews, Rico Petrocelli, and Jim Rice. It's a great time."

Table 4. Red Sox Spring Training Attendance for Selected Years

		ATTENDANCE	
YEAR	LOCATION	AVG.	TOTAL
1977	Winter Haven, FL	2,924	29,240
1983	Winter Haven, FL	3,481	41,769
1988	Winter Haven, FL	4,050	60,747
1993	Fort Myers, FL	6,428	96,421
1998	Fort Myers, FL	5,560	94,517
2003	Fort Myers, FL	7,667	115,000

(Statistics courtesy of MLB.com)

SPRING TRAINING 2004

One of the special moments of any spring training is the discovery of a future star. In 1964 at spring training in Scottsdale, after one season in Class D ball, Tony Conigliaro opened eyes with his slugging. Manager Johnny Pesky told the press, "Of course I'm going to bring him up to the big leagues. They would kill me in Boston if I didn't!"

In the current age of sophisticated scouting and in-depth reporting, there is less chance for an unknown to burst upon the scene. Yet when Shea Hillenbrand reported to the Red Sox camp in 2001, he was assigned uniform number 71 and he was projected to be assigned to triple-A Pawtucket for his sixth minor league season. Jimmy Williams, who was the Red Sox manager at the time, spotted the talent in the 25-year-old rookie, and made him the team's regular third baseman. Shea Hillenbrand was an American League All-Star in 2002.

It is too soon to tell if a future star will emerge from the 2004 Boston Red Sox spring training camp. But one thing is certain. Hope *springs* eternal for every Red Sox fan.

FREDERIC REAMER

The Accuracy of Preseason Forecasts

Each spring produces a new crop of preseason forecasts by baseball columnists and other sports pundits. About a week or two before the first official pitch is thrown, prognosticators survey rosters, injury reports, and trade rumors, and speculate about the order of finish in each American and National League division. Page one of the typical newspaper sports section presents side-by-side comparisons of columnists' predictions.

Rarely, however, does one find any serious end-of-season accounting of the accuracy of these forecasts. An exception was the ESPN.com column (October 30, 2002), which reprinted the forecasts made by ESPN.com's baseball staffers at the start of the 2002 season and compared them with the season's final standings. As a measuring rod, ESPN.com assigned one point for each place missed in the final standings.

After reading the ESPN.com column, I realized that an adaptation of a well-known statistic would provide a much more sensitive and robust indicator of the accuracy of preseason forecasts and facilitate comparisons among various staffers and publications. I think this approach can be used broadly whenever sports columnists predict the final standings in any sport.

The statistic on which my approach is based is known as *Spearman's rho*. It is one of many statistics that researchers use to analyze quantitative data. This particular statistic is known in the trade as a nonparametric measure and provides a clear, direct indication of the extent to which there is a linear relationship, or correlation, between two sets of ranked (or what are called ordinal) scores. Put more simply, Spearman's rho allows one to determine the extent to which two sets of ranks of the same phenomenon - in this case, the preseason and post-season ranks of each baseball division's teams—are similar. Did the final standings resemble the preseason forecast or not?

Based on a mathematical formula that looks intimidating but is very easy to calculate, Spearman's rho is a number, or coefficient, that is very intuitive and easy to interpret. The final coefficient is a number between +1.0 and -1.0. If the two sets of ranks (e.g., preseason and final standings) are identical, Spearman's rho would be +1.0, i.e., the postseason order of finish is identical to the preseason forecast. If the two sets of ranks are completely reversed, the Spearman's rho would be -1.0, i.e., the team that the columnist thought would finish first actually finished last, the team the columnist thought would finish last actually finished first, and so on up and down the line. Spearman's rho will equal zero, or will be close to zero, if the results are very mixed, i.e., some predictions were on the mark, some were close, and some were way off the mark.

For illustrative purposes, I compared the accuracy of the preseason forecasts (2002 season) of five ESPN.com staffers: Jayson Stark, Rob Neyer, Jim Caple, Matt Szefc, and Sean McAdam. I calculated the Spearman's rho coefficients for each staffer's predictions for each of the three divisions in both leagues. Thus, I calculated six Spearman's rho coefficients for each staffer. Positive coefficients are better than negative, i.e., they indicate more accurate preseason forecasts. The maximum (and best possible) total score for each staffer was +6.0 (the sum of a coefficient of +1.0 for each of the six divisions). The worst possible total score was -6.0. In order to come up with a summary score, I added each staffer's six coefficients (whether positive or negative) and divided the total by +6.0 (the best possible outcome). For example, Matt Szefc's six Spearman's rho coefficients were:

AL EAST: +1.0 (Perfect prediction of the final standings.)

AL CENTRAL: +.6 (A pretty good result—what hurt Szefc most was that he predicted the Twins, who finished in first place, would come in third.)

FREDERIC REAMER, *Ph.D., is professor in the graduate program, School of Social Work, Rhode Island College. His research and teaching address issues related to criminal justice, professional ethics, and public policy. His most recent book is* Criminal Lessons: Case Studies and Commentary on Crime and Justice *(Columbia University Press).*

AL WEST: +.4 (A fair, but not great result—Szefc predicted the Mariners to come in first, but they came in third.)

NL EAST: 0.0 (The results were quite mixed—Szefc accurately predicted the Braves' first-place finish, but he predicted the Expos, who finished second, would come in last; he predicted the Mets would come in second, but they came in last.)

NL CENTRAL: +.54 (A good result—what hurt Szefc the most was that he predicted the Cubs, who finished fifth, would come in second.)

NL WEST: +.6 (A good result—Szefc predicted the Padres, who finished in last place, would come in third.)

The sum of these coefficients is +3.14 (out of a maximum total of +6.0).

An easy way to compare the accuracy of each staffer's predictions is to calculate a final score based on the sum of each division's Spearman's rho coefficients (adding up both positive and negative coefficients) divided by +6 (the largest possible total if all preseason forecasts perfectly match the final standings). A set of perfect predictions would generate a final score of +1.00 (+6 divided by +6); a set of perfectly inaccurate predictions (the final results are the reverse of the predictions) would generate a final score of -1.00 (-6 divided by +6). Positive final scores are better than negative final scores; scores closer to +1.0 are the strongest and scores closer to -1.0 are the weakest. Scores around zero are truly mediocre.

One can also compare the staffers' final scores with a final score corresponding to what Phil Birnbaum calls "naïve predictions," that is, the extent to which the final standings at the end of one season (in this case 2001) accurately forecast the final standings at the end of the following season (2002). To compute the final score for naïve predictions, one can calculate a Spearman's rho coefficient corresponding to the naïve predictions for each American League and National League division and calculate a final score by adding up these six coefficients and dividing by +6 (the largest possible total if all of the naïve preseason forecasts, based on the final standings at the end of the preceding season, match the final standings at the end of the following season).

In Szefc's case, the final score is +0.52, indicating a good but not outstanding set of predictions (+3.14 divided by +6.0). One can then compare this final score with the scores of all other staffers who made preseason predictions; the higher the score the more accurate the prediction. For the ESPN.com staffers and naïve predictions, the results for each league[1] and overall are:

	AL	NL	SCORE
Jayson Stark	.70	.53	.62
Sean McAdam	.73	.45	.59
Rob Neyer	.63	.47	.54
Matt Szefc	.67	.38	.52
Jim Caple	.67	.35	.51
GROUP AVG. SCORE	.68	.44	.56
NAÏVE PREDICTIONS	.67	.56	.62

This statistical strategy facilitates easy comparisons among staffers and the naïve predictions. Jayson Stark was the most accurate prognosticator and Jim Caple was the least.[2] Overall, the group's preseason forecast was good, but not outstanding (+.56). Also, the group's preseason forecasts for the American League (+.68) were much more accurate than for the National League (+.44). The naïve predictions for the American League (+.67) were virtually identical to the staffers' group average (+.68), but the naïve predictions for the National League (+.56) were considerably better than the staffers' group average (+.44). The overall final score for the naïve predictions (+.62) was somewhat higher than the final score for the group of staffers (+.56). None of the staffers had a higher final score than the final score based on the naïve predictions.

This statistical procedure—the *Reamer Prediction Accuracy Index*—provides a straightforward, precise, and intuitive measure of the accuracy of preseason predictions in any sport and at any level.

SPEARMAN'S RHO

The formula for Spearman's rho, which can be found in any standard statistics textbook, is:

$$1-\left(\frac{6\Sigma D^2}{n(n^2-1)}\right)$$

where D is the difference between each pair of ranks and n is the number of items (i.e., teams) ranked. In our case, each team in the preseason forecast would be ranked from first to last (rank=1, rank=2, and so on). These same numbers or ranks would then be listed based on the actual post-season finish and D is calculated by subtracting the post-season rank from the preseason rank. For example, if the predicted order of finish were: Yankees, Red Sox, Blue Jays, Orioles, and Devil Rays, the Yankees would be assigned rank #1, the Red Sox rank #2, the Blue Jays rank #3, and so on. If the Blue Jays finished in first place, D would equal +2, which is the result of 3 (the predicted place in the final standings) minus 1 (the actual place in the final standings). Some D scores will be positive (when teams do better than predicted), and some will be negative (when teams do worse than predicted). D will equal zero if the team's place in the final standings is exactly as predicted, no matter where in the standings the team falls. Each D score is squared (to avoid the problem of working with positive and negative numbers), and the sum of these squared D scores is multiplied by 6 (the mathematical reasons for this are complex). This result is divided by the mathematical product of the number of ranked teams (n) and the total of the number of ranked teams squared (n^2) minus 1 (again, the mathematical reasons for this are complicated and can be read in any traditional statistics textbook). This number is then subtracted from 1, and the result is the Spearman's rho coefficient, which ranges from +1.0 (a perfect direct correlation) to –1.0 (a perfect inverse correlation).

For an overview of Spearman's rho, see R. P. Runyon and A. Haber, *Fundamentals of Behavioral Statistics*, 2nd ed. (Reading, MA: Addison-Wesley, 1971), pp. 102-104.

NOTES

1. The American and National League scores were calculated by adding the three Spearman's rho coefficients for each league (one coefficient for each division) and dividing by 3 (the largest possible total if the preseason forecasts match the final standings).
2. The results produced by this statistical procedure are different from the results produced by ESPN.com's analysis, which was based only on how far each *individual* team's actual finish was from the predicted finish (as reported in the ESPN.com column, each place missed in the final standings was worth one point). The ESPN.com approach does not take into consideration the overall *pattern* among the ranked scores when the two sets of ranks are compared. The advantage of my approach is that it takes into consideration the overall *pattern* of ranks and compares the *total* preseason forecast with the *total* post-season results.

CURVEBALL AT THE LETTERS **by Bill Nowlin**

In the only major league boxscore in which he'll ever appear, Bill Schlesinger can't even find his whole name. It was shortened to Schles'ger. The space taken by the apostrophe replaced the "n"—leaving his name actually shortened only by the width of the typset letter "i." Conigliaro got his full name in the same boxscore, but Carl Yastrzemski had to settle for "Y'trz'ski." The boxscore appeared in the *New York Times*. Since the game was a West Coast game, there was no boxscore in the *Boston Globe*—nor did they run one the following day.

And the record books show Bill's name as "Rudy" Schlesinger . . .

Ranking Baseball's Best Single-Season Home Run Hitters

Seventy-three home runs in a single season would have been considered absurd only a few years ago, yet Barry Bonds accomplished this astounding feat in 2001. Bonds' record-breaking season, eclipsing the mark of 70 set by Mark McGwire only four years earlier, raises the question of which ballplayers truly own the greatest single-season home run performances in baseball history. Although Bonds and McGwire shattered the previous standards, their numbers may be somewhat tainted as home run totals in recent seasons have experienced significant inflation.

Prior to 1998, only two players had ever hit 60 home runs in a single season. The 60 home run barrier, virtually unattainable for decades, was surpassed six times between 1998 and 2001. This rapid rise in the number of home runs creates complications in comparing the statistics of Babe Ruth to those of Barry Bonds or other recent players. In order to compare home run statistics of recent seasons to those from years ago in a straightforward manner, a method that accommodates for the fluctuations in home run totals throughout baseball's history is desirable. This is done by calculating the number of standard deviations an individual's home run statistics were above the average player's statistics for that particular season.

Accounting for the disparity in home run totals created by the era in which a hitter played is an effective way to compare various players and determine the best single season home run performances of all time.

THE *Z*-SCORE CALCULATION

A z-score is a calculation of how many standard deviations a statistic is above or below the mean. The standard deviation is a measurement of dispersion from the mean. The following formula is used to produce the z-score:

$$\frac{x - \mu}{\sigma}$$

where:

x = the statistic of the individual player whose z-score is being calculated
μ = the mean for the entire league
σ = the standard deviation

Instead of examining home run totals directly, we use home run rates in the z-score calculation. This is the percentage of at-bats that result in home runs. The use of home run rates places all players on an equal level, so a player with a high number of at-bats will not receive a favorable z-score simply because he had more opportunities to hit home runs. This does not penalize Bonds, Ruth, or other sluggers who frequently received bases on balls because pitchers refused to pitch to them and risk the more serious damage of being taken deep. The use of home run percentages has the added benefit that a player who missed games due to injury is not disadvantaged. Since every player is on a level playing field, home run rates produce a more accurate ranking of the greatest single-season home run performances.

STRENGTHS AND WEAKNESSES OF THE METHOD

The z-score is best considered a measure of domination, since it only determines how well a hitter performed with respect to his contemporaries within the same season. This is perhaps both the greatest strength and the major weakness of using this method to rank home run hitters. The primary strength is that the era in which a hitter played should have no influence on where he is ranked. This is essential for developing an accurate ranking since a variety of factors have severely inflated home run totals in recent years. These factors include:

- Smaller ballparks
- Advances in training and medicine
- A smaller strike zone and lower pitching mound

KYLE BANG *graduated from the the University of Minnesota—Duluth in 2003. This is his first article for a SABR publicaton.*

- Greater protection such as batting helmets and elbow guards
- Better lighting

Conversely, players from earlier eras had advantages over current players. Among these are:

- Fewer relief pitchers
- Less travel
- Less competition due to the absence of African American and Latin players

These factors will not affect the z-score since players are only measured against contemporaries that played under the same conditions.

However, there are a few shortcomings. It cannot adjust for home run discrepancies created by variations in the sizes of ballparks within a single season, and this will have an evident impact on the list. Some parks are simply easier to hit home runs in than others. A modern example involves the Colorado Rockies. Players for the Rockies rank artificially high due to inflated home run rates caused by the high elevation and thin air in Denver. Another flaw that can skew rankings is an abnormal talent level during a particular season. If there was an unusual abundance or absence of prodigious power hitters, every z-score in the league could be understated or overstated because of a very high or low league home run rate. This appears to have taken place at several points throughout history, most notably during the 1950s and 1960s, when many legendary sluggers were active.

RANKING THE GREATEST SINGLE-SEASON PERFORMANCES

The mean home run rate and standard deviation for each season are computed using the statistics of all players with at least 400 at-bats, and the z-score calculation is performed for the individual players with the highest home run rates in each respective season. The resulting z-score is used to develop a list that ranks the most dominant single-season home run totals.

Table 1 lists the top 25 single-season home run performances according to z-score. The z-score calculation has been made for the top performers in every season from 1920 to 2002. Seasons prior to 1920, commonly referred to as the Deadball Era, are excluded from calculations because the enormous size of most of the era's ballparks, combined with the "dead" baseball, made home runs rare. The years 1943 to 1945 are also excluded, when many of the game's best hitters served in World War II. A list of the 100 greatest seasons can be found at www.d.umn.edu/~jgallian/zscore.html.

Table 1. Most dominant HR performances 1920-1942, 1946-2002

RANK	PLAYER	YEAR	Z-SCORE
1	Babe Ruth	1920	7.971
2	Babe Ruth	1921	6.619
3	Babe Ruth	1926	6.415
4	Babe Ruth	1927	5.568
5	Babe Ruth	1924	5.450
6	Babe Ruth	1928	5.304
7	Barry Bonds	2001	5.138
8	Mark McGwire	1998	4.691
9	Jimmie Foxx	1933	4.638
10	Babe Ruth	1923	4.571
11	Babe Ruth	1922	4.474
12	Cy Williams	1923	4.436
13	Jimmie Foxx	1932	4.340
14	Rogers Hornsby	1925	4.244
15	Babe Ruth	1931	4.167
16	Hank Greenberg	1938	4.074
17	Lou Gehrig	1936	3.994
18	Babe Ruth	1933	3.985
19	Mike Schmidt	1981	3.890
20	Mark McGwire	1999	3.857
21	Hank Greenberg	1946	3.849
22	Ted Williams	1942	3.843
23	Babe Ruth	1932	3.824
24	Lou Gehrig	1927	3.820
25	Mark McGwire	1996	3.808

RUTH AND HIS CONTEMPORARIES

The results prove that Babe Ruth is undoubtedly the most dominant home run hitter ever in terms of z-score, as no other player even approaches Ruth's numbers. Ruth occupies the top six positions on the list and 11 of the top 25. His 1920 season produced an absolutely astronomical z-score of 7.971, drastically outdistancing the next highest z-score, Ruth's 1921 season of 6.619. Ruth obliterated the home run record by hitting 54 home runs in 1920, nearly doubling his old record of 29 set the previous season. Ruth utterly dominated the early 1920s, but his numbers became less stunning as the decade progressed and several players began to rise and challenge Ruth by producing comparable home run figures. Ruth does possess the top rate for each year during the decade except for 1925, when he missed a significant portion of the season due to an illness. Ruth ranks among the top 50 for all 13 seasons of his career in which he had at least 400 at-bats.

The 1920s and 1930s is the golden age of dominant power hitters. Besides Ruth, the best seasons of Jimmie Foxx, Lou Gehrig, and Hank Greenberg rank very high on the list. Foxx dominated the early 1930s, as his 1932 and 1933 seasons both rank among the top 15. Foxx's 48 home runs and home run percentage of 8.38 in 1933 placed ninth on the list, trailing only Ruth and the record-shattering performances of McGwire and Bonds in 1998 and 2001, respectively. Hank Greenberg appears near the top of the list despite missing nearly five entire seasons at the peak of his career due to his participation in World War II. Both Greenberg's 1938 and 1946 seasons are among the 25 best seasons.

The major surprise from Ruth's era is the 1923 season of the Phillies' Cy Williams. Williams tied Ruth with 41 home runs and had a home run rate just slightly below Ruth. Williams had several outstanding seasons throughout the 1910s and 1920s, but his home ballparks made a significant contribution to his statistics. Williams began his career with the Chicago Cubs prior to the construction of Wrigley Field and was then traded to the Philadelphia Phillies. In each instance, he took advantage of extremely short fences in right field as a left-handed pull hitter.

CURRENT PLAYERS

Two of today's superstars fared exceptionally well when the *z*-score was applied to their statistics. The historical seasons of Barry Bonds and Mark McGwire rank seventh and eighth, respectively, following only the six best seasons of the unparalleled Ruth. Although home run rates have skyrocketed to astounding new heights over the past few seasons, both Bonds and McGwire are still off the charts. Bonds' 2001 *z*-score of 5.138 places him alongside Ruth as the only players to surpass five standard deviations above the mean home run rate. Bonds had an absolutely ludicrous home run rate of 15.34, meaning he would have hit .153 solely with home runs. McGwire, who consistently places in the top 50, is the only other recent player found among the all-time leaders. Based on *z*-score, McGwire is arguably the second greatest power hitter in baseball history, behind only Babe Ruth. McGwire makes five appearances in the top 50, and he would have more had he not struggled with injuries for the bulk of his career. For example, McGwire's incredible 32 homers in only 236 at-bats in 2000 gave him a home run percentage of 13.56, just slightly behind his record pace of 1998. This would have been the third highest home run rate ever had he not missed half of the season battling injuries.

The remainder of active sluggers do not produce impressive *z*-scores like Bonds and McGwire. Sammy Sosa, who possesses three of the top six single-season home run totals of all time, failed to place any of his performances among the top 50 in terms of *z*-score. His best *z*-score of 3.252 in 2001 ranked 52nd. Sosa's memorable 1998 season is hurt by his unusually high number of at-bats for a power hitter, which lowers his home run percentage. Also, Sosa amazingly did not lead the league in home run percentage in any of these three seasons, guaranteeing that at least one player must be ranked ahead of him from all three years. The *z*-score, a measurement of domination over competitors, shows that Sosa's statistics are not as incredible as they may initially appear.

Sosa is not alone in this category. The home run statistics of most active players are considerably less impressive after the *z*-score calculation is performed. The dramatic rise in the mean home run rate, depicted in the graph on the following page, has lowered *z*-scores. Consequently, home run totals that initially seemed remarkable are just typical statistics. One of the most telling figures that demonstrates this inflation is that 36 of the top 100 single-season home run totals have occurred since 1996. A prime example of a low *z*-score resulting from high inflation is Luis Gonzalez's 57 home runs in 2001. He ranks twelfth on the single-season home run list, but his *z*-score of only 2.4 standard deviations above the mean does not come close to cracking the list of the top 100 *z*-scores. Obviously, when compared only to his counterparts of 2001, Gonzalez cannot compete with the home run statistics of Bonds or even Sosa.

Ken Griffey, Jr., who is frequently mentioned in the same breath with baseball's most legendary hitters, is another active player with zero appearances in the top 100. Griffey did have some outstanding seasons, but none where he is distinctly separated from the competition, which is essential for achieving a high ranking in a method that measures domination. In general, *z*-score accurately depicts where each current player truly ranks among the great home run hitters. Barry Bonds and Mark McGwire, the two players who do possess

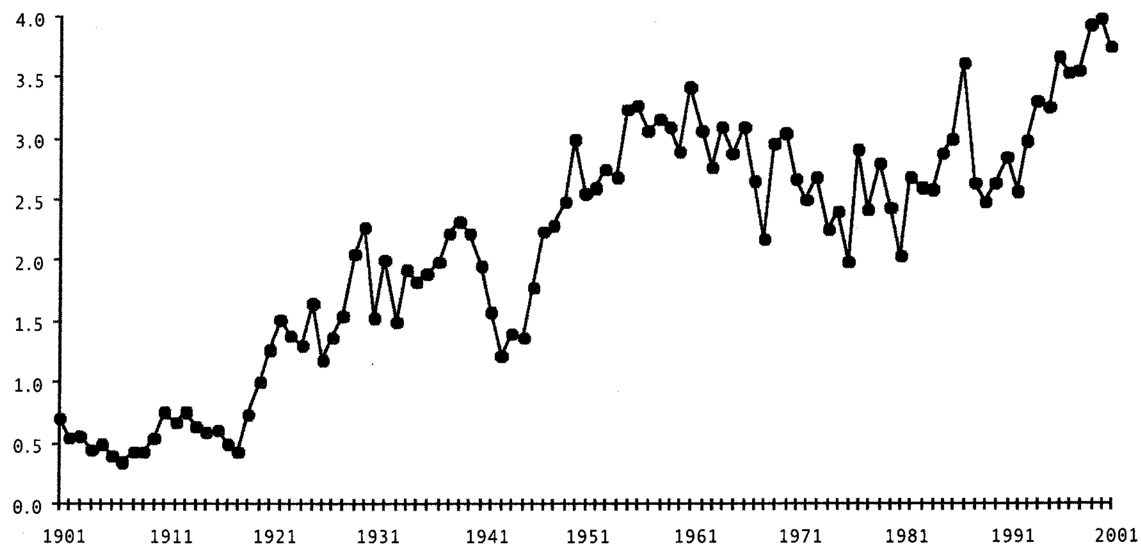


Figure 1. Baseball's increasing home run rates, 1902-2001

dominating home run totals and rates, rank very high on the list, and the remaining players whose figures are elevated due to home run inflation are found far down the list in their proper location.

PROBLEMS WITH THE METHOD

The ranking based on z-score is a vast improvement from the basic home run list, where it is virtually impossible to compare any statistics from different eras. However, this method does have its flaws. A problem that arises throughout the 1950s and early 1960s is the best example of the deficiencies of this method. With 61 home runs in 1961, Roger Maris held the single-season home run record for 37 years until it was shattered by McGwire. Maris's 1961 performance ranks 77th on the list, between players such as Hal Trosky and Bob Horner. Maris totaled nearly 600 at-bats for the season, making his home run percentage relatively low compared to his actual home run total because of his uncommonly high number of at-bats for a power hitter. Nevertheless, ranking 77th is peculiarly low for a season of 61 home runs under any circumstances. Home run rates did soar to previously unseen levels in 1961, partly due to the remarkable performances of Mickey Mantle and Roger Maris and partly because of the American League expansion. Both home run percentage and total home runs in 1961 were actually comparable to today's figures, as the 1961 season has the highest overall league home run rate excluding 1987 and the past seven seasons. Even excluding Mantle and Maris, home run rates in 1961 and throughout the entire era were abnormally high. Consequently, no player from the 1950s or 1960s produces a very high z-score.

THE TALENT FACTOR

It's strange that the 1950s and 1960s, an era famous for its wealth of power hitters, did not produce a single z-score that ranks high on this list. Harmon Killebrew's 1963 season and Willie Mays' 1965 season are the only two seasons between 1950 and 1970 that rank among the top 50. Throughout these two decades, a group of nine or ten players consistently separated themselves from the remainder of the league. This group includes legendary sluggers like Mickey Mantle, Ted Williams, Willie Mays, Hank Aaron, Harmon Killebrew, Ralph Kiner, Ernie Banks, Rocky Colavito, Frank Robinson, and Eddie Mathews.

These legends were clustered together at the top of the home run chart, raising the league home run rate for the season by a considerable amount. Since the z-score method favors a situation in which only one or two players distance themselves from the pack, no player between 1950 and 1970 is found near the top of the list. With the overwhelming talent that existed during this time, no individual player was capable of notably distinguishing himself from the remainder of the league and producing a high z-score. This is the biggest weakness of using z-scores, since an unusually strong year or era or an unusually weak year or era will distort figures because of the abnormal overall home run percentage. This happened as home run rates peaked during the 1950s and 1960s, and the effect seems to disappear as the 1960s came to a close and rates declined. Home run rates would not return to the levels seen during this era until the mid 1990s.

The sharp decline in home runs as baseball entered the 1970s might be partially due to the opposite effect of what happened in the 1950s and 1960s. The 1970s and 1980s are not remembered as an era with an abundance of talented power hitters. Home run rate leaders include mediocre players such as Dave Kingman, Gorman Thomas, Ron Kittle, and Rob Deer. These players took vicious swings and occasionally hit tape-measure blasts, but they also compiled astonishing strikeout totals. The one legitimate slugger of the time was Mike Schmidt, whose best season of 1981 places 19th on the list. Schmidt's extraordinary performance is frequently overlooked since he tallied only 31 home runs in this strike-shortened season. Several z-scores from this era rank much higher than one would expect, including Schmidt's 1981 season and Dave Kingman's 1976 season. Kingman's seemingly mundane 37 home runs and 7.81 home run percentage rank 32nd all-time. The years 1976 and 1981 happen to be the two seasons with the lowest overall home run rates since World War II, and in each case the rate plummeted from the previous season.

The overall talent level is definitely the chief cause of inaccurate rankings. The list appears accurate for the past 15 years and from the 1920s through World War II, but the degree of clustering at the top of the home run leader board from 1950 through the mid 1980s seems to distort the z-scores for these seasons. For instance, Mickey Mantle makes only two appearances in the top 100 despite being widely considered one of baseball's greatest home run hitters. Still, in spite of this slight distortion created by the talent factor, the z-score method is an excellent technique for accurately determining the greatest single-season home run performances.

CONCLUSIONS

Several conclusions can be drawn from this study. The obvious conclusion is that Babe Ruth is far and away baseball's most dominant home run hitter. To this day, we often hear sports announcers use phrases such as "Ruthian numbers." Besides Ruth's dominance, here are other important conclusions from this research:

- The historic seasons of Barry Bonds and Mark McGwire do not approach Ruth's numbers, but they are still among the finest performances ever despite the major inflation home run statistics have experienced in recent years.

- The era with perhaps the most elite home run hitters fared miserably by this method because the league home run rate rose high enough that it was nearly impossible for any individual player to produce an outstanding z-score.

A question that is raised almost every day during baseball telecasts and radio talk shows as current superstars continue to chase and set new records can now be answered. How do today's players compare to Babe Ruth? Although it is impossible to say how Ruth would perform under today's conditions, he is definitely in his own league in terms of dominance of his contemporaries, but Barry Bonds and Mark McGwire have come closer to matching Ruth's dominance than any other player has. Ruth consistently doubled and nearly tripled the home run totals of his closest competitors, a feat nobody else has even approached. He totally changed the face of the game and was the first player to make the home run an integral part of baseball. It is a guarantee that no player will ever dominate the sport the way Babe Ruth did. If Barry Bonds' 73 home runs do not approach the level of dominance that Ruth experienced in his best seasons, it is reasonable to argue that Babe Ruth will always stand alone as the greatest home run hitter in baseball history.

JEROME P. REITER

Should Teams Walk or Pitch to Barry Bonds?

In 2001, Barry Bonds of the San Francisco Giants had arguably the greatest individual season in the history of major league baseball. He set the record for home runs in a season with 73. He hit for the highest slugging percentage ever at .863, breaking Babe Ruth's 1920 mark of .847. He knocked in 137 runs, good for fourth best in the National League, and his batting average was .328, good for seventh best in the league. Bonds achieved these gaudy statistics despite being walked by opposing pitchers a major league record 177 times, besting Ruth's 1923 record of 170. Not surprisingly, Bonds was voted the Most Valuable Player in the National League by an overwhelming margin. It was the fourth time he won the MVP award.

Bonds did not slow down in the next year. He began the 2002 campaign with an incredible display of power, hitting five home runs in the Giants' first four games. His year-end statistics were remarkable: a major league–leading .370 batting average, a National League second-best 46 home runs, and a National League sixth-best 110 runs batted in. He reached base in 58.2% of his plate appearances, eclipsing the previous record of 55.1% set by Ted Williams in 1941. He passed Frank Robinson for fourth all-time in career home runs, with only Willie Mays, Babe Ruth, and Hank Aaron ahead of him. Perhaps most amazingly, Bonds finished with 198 walks, shattering the record he set just last season. Once again Bonds was voted MVP in the NL.

Bonds maintained this remarkable productivity in 2003, despite missing about 30 games. He was among the top three in batting average (.341) and home runs (45) in the National League, and he had a slugging percentage (.749) 82 points higher than Albert Pujols' second best mark. Bonds hit his 45 home runs in 390 official at-bats; in contrast, the National League home run champion, Jim Thome, needed 587 official at-bats to hit 47 home runs. As Bonds' low at-bat total indicates, he was again walked with unusual frequency: he finished with 148 walks, averaging more than one per game. Bonds was voted MVP for the third straight year, the first player to win three consecutively.

JERRY REITER *is an assistant professor at the Institute of Statistics and Decision Sciences at Duke University. His areas of research include statistics in government, social sciences, and sports. He is an avid Red Sox fan and doesn't believe in The Curse.*

The inflated walk totals in 2001, 2002, and 2003 reflect increased use of a strategy for dealing with Bonds' awesome power, namely not to let him have a chance to hit the baseball and instead to walk him. Why risk pitching to a player who averaged one home run every 6.5 at bats in 2001 and every 8.7 at bats in 2002 and 2003? However, walking Bonds is not fail-safe. Putting him on base via walk could actually help the Giants, since runners on base greatly increase a team's chance of scoring runs. Plus, as prolific a batter as he is, over his career Bonds has made an out roughly 70% of the times he is not walked. Why not pitch to him since outs are the most likely outcome?

Thus, we are confronted with an interesting question of baseball strategy: is it better to walk Barry Bonds or to pitch to him? Clearly, the answer depends on the game situation when Bonds steps to the plate. For example, walking Bonds has different consequences when there are zero outs and runners on first and second as opposed to when there are two outs and a runner on third. The answer also depends on the outcomes the opposing manager is concerned about. The manager who seeks to prevent even a single run—for example in late or extra-inning situations in which one run can result in a lost game—evaluates walks differently than does the manager who will concede one run to reduce the possibility of multiple runs.

In this article I examined data from the 2001, 2002, and 2003 seasons to investigate if and when it is better strategy to walk rather than pitch to Barry Bonds. I focused primarily on two game situations: when there is no one on base, and when there is a player on first base only. The conclusions suggested by the data are somewhat surprising: in these two situations, walking Bonds generally is not more effective at preventing runs than letting him hit. In fact, the data even suggest that it is preferable in some situations to let Bonds swing away.

DATA USED IN ANALYSES

To assess the two strategies, I examined data on Bonds' plate appearances in the 2001 through 2003 seasons. The data were retrieved from www.cbs.sportsline.com, which has links to pitch-by-pitch game logs for each game in 2001, 2002, and 2003. Data for a few games were unavailable because of invalid web links; these games are excluded from the analyses. This should not skew results since these games are missing completely at random, that is, they are missing for reasons unrelated to the variables measured.

SOME CHARACTERISTICS OF BONDS' PLATE APPEARANCES

There are 24 possible game situations when a player steps into the batter's box. These are obtained by crossing the three possible out values and the eight possible configurations of players on base. Table 1 displays the number of walks and plate appearances by Bonds in 2001, 2002, and 2003 in each of these 24 game situations. In this table, and for all analyses, intentional and unintentional walks and hit by pitches are included in the walk totals.

In most games in 2001, Bonds batted third in the San Francisco Giants batting order, so that he often came up with two outs and the bases empty or with one out and a runner on first base. In 2002 and 2003, Bonds typically batted fourth in the order, often leading off innings or batting with runners on base and two outs. In all three seasons, most of Bonds' walks were issued with two outs. This is understandable: with two outs, some managers feared Bonds' ability to get extra-base hits more than his teammates' ability to score him from first base. Managers were more reluctant to walk Bonds with zero or one outs, perhaps because multiple players had the chance to advance him home. Bonds walked most frequently when runners were in scoring positions and first base was unoccupied. Walking players in these situations is a common baseball strategy.

Game situations other than those in the "None On" and "First Only" categories have few observations. Hence, the analyses that follow focus primarily on the None On and First Only categories.

Bonds is in the heart of the Giants batting order, so that we expect innings in which he appears at the plate to be the most productive for the team. The relevant outcome is not the total number of runs in each inning, rather it is the total number of runs scored in the inning after the first pitch to Bonds. Hence, for purposes of analyses, runs in an inning is redefined to be the number of players crossing the plate in that inning. Figure 1 shows the frequencies of the run totals for all innings in which Bonds stepped to the plate. In more than 60% of the innings, the team scores zero runs. Scoring more than two runs in an inning is rare. This suggests that fear of large run totals should not be a strong factor in the decision to walk or pitch to Bonds.

WALKING BONDS TO PREVENT THE GIANTS FROM SCORING

When games are close in late or extra innings, the goal is to not allow runs. This goal motivates the primary question: does walking rather than pitching to Barry Bonds reduce the chance that at least one run scores?

For some game situations, baseball strategy dictates that walking Bonds is the smart choice. These include the situations in which first base is unoccupied and there are runners in scoring positions. Consider the situations on an out-by-out basis. With two outs, walking Bonds is preferable since he has a higher batting

Table 1. Characteristics of Bonds' plate appearances (walks/appearances)

	YEAR	0 ON	1B	2B	3B	1B+2B	1B+3B	2B+3B	LOADED
0 Out	2001	20/111	3/30	1/9	0/1	1/14	1/7	1/3	0/1
	2002	32/116	7/25	2/4	1/2	2/9	1/3	2/2	0/3
	2003	27/150	4/23	2/5	0/0	4/14	0/0	4/4	0/4
1 Out	2001	17/88	19/84	12/23	3/12	5/15	2/5	2/2	0/3
	2002	22/86	13/45	15/21	12/13	4/18	2/7	4/4	0/4
	2003	12/60	11/41	5/10	1/1	2/19	4/8	9/10	1/5
2 Out	2001	41/151	14/34	13/18	5/7	9/15	4/9	1/1	1/5
	2002	36/117	16/48	14/21	1/6	5/15	7/11	1/1	0/4
	2003	15/76	17/52	16/26	8/11	9/17	1/3	5/6	0/2

average than those players batting immediately after him. That is, Bonds is more likely to get a hit, thereby driving home the run, than those other players. With one out, walking Bonds sets up a double play that potentially can end the inning without any runs scoring. With zero outs, walking Bonds sets up force plays that prevent the lead runner from advancing further.

For other situations, it is clear that a walk is not the smart choice. When the bases are loaded, a walk automatically gives the Giants a run. With runners on first and second and less than two outs, a walk advances a runner to third base, where he can score on a fly out or a well-placed ground out.

For the remaining situations, it is not clear from baseball strategy whether walking or pitching to Bonds is the smart choice. Hence, it is useful to examine data for evidence of the success of one strategy over the other. This can be done by comparing the percentages of innings in which the Giants score at least one run when Bonds walks versus when he bats. This comparison is displayed in Figure 2 for the None On and First Only situations.

For comparisons, all three years of data are combined in single percentages. Pooling the data across years simplifies comparisons of the strategies. Additionally, the combined percentages are based on larger numbers of innings than the annual percentages, which improves our ability to differentiate the effectiveness of the strategies. A drawback to pooling the data is that it masks any differences across years.

At first glance, the combined percentages suggest competitive advantages for each strategy. With none on and at least one out, walking Bonds seems more effective than pitching to him. This may be because in these situations, the risk that Bonds hits a home run outweighs the risk that he scores when put on first base. With none on and no outs, walking Bonds seems less effective than pitching to him. This suggests that avoiding Bonds' home run power is outweighed by beginning an inning with a free pass. With one man on base, pitching to Bonds seems to be the better strategy. The walk advances the runner on first to scoring position, and the risk of that runner scoring may outweigh the risk of Bonds driving in the runner from first.

These percentages are based on a limited number of plate appearances. Suppose there is no difference in the true probabilities of the Giants scoring when

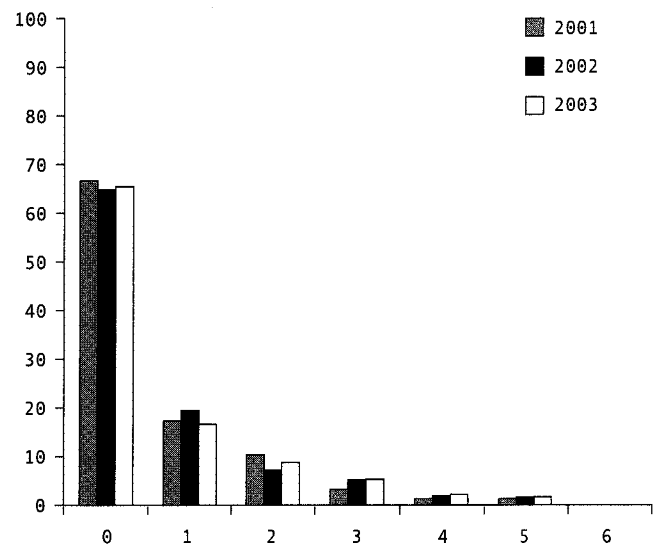


Figure 1. Percentages of number of runs in innings when Bonds appears at the plate

Bonds walks or hits. Could these apparent differences be plausibly explained by random chance? To answer this question, we conceive of a hypothetical population of Bonds' plate appearances under the same conditions that existed in 2001 through 2003, and we consider the plate appearances in 2001, 2002, and 2003 a random sample from this hypothetical population. Under this framework, the answer to our question is "not likely" for some situations and "entirely plausible" for others. When we combine the three years of data, the p-values for two-tailed statistical hypothesis tests are small for None On Zero Outs (p-value = 0.08), for None On One Out (p-value = .04), and for First Only Two Outs (p-value = .04). Hence, if walking and pitching to Bonds are equally effective, in these three situations we expect to see differences in the combined percentages as large (or larger) than those in 2001 through 2003 at most 8% of the time, indicating chance error is not a likely explanation of these differences. The p-values associated with the tests for None On Two Outs and for First Only One Out are both much greater than 0.10, indicating walking and pitching to Bonds in these situations could be equally effective. There is little data for First Only Zero Outs, although a simple examination of the differences suggests pitching to Bonds is a better option than walking him in that scenario.

Before accepting these conclusions, we should make sure that our comparisons are fair to both strategies. That is, up to the point when Bonds steps to the plate, the innings in which he walks should have similar characteristics to those in which he hits. In general, when comparing two strategies, fairness can be assured by assigning the strategies to the experimental units at random. We do not have this setup when analyzing the data. Decisions to walk or pitch to Bonds were made by managers rather than at random. Hence, we have to check whether the walk-innings differ from the hit-innings in ways that could affect the chances of scoring runs.

Other than game situation, the primary variable that affects the number of runs scored is the opposing pitcher's quality. Weak pitchers are likely to give up more runs than strong pitchers, regardless of whether they walk or pitch to Bonds. A pitcher's quality can be measured by his earned run average (ERA) over his career, which roughly equals the total number of all runs allowed by the pitcher divided by the number of innings he has pitched. Based on investigations within each game situation, the distributions of ERA are similar for the innings Bonds walks and innings he does not walk. An example of this similarity is displayed in Figure 3, which shows distributions of opposing pitchers' ERAs for None On situations in 2003. The overlap in the distributions for walk-innings and hit-innings is reproduced in other game situations and years. Hence, any effects of ERA on runs are approximately equally present in the side-by-side percentages of Figure 2. The comparison is fair with respect to ERA.

Another potentially important factor in the decision to walk or face Bonds is the quality of the player hitting after Bonds. In roughly 85% of his at bats in 2001, Bonds was followed by Jeff Kent, and walks were similarly distributed in games when Kent or someone else batted after Bonds. In 2002, Kent followed in roughly 52% of plate appearances; Benito Santiago followed in roughly 28% of plate appearances; Reggie Sanders followed in roughly 18% of plate appearances; and, other players accounted for the remaining 2%. Bonds walked in roughly 33% of his at bats when followed by Kent, 33% when followed by Santiago, and 37% when followed by Sanders. In 2003, Bonds was followed by Edgardo Alfonso in roughly 28% of plate appearances, by Jose Cruz Jr. in roughly 30% of appearances, by Benito Santiago in roughly 29% of appearances, and by a mix of others in 13% of appearances. Bonds walked 25% of the time when followed by Alfonso, 31% of the time when followed by Cruz Jr., 29% of the time when followed by Santiago, and 30% of the time when followed by others. Based on the three years of data, we can conclude that Bonds was walked with similar frequency regardless of who was on deck. Hence, comparisons of walk-innings and hit-innings within each game situation should not be affected greatly by differences in the players batting after Bonds.

Other variables examined include whether the game is in San Francisco or at other stadiums, the score, the inning, and the game number in the season. For these variables, there are one or two game situations for which the variables' distributions differ in the walk-innings and the hit-innings. These differences should have minimal effect on the comparisons, because these variables do not have strong relationships with the probability of scoring runs. Players of Bonds' caliber, and those who bat after him, try equally hard to score

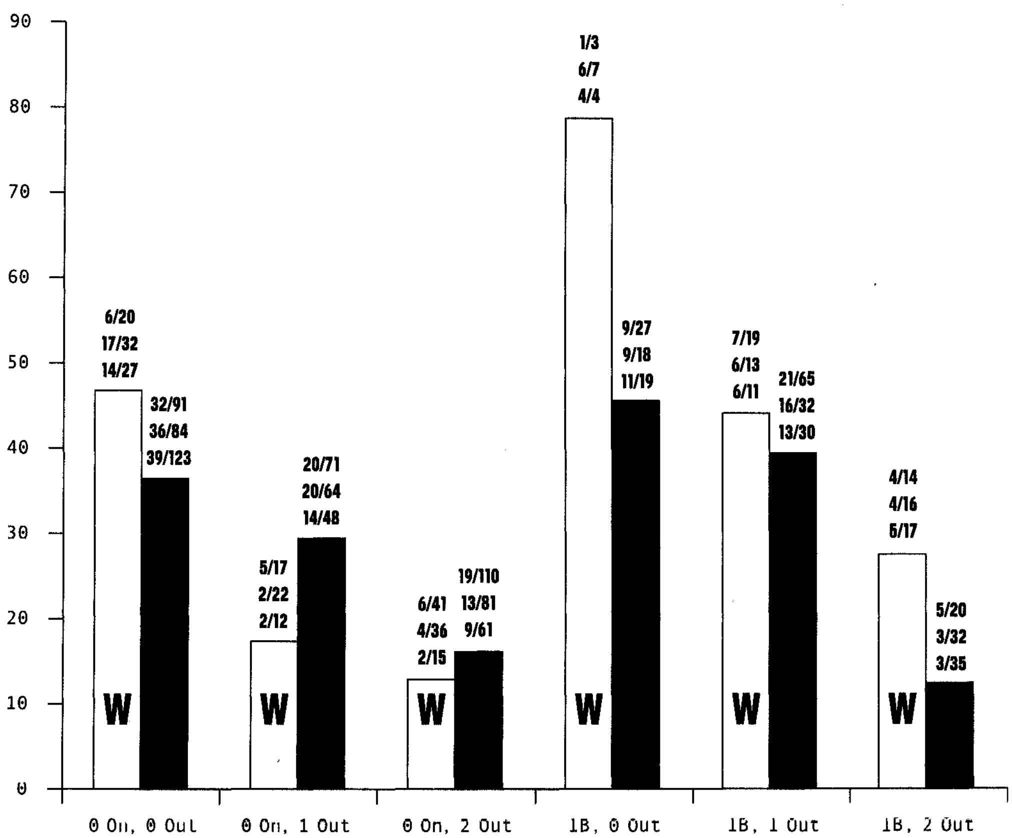


Figure 2. Percentage of innings the Giants scored at least one run after the first pitch to Bonds. Bars with a W show innings in which Bonds walked. Within each situation, the left bar shows the percentage when Bonds is walked, and the right bar shows the percentage when Bonds is pitched to. Each bar represents the combined percentage obtained by pooling the three years of data. The annual percentages are above the bars, going from 2001 at the top to 2003 at the bottom. For example, in 2001 the Giants scored in six of the 20 innings in which Bonds was walked with none on and no outs; in 2002 they scored in 17 of 32 such innings; and, in 2003 they scored in 14 of 27 such innings. Thus, the Giants scored when Bonds was walked with none on and no outs a combined (6+17+14)/(20+32+27) = 47% of the time.

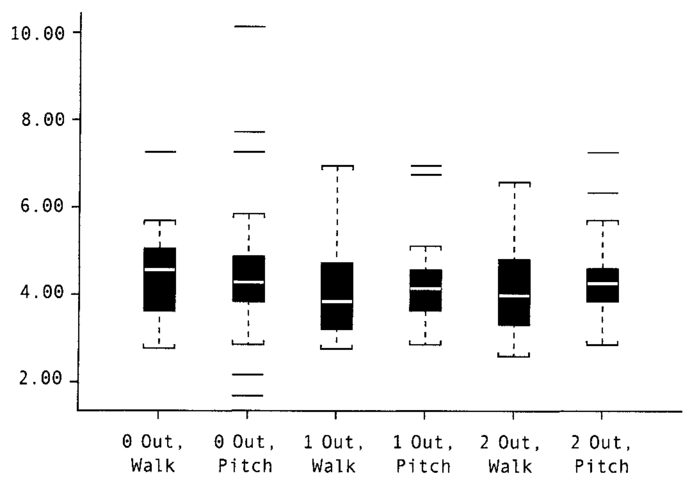


Figure 3. ERA comparisons for None-On Scenario, 2003

runs regardless of the stage of the game and the time of the season.

Hence, the conclusions stand. When the objective is to prevent any runs from scoring, the data suggest walking Bonds may be preferable when there is no one on and at least one out, and pitching to him may be preferable when there is a runner on first base or when Bonds leads off an inning.

DOES WALKING BONDS HELP AVOID BIG INNINGS?

Does walking Bonds have an effect on the chance of scoring runs in general? That is, do we expect bigger run totals when Bonds is walked as opposed to when he hits? We next use the data to investigate this question for the None On and First Only game situations.

Let the variable y represent the number of runs scored in an inning. For example, y=2 means two runs are scored. Ideally, we'd know the probability the Giants score y runs for each value of y in each game situation under both strategies, so we could compare the two strategies in any situation by comparing their probabilities. Of course, we don't know these probabilities, so we use the data to learn about them.

As Figure 1 shows, it is relatively rare that y>3. In fact, y>3 for only about 1% of all innings in the None On situation and about 3% of all innings in the First Only situation. This suggests that we can simplify analyses without losing much information by collapsing runs into four categories: y=0, y=1, y=2, and $y\geq3$.

Natural estimates of the probabilities for these four run categories are the proportions of runs that fall in each category in Bonds' combined plate appearances from 2001, 2002, and 2003. These proportions are displayed in Table 2. The table also includes the averages and standard deviations of the number of runs. For positive run categories, there is an interesting trend in the data. Let x be the number of runners on base when Bonds steps to the plate. When Bonds hits, the Giants are more likely to score exactly (x+1) runs than when he walks. On the other hand, when Bonds walks, the Giants are more likely to score (x+2) or more runs than when he hits. The averages of runs are typically as small or even smaller when Bonds hits than when he walks, the one exception being when the bases are empty with one out. Overall, these patterns suggest that it might be preferable to pitch to Bonds rather than to walk him in the None On and First Only

Table 2. Total innings (N), proportions for run categories, and averages and standard deviations of runs scored, for 2001, 2002, 2003, and all three years combined

		NO WALK						WALK					
SITUATION	**YEAR**	**N**	**0**	**1**	**2**	**3+**	**AVG (SD)**	**N**	**0**	**1**	**2**	**3+**	**AVG (SD)**
0 On, 0 out	2001	91	.65	.25	.09	.01	.49 (0.9)	20	.70	.20	.10	.00	.40 (0.7)
	2002	84	.57	.25	.11	.07	.71 (1.1)	32	.47	.22	.22	.09	1.03 (1.3)
	2003	123	.68	.20	.07	.05	.50 (0.9)	27	.48	.15	.26	.11	1.04 (1.2)
	Comb.	298	.64	.23	.09	.04	.56 (1.0)	79	.53	.19	.20	.08	.87 (1.1)
0 On, 1 out	2001	71	.72	.21	.06	.01	.39 (0.9)	17	.70	.12	.12	.06	.53 (0.7)
	2002	64	.69	.23	.03	.05	.48 (0.9)	22	.91	.05	.00	.04	.18 (0.7)
	2003	48	.71	.25	.02	.02	.35 (0.6)	12	.83	.00	.17	.00	.33 (0.8)
	Comb.	183	.71	.23	.044	.02	.41 (0.8)	51	.82	.06	.08	.04	.33 (0.8)
0 On, 2 out	2001	110	.83	.17	.00	.00	.17 (0.5)	41	.85	.05	.08	.02	.26 (0.9)
	2002	81	.84	.12	.04	.00	.20 (0.5)	36	.89	.05	.03	.03	.25 (0.9)
	2003	61	.85	.11	.02	.02	.20 (0.5)	15	.87	.00	.07	.06	.27 (0.8)
	Comb.	252	.84	.14	.02	.01	.19 (0.5)	92	.87	.04	.06	.03	.26 (0.8)
1B, 0 out	2001	27	.67	.07	.15	.11	.74 (1.2)	3	.67	.00	.00	.33	2.00 (3.5)
	2002	18	.50	.33	.06	.11	.83 (1.1)	7	.14	.43	.14	.29	1.86 (1.7)
	2003	19	.42	.26	.21	.11	1.68 (1.7)	4	.00	.50	.25	.25	1.75 (1.0)
	Comb.	64	.55	.20	.14	.11	1.04 (1.4)	14	.21	.36	.14	.29	1.86 (1.8)
1B, 1 out	2001	65	.68	.14	.15	.03	.72 (0.9)	19	.63	.16	.05	.16	.79 (1.3)
	2002	32	.50	.25	.12	.13	1.00 (1.4)	13	.54	.23	.08	.15	.85 (1.1)
	2003	30	.57	.17	.13	.13	1.00 (1.5)	11	.46	.18	.18	.18	1.09 (1.2)
	Comb.	127	.61	.18	.13	.08	.86 (1.2)	43	.56	.19	.09	.16	.88 (1.2)
1B, 2 out	2001	20	.75	.00	.20	.05	.55 (1.0)	14	.71	.07	.07	.15	.78 (1.5)
	2002	32	.91	.06	.03	.00	.13 (0.4)	16	.75	.13	.06	.06	.56 (1.2)
	2003	35	.91	.06	.03	.00	.11 (0.4)	17	.70	.12	.06	.12	.64 (1.2)
	Comb.	87	.87	.05	.07	.01	.22 (0.6)	47	.72	.11	.06	.11	.65 (1.2)

situations, although opposing teams face risks when using either strategy.

As before, we should consider random variation when interpreting these sample proportions and sample averages. Suppose there is no difference in the effectiveness of walking or pitching to Barry Bonds in reality. Could the differences between walk-innings and hit-innings observed in 2001-2003 be plausibly explained by random chance? Let's again conceive of Bonds' combined 2001-2003 plate appearances as a random sample from a hypothetical population of his plate appearances under the current conditions in the league. We seek to learn about the differences in average runs in this hypothetical population when walking versus pitching to Bonds. For the game situations (i) a runner on first only and zero outs and (ii) a runner on first only and two outs, the p-values for two-tailed statistical hypothesis tests are both around 0.02, small enough values to cast doubt on chance error as an explanation for the differences in the sample averages in these situations. For these situations, the data provide evidence that favors pitching to Bonds. For the other game situations, the p-values of the two-tailed statistical hypothesis tests are all greater than .10, making it hard to rule out chance errors as explanations of the differences in the sample averages. For these other situations, there is not enough evidence to determine conclusively that one strategy results in fewer runs on average than the other strategy does.

CONCLUDING REMARKS

There have been an incredible number of walks issued to Barry Bonds in the last three years. Given his prodigious home run power, it is understandable why managers fear pitching to him. However, the data from 2001 through 2003 suggest that there is little difference in opposing teams' ability to prevent runs when walking Bonds versus when letting him hit. In fact, the data suggest that it may be better to pitch to Bonds than to walk him in some game situations.

ACKNOWLEDGMENTS

The author thanks David Tung, John Lee, and Jonathan Bigelow for assistance with collecting the data.

Best Ten-Year Performers

Lou Gehrig, born 101 seasons ago, would never have made the claim, but let's credit him with the best 10-season batting record in major league history if you don't mind.

It happened from 1927 through 1936, when baseball offense had a heyday. In a compilation of the finest 10-season performers from 1901 through 2003, Gehrig places in seven of 10 position player areas and is first in three of them.

Searching for consistency at a high level, the study considered those who played for 10 seasons in the American and/or National leagues within the period listed. Years missed were mostly due to military service. The focus was on the top 10 players rather than performances where, for instance, Babe Ruth holds the 12 best slugging percentages over 10-year periods: 1920-29, 1919-28, etc.

But back to Gehrig's stunning decade. He compiled 10-year bests in runs, runs batted in, and total bases during the period. His slugging percentage is third best for a player in ten years of play, and he's sixth on the hit list and ninth in both home runs and doubles.

During the period the New York Yankee first baseman hit 20 of his career record 23 grand slams. In 1932 he was the first major leaguer of the 20th century to hit four home runs in a game. Only in doubles did he not have his best 10-year stretch in 1927-36.

Imagine averaging more than 141 runs scored, 152 RBI, and 381 total bases for 10 consecutive years of play. In addition, Gehrig batted .350 (12th best ever) and averaged 202 hits, 40 doubles and 39 home runs, and posted a .662 slugging mark. And as if all that weren't enough, Larrupin' Lou had 117 triples and 80 stolen bases over the decade.

Gehrig began and ended that 1927-36 time frame with a pair of American League Most Valuable Player awards, and he won the Triple Crown in 1934. Amazingly, Gehrig accomplished much of this while following one of history's greatest base cleaners in the lineup, Babe Ruth. Gehrig's remarkable RBI average for 1927-36—152 plus—would have won the American League crown every year from 1949-1997. From 1941 through 1996 only Jim Rice (1977-78) and Don Mattingly (1986) won AL total base titles with more than Gehrig's 381 average. And since 1949 just one AL run-scoring leader—Rickey Henderson (146 in 1985)—has exceeded Gehrig's 141.7 average from 1927 to 1936.

Twenty-four times the record book shows Gehrig's numbers in bold type as the AL leader in some statistical category during the period. He led in RBI five times; runs, four times; home runs, three times; doubles and slugging, twice; and even batting once. He also once had a record-tying three triples in a game that was rained out before it became an official contest and thus wasn't recorded. For his career, Gehrig averaged the most runs and RBI per game of any 20th-century player.

Gehrig was only 33 years old when his remarkable decade ended in 1936. Seemingly, he would have at least a few more years of stardom. He had a good 1937 season, but in 1938 his batting average fell below .300 for the first time since 1925, and it was clear that something physical was wrong. He lacked his usual strength, and pitches he would have murdered were only flyouts.

After a feeble start in 1939, Gehrig, the Yankees' captain, presented the lineup card to the umpires for the ninth game of the season, but his own name wasn't on it. His streak of 2,130 consecutive games had ended, and shortly thereafter he was diagnosed as having a very rare degenerative disease: amyotrophic lateral sclerosis (now commonly called Lou Gehrig's disease). He would never play baseball again.

New York honored Gehrig with a day on July 4, 1939, and his teary words of thanks have become part

SCOTT NELSON, *when not working on sports or family history publications, can be found filming progress on the log home a daughter and husband are building next to the family lake home in Northern Minnesota.*

of baseball lore. Less than two years later, just days before his 38th birthday, Gehrig died.

The Iron Horse's marks and most of the others from the first 60 years of the past century are more notable, too, in light of baseball's switch from 154-game to 162-game schedules in 1962. In a 10-year period, the difference adds up to 80 games, nearly half of a current season.

RUTH DID OKAY, TOO

Teammate Babe Ruth, listed six times, has the best 10-season mark in slugging (.740), but was surpassed on the home run list by Sammy Sosa in 2003, 469 to 467. Sosa, a Chicago Cub the past 12 seasons, and Rogers Hornsby are the only batting leaders on these lists who played most of their games in the National League. And Hornsby exceeded Ruth in making seven of 10 hitter classifications, including a first in hits with 2,085 for four different teams in 1920-29. Jimmie Foxx made it in five and Ty Cobb, Al Simmons, and Stan Musial in four, while Paul Waner, Willie Mays, and Sosa do it in three areas.

Sosa, Barry Bonds, Ken Griffey Jr., and Rafael Palmeiro give active home run hitters their best showing in any batting category. Twenty-four entries have been made in the new century. Not surprisingly, six of the 10-year leaders are also career record holders: Cobb (batting average), Tris Speaker (doubles), Sam Crawford (triples), Ruth (slugging), Henderson (stolen bases), and Walter Johnson (earned run average).

However, Hank Aaron, the career record holder in total bases, RBI, and home runs, is only sixth, 11th, and 12th, respectively, among 10-year performers.

A pair of Johnsons and Christy Mathewson are pitching leaders in the study. Randy Johnson leads in winning percentage and strikeouts, while yesterday's Walter Johnson and Matty appear on three of five pitcher lists. Mathewson averaged nearly 28 wins a year from 1903 through 1912, with Walter not far behind.

With his 24-5 record in 2002, Randy Johnson leaped way out front with the best winning percentage among hurlers having at least 100 victories in a 10-season period. Dennis Eckersley holds the saves record.

When Randy Johnson passed Nolan Ryan on the 10-year strikeout list in 2002, it meant only Walter Johnson in this compilation is a career or season record holder in the same category among pitchers.

Eleven players active in 2003 are on the 10-season lists, led by leaders Henderson and Randy Johnson. Henderson's 838 stolen bases in 1980-89 give him a 20% edge over the next best record, the largest margin in our study. Henderson and Hank Aaron are the only players holding three major league career records--his are in stolen bases, runs, and walks.

Nearly half of those listed played their entire career with one club. In contrast, Henderson in his 25-year career has played for nine different teams, including Oakland (four times) and San Diego (twice). Because of the more recent focus on home run hitters and relief pitchers and the dearth of low ERAs since 1920, the study includes an extra list in each of those categories.

HIGHLIGHTS

Batting Average--Tony Gwynn only entry in Top 10 since Ted Williams 50 years ago.

Hits—Kirby Puckett is the last entrant, 10 years ago. Puckett fell four short of 2,000 when the 1994-5 strike wiped out 67 Twins games.

Doubles—Wade Boggs (1983-92) and Pete Rose (1971-80) are most recent entries. Rose is 10th and missed 400 by one.

Triples—Sam Crawford's 185 (1906-15) leads by a wide margin and is the oldest leader record among batters in the study. Paul Waner was last to make the list in 1935 in an all-but-forgotten hitting category.

Home Runs—Though Hank Aaron, Frank Robinson, Reggie Jackson, Mike Schmidt, and Mickey Mantle each hit more than 500 and are in the top 11 career-wise, none make the list in a 10-summer stretch.

Total Bases—Only 29 times has a batter totaled 400 or more, and Lou Gehrig did it on five occasions. No one else more than three.

Runs—Rickey Henderson and Ty Cobb are one-two career leaders, but neither one, in his 10 best years, crossed the plate enough for this list.

RBI—Though he had just 103 in the last season of his injury-shortened career (2000), Albert Belle is ninth among the best 10-year performers. Until Belle and Sosa came along, Aaron was last to qualify in 1966.

Slugging Average—Stan Musial missed .600 by the narrowest of margins with a .5995 average (1948-57).

Stolen Bases—Only two of the leaders made their mark before 1969.

Wins—Christy Mathewson won 278 games in 1903-1912, and that's held up as the oldest 10-year best. Nobody's made the Top 10 since Juan Marichal 33 years ago.

Percentage—Randy Johnson, Pedro Martinez, and Greg Maddux made the 10 best in this century. Mathewson is the lone entry on this list with more than 200 wins in a 10-year period.

ERA—How baseball has changed: Reliever Hoyt Wilhelm is the sole hurler listed since 1920. So, we also listed the best 10-year ERAs since then.

Saves—All made the list since 1992.

Strikeouts—All but Walter Johnson and Sandy Koufax reached their 10-season peak in the last 32 years.

SOURCES

Big League Baseball Electronic Encyclopedia, by Franklin Electronic Publishers.

McConnell, Bob and David Vincent, eds. *SABR Presents The Home Run Encyclopedia*; New York, NY, Macmillan Publishing, 1996.

The Sporting News Complete Baseball Record Book. 2003, St. Louis, MO., Sporting News Publishing, 2003.

STATS Inc. All-time Major League Sourcebook. James, Dewan, Zminda, Callis, Munro, Stats Publishing, 1999.

Thorn, John and Pete Palmer, eds. *Total Baseball*; New York, NY, Harpercollins, 1993.

Table 1. Best ERAs Since 1920. The letters (c) and (s) indicate player is ML career and/or single-season record holder since 1901. Some careers began before 1901.

STARTERS		
Maddux	1992-2001	2.46
Seaver	1968-77	2.46
Palmer	1967-76	2.49
W. Wood	1962-72†	2.52
Marichal	1960-69	2.57
Gibson	1964-73	2.58
P. Martinez	1994-2003	2.58
Messersmith	1966-75	2.65
Peters	1959-68	2.65
Tudor	1984-93	2.66
†Missed 1966 season.		
RELIEVERS (400 IP MINIMUM)		
Wilhelm	1961-70	2.10
Gossage	1977-86	2.27
J. Brewer	1965-74	2.45
Lyle	1968-77	2.45
Rivera	1995-2003	2.49
SAVES		
Eckersley	1988-97	370
Smith (c)	1986-95	358
Hoffman	1993-2002	352
Wetteland	1991-2000	329
Reardon	1983-92	315
Myers	1989-98	315
Nen	1993-02	314
Henke	1986-95	295
R.Hernandez	1993-02	308
Montgomery	1989-98	291
STRIKEOUTS		
R. Johnson	1993-2002	2,928
Ryan (c,s)	1972-81	2,756
Seaver	1969-78	2,381
Koufax	1957-66	2,336
P. Martinez	1994-03	2,299
Gibson	1962-71	2,295
McDowell	1964-73	2,253
Lolich	1965-74	2,245
W. Johnson	1909-18	2,236
Carlton	1974-83	2,225

Table 2. Top 10-year performers. The letters (c) and (s) indicate player is ML career and/or single-season record holder since 1901. Some careers began before 1901. An asterisk indicates that the player missed season(s) due to military service.

BATTING AVERAGE (3,000 AB)

Cobb (c)	1910-19	.387
Hornsby	1920-29	.382
Heilmann	1921-30	.367
Speaker	1916-25	.360
Simmons	1925-34	.359
J. Jackson	1911-20	.357
Ruth	1920-29	.355
Lajoie (s)	1901-10	.352
T.Williams	1941-53*	.351
G. Sisler	1917-27†	.351
T. Gwynn	1992-2001	.351

†Missed 1923 season

HITS

Hornsby	1920-29§	2,080
P. Waner	1927-36	2,074
Rose (c)	1968-77	2,067
Musial	1943-53*	2,056
G. Sisler (s)	1917-27§	2,040
Gehrig	1927-36	2,022
S. Rice	1921-30	2,010
Simmons	1925-34	2,005
Cobb	1908-17	2,001
Puckett	1984-93	1,996

§Missed 1923 season

DOUBLES

Speaker (c)	1918-27	446
Medwick	1933-42	441
Musial	1943-53*	432
Gehringer	1929-38	416
P. Waner	1927-36	414
Heilmann	1921-30	412
Boggs	1983-92	408
Hornsby	1920-29	405
Gehrig	1926-35	400
Rose	1971-80	399

TRIPLES

Crawford (c)	1906-15	185
Cobb	1908-17	164
J. Jackson	1911-20	163
P. Waner	1926-35	156
Combs	1925-34	150
Wagner	1903-12	143
Hornsby	1916-25	138
Konetchy	1908-17	138
Roush	1916-25	138
S. Rice	1921-30	137

HOME RUNS

Sosa	1994-2003	469
Ruth	1920-29	467
Barry Bonds (s)	1993-2002	437
Foxx	1930-39	415
McGwire	1990-99	405
Killebrew	1961-70	403
Griffey Jr.	1991-2000	400
Palmeiro	1994-2003	396
Gehrig	1927-36	390
Mays	1957-66	390

OTHERS WITH 350 OR MORE HOME RUNS

Aaron (c)	1962-71	386
Mathews	1953-62	374
Belle	1991-00	373
Thome	1994-2003	371
Mantle	1955-64	370
Schmidt	1974-83	370
Kiner	1946-55	369
Bagwell	1994-2003	366
Gonzalez	1992-2001	365
Banks	1955-64	355

RUNS

Gehrig	1927-36	1,417
Ruth (s)	1920-29	1,365
T. Williams	1939-51*	1,273
Foxx	1930-39	1,244
Gehringer	1929-38	1,224
Hornsby	1920-29	1,195
Musial	1944-54*	1,193
Mantle	1953-62	1,186
Mays	1955-64	1,184
Bagwell	1994-2003	1,160

RUNS BATTED IN

Gehrig	1927 36	1,527
Foxx	1929-38	1,415
Ruth	1923-32	1,377
Simmons	1925-34	1,277
J. DiMaggio	1936-48*	1,277
T. Williams	1939-51*	1,261
Sosa	1994-2003	1,216
Ott	1929-38	1,206
Belle	1991-2000	1,199
Greenberg	1933-46*	1.190

TOTAL BASES

Gehrig	1927-36	3,815
Ruth (s)	1920-29	3,613
Foxx	1930-39	3,580
Mays	1954-63	3,525
Musial	1946-55	3,508
Aaron (c)	1955-64	3,483
Hornsby	1920-29	3,470
Sosa	1993-2002	3,344
J. DiMaggio	1936-48*	3,304
Simmons	1925-34	3,290

SLUGGING AVG. (3,000 AB)

Ruth (c)	1920-29	.740
Bonds (s)	1994-03	.670
Gehrig	1927-36	.662
Foxx	1930-39	.652
T. Williams	1941-53*	.649
Hornsby	1921-30	.642
McGwire	1991-2000	.641
Greenberg	1934-46*	.631
L. Walker	1994-2003	.615
Mantle	1955-64	.614

STOLEN BASES

Henderson(c,s)	1980-89	838
Coleman	1985-94	698
Brock	1965-74	670
Cobb	1908-17	630
Raines	1981-90	627
Wills	1960-69	535
E. Collins	1909-18	523
W. Wilson	1978-87	521
Morgan	1969-78	509
O. Nixon	1988-97	498

WINS

Mathewson	1903-12	278-109	.718
W. Johnson	1910-19	265-143	.649
Alexander	1911-20	235-114	.673
Young (c)	1901-10	218-137	.614
Grove	1927-36	217-86	.715
Plank	1903-12	214-117	.647
Spahn	1953-62	205-118	.635
Marichal	1962-71	202-97	.676
Feller	1939-51*	199-110	.644
Roberts	1949-58	199-147	.575

WINNING PCT. (MIN.100 WINS)

R. Johnson	1993-2002	175-58	.751
Chandler	1937-46	100-38	.725
Grove	1930-39	199-76	.724
P.Martinez(c)	1994-2003	156-61	.719
Mathewson	1904-13	273-107	.718
Ford	1950-61*	158-63	.715
Leever	1901-10	158-64	.712
Maddux	1993-2002	178-77	.698
Guidry	1976-85	154-67	.697

EARNED RUN AVERAGE (MIN.1000 IP)

W. Johnson (c)	1910-19	1.59
Walsh	1907-16	1.74
M.Brown	1903-12	1.82
Mathewson	1904-13	1.87
J. Wood	1908-19†	2.00
Young	1901-10	2.05
Alexander	1911-20	2.07
Wilhelm	1961-70	2.10
Waddell	1901-10	2.11
Bender	1905-14	2.17

†Missed two seasons

TOM HANRAHAN

Highest Future Value

The good hot stove question is: What ballplayer, at the peak of his career, could you most solidly say, "I wouldn't trade him for *anybody*"?

Stated another way, if we take all major leaguers and track their careers as they progress (with no future knowledge), which one would have been perceived as having the highest future value (or value in a potential trade) of all time. It would have to be a player early enough in his career that he still had many years of service remaining, but also far enough into his career that he had established a high level of performance.

The metric I used to answer this question was the Bill Jamesian uber-stat: Win Shares. In fact, this is precisely the type of query that Win Shares is qualified to answer. As a measure, it is comprehensive, easy to use, and well accepted in sabermetric circles.

For those unfamiliar with Win Shares (WS), it is an attempt to take all of a player's contributions in a season and assign them a value in terms of how many wins he contributed to his team. Offensive rates (on-base average, slugging, stolen base success), playing time, fielding stats, defensive positions, and clutch play are all taken into account. The denomination of WS is actually "wins divided by 3". As examples, Babe Ruth has the highest total of WS ever; with 758 WS, he was worth 758 ÷ 3 = 253 wins in his career to his teams, above a poor replacement outfielder/pitcher. In 2001, when Barry Bonds hit 73 home runs and set many other records, he collected 54 WS, which is one of the highest single-season totals ever. A complete season-by-season listing of the WS all of these players earned can be found in the book entitled *Win Shares*.

To measure a player's current value or what I will call established value (EV), I used a weighted average of WS over the most recent four years: 40% of the current year, 30% the previous year, 20% of the year before that, and 10% in year minus three. I then researched every hitter I could find who had strong seasons at young ages to see who had the highest EV at a given age. For this exercise, I did not compare pitchers with hitters; I reasoned that the comparatively larger number of career-threatening injuries to young pitchers would never allow a bright shining light such as a young Doc Gooden to be thought of as even trade bait for an equally young and brilliant Alex Rodriguez. So, my going-in assumption was that no pitcher was ever perceived as so good *and* such a horse *and* so injury-free that he might compare with a young, established superstar hitter.

Many players, even future Hall of Famers, did not play in the majors prior to age 20 or 21. Since a player can easily earn 10 to 15 WS in a season by playing full-time even when not playing very well, I decided to assign a minimum number of WS to any early missing player-seasons, assuming that if the major league club had called up the player, he would have performed at a certain level. This keeps the early list from being dominated by those players who were fortunate to have been brought up to the big leagues at 18 or 19. It still does not adequately make up for those who served in the military at an early age; Willie Mays may have established one of the top EVs at ages 23 through 25 if he hadn't missed time. The minimum assigned WS are as follows: 6 at age 18, 10 at age 19, 12 at age 20, and 14 at age 21 and following.

Table 1 shows the WS earned by age for each player who had one of the six highest established values at some point prior to age 27. There were some players who played in the majors at age 18, but since none of them in this table earned more than 6 WS at that age, the "age 18" column was dropped. Other players did earn more than 6 WS at age 18, but they did not become big enough stars later on, and so did not factor into the analysis. Nineteenth-century performances were not considered.

Traditionally, a player's baseball age is how old he was on July 1 of the year. Later in the analysis, I will

TOM HANRAHAN *is an avionics systems engineer at the Naval Air Warfare Center in Maryland. He used to have other hobbies until baseball research and his family absorbed his time.*

note which players are hurt or helped by this arbitrary date, and who else might have had the highest EV if another date was used.

Example: Rogers Hornsby. He did not play MLB at age 18, and earned no WS for his few games at age 19. So, at age 19, his Established Value is computed as

$$EV(19) = (.4 \times 10) + (.3 \times 6) = 5.8 \text{ WS}$$

Where 10 and 6 are the minimum WS assigned for ages 18 and 19.

At age 20 (the year 1916), Hornsby suddenly blossomed. He was one of the best hitters in the league, finishing fourth in batting, slugging more than 100 points higher than any of his teammates, and split time between SS and 3B. For this he earned 28 WS. His EV at age 20 is $EV(20) = .4\times28 + .3\times10 + .2\times6 = 15.4$.

At age 21, Hornsby was the St. Louis Cardinals full-time shortstop, and led the league in slugging, total bases, OPS, and triples. He earned 37 WS. His $EV(21) = .4\times37 + .3\times28 + .2\times10 + .1\times6 = 25.8$. This is the second highest EV(21) ever.

To put 25.8 WS in perspective, Bill James has written that 30 WS is an "MVP candidate season." So here we have a 21-year-old who has established that he is already almost a consistent MVP candidate. Probably not someone you are going to trade for most any other player.

Table 2 shows the six highest EV at each age from 20 through 26. No attempts have been made to attempt to adjust for the vast differences in playing conditions, quality of play, or schedule. It should be noted that only four of the 15 players mentioned in the table were youngsters in the post-integration era (Mickey Mantle, Henry Aaron, Al Kaline, and Alex Rodriguez). One player who is not on listed on the table, but is on a rapid ascent toward young greatness, is Albert Pujols. After the 2003 season, he has an EV(23) of 31.8, having earned 29, 32, and 41 WS in his first three seasons. Another season like the others will put Mr. Pujols in elite company.

The players listed in italics are those whose birthdays were in the months of January through June. They are on the young side of each year in question. Ty Cobb has the highest EV(21) of 30.5 WS, but Ty was really an "old" 21; he turned 22 shortly after the season ended. Mel Ott (born in March) was actually about

Table 1. Win Shares by age for top young stars

PLAYER	19	20	21	22	23	24	25	26
Aaron, H		13	29	30	35	31	38	
Cobb, T	16	41	36	44	45	47	40	31
Collins, E			11	43	39	35		
DiMaggio, J			25	39	30	34	31	41
Foxx, J	6	22	34	34	24	40	41	
Hornsby, R		28	38	20	26	38	41	47
Kaline, A	7	31	26	20				
Magee, S	11	28	31	37	26			
Mantle, M	13	32	26	36	41	49	51	39
Mays, W		19	5	0	40	40	27	34
Musial, S		3	28	39	38	0	44	
Ott, M	20	31	28	26	33	31	38	
Rodriguez, A	2	34	22	30	23	37	37	35
Ruth, B		23	37	36	44	43	51	53
Vaughan, A		21	34	36	39	35		
Williams, T		32	30	42	46	0	0	0

Note: I added 10% to the WS earned by players in the war-shortened 1918 season (Ruth and Hornsby). Blanks in the right side of the table do not mean the player did not play in those years; only that the WS the player earned did not improve his EV.

Table 2. Highest EV at each age

PLAYER	20 YRS	PLAYER	21 YRS	PLAYER	22 YRS	PLAYER	23 YRS	PLAYER	24 YRS	PLAYER	25 YRS	PLAYER	26 YRS
Cobb	22.4	Cobb	30.5	Cobb	38.2	Cobb	42.5	Cobb	44.7	Mantle	46.9	Ruth	49.5
Ott	19.6	Hornsby	25.8	Williams	33.2	Williams	40.2	Mantle	41.7	Ruth	45.8	Mantle	44.8
Mantle	17.9	Ott	25.0	Ruth	31.5	Ruth	38.4	Ruth	41.5	Cobb	43.5	Hornsby	41.3
A Rod	17.8	Ruth	24.3	Magee	30.9	Mantle	35.6	Vaughan	36.3	Foxx	36.6	Cobb	38.3
Williams	17.0	Williams	24.2	Vaughan	29.8	Vaughan	35.3	Collins	35.7	Hornsby	35.0	DiMaggio	35.5
Kaline	16.6	Magee	23.6	Mantle	29.8	Musial	33.7	Foxx	33.4	Aaron	34.8	A Rod	34.8
AVG	18.6		25.6		32.3		37.6		38.9		40.4		40.7

three months younger than Cobb at the same "age." If I had chosen January 1 instead of July 1 as the cutoff date, Ott would be at the top of the age 21 group.

Table 2 shows a sharp rise in EVs in general from ages 20 through 23, and then small increases after that. Babe's EV(26) of 49.5 EV was the highest he ever reached. It is not the highest EV of all time, however. Honus Wagner earned 46, 43, 44, and 59 WS from ages 31 through 34, giving him an EV(34) of an even 50.0. The highest "old player" EV is owned by Barry Bonds, who after the 2003 season has an EV(38) of 44.3.

How should we turn EV into "future value"? Future value should simply be a function of EV and age. I say simply because that is only two parameters, but the function would be difficult to agree on. Extensive research would have to be done on MLB player success and the aging process. Rather than attempting to create one grand number that attempts to definitively answer the question, I would encourage the reader to eyeball the figures in Table 2 with me.

As good as the young Mel Ott was at ages 20 and 21, it does not seem reasonable to conclude that a manager would refuse hypothetically to trade him for one of the vastly superior players who were only a few years older.

The real battle is between the young Ty Cobb, ages 22 through 24, and the more established Babe Ruth. The peak in Cobb's EV curve occurs at age 22. He had just won three consecutive batting titles (1907 through 1909), as well as having led the league in slugging, hits, total bases, and RBI three years in a row. In contrast, the Babe continued to add significantly to his established value by whacking 50+ home runs (more than most teams) and shattering numerous other records at ages 25 and 26 (1920 and 1921). How should we compare Cobb's EV(22) of 38.2 with Ruth's EV(26) of 49.5, remembering that Ruth's birthday was two months after Cobb's, giving him a small age disadvantage?

Advantage for the Babe: His EV(26) is 30% higher than Ty's EV(22).

Advantages for Ty: He was almost four years younger, which meant likely a longer career ahead of him, plus much more room for the growth that a 22-year-old player would expect to have.

It is clear to me that the combination of these two factors is larger than the difference in their established values.

CONCLUSION

If Win Shares were the perfect value-measurement tool, and *if* there were no subjective factors to consider, and *if* no adjustments need to be made for changes in quality of play over time, then Tyrus Raymond Cobb, age 22 in the fall of 1909, would appear to have greatest future value than any player at any point in major league history. Conclusions relevant to baseball today: Alex Rodriguez has been the best young player to grace the major leagues in at least the past 40 years—but watch out for Albert Pujols.

RON SELTER

Sportsman's Park's Right-Field Pavilion and Screen

The first ballpark on the site (Grand Boulevard and Dodier Street) of what would later be known as Sportsman's Park was Grand Avenue Grounds. That park was first used for major league baseball as the home of the American Association St. Louis Browns in 1875. The next ballpark on this site, Sportsman's Park I, was used in the 1882–92 time period by the St. Louis franchises in the American Association and the NL. During its seasons of use by the American Association St. Louis Browns, the park was noted for its Beer Garden in RF. Until the 1888 World Series the Beer Garden, in addition to being popular, was also in play!

The AL had a franchise in Milwaukee during their 1901 inaugural season. The franchise, upon moving to St. Louis for the 1902 season, became the St. Louis Browns and acquired the site of Sportsman's Park I. For their first season in St. Louis, the Browns built a new wooden ballpark (denoted as Sportsman's Park II), replacing the existing rundown stands with a new grandstand and bleachers. The field was oriented with home plate in the northwest corner of the plat (or land parcel). The stands consisted of (1) a single deck covered grandstand that curved behind home plate and stretched between first base and third base, (2) bleachers that ran along the RF and LF foul lines and converged with the foul lines in the LF and RF corners, and (3) a separate set of uncovered wooden bleachers in LF that ran from the LF foul line nearly to the clubhouse that was in CF. The RF fence was parallel to the LF foul line and ran until meeting the clubhouse in CF. There was no seating beyond the RF fence—in fact this area was a peach orchard and not part of the ballpark at all.

The playing field (fair territory) was nearly exactly rectangular in shape—the only exception being the diagonal CF scoreboard, which masked most of the clubhouse from the playing field. The estimated dimensions of the park were: LF 330, CF at the scoreboard 430, RF 315, and the backstop 60 ft from home plate.[1]

In its third incarnation (Sportsman's Park III) the ballpark was drastically altered after the 1908 season. The playing field was reoriented, with home plate and the infield now in the southwest portion of the site. In addition, the southern boundary of the plat was extended towards Dodier Street. The site now amounted to 6.3 acres—roughly typical in size for the classic era ballparks. The Browns used this extra area on the south to build a new state-of-the-art (for 1909) double-deck steel and concrete grandstand. This involved removing the prior park's 1B bleachers to make room for the new grandstand. Never ones to waste money, the Browns' management retained both the former 3B and LF wooden bleachers as well as the curved covered grandstand (which had faced the infield) from its previous layout. The former LF bleachers became the RF bleachers in the new park, the former curved grandstand became the 3B-LF pavilion, and the former 3B bleachers became the new LF bleachers. The LF bleachers were angled at less than 90° to the LF foul line.

For the 1912 season, the curved pavilion in LF and the angled LF bleachers that were carryovers from Sportsman's II were replaced by a new 3B pavilion and a single rectangular set of LF bleachers. At the same time a new 1B pavilion was built. Now the LF fence and bleachers were 90° to the LF foul line. Contemporary newspaper game accounts contain references to a roof in RF during the 1912 season. In the first game of a doubleheader played on September 27, 1912, and again in the second game of another doubleheader the next day, Gus Williams of the Browns hit homers onto the roof in RF.[2] However, several other game accounts during the 1912 season mention homers hit into the RF bleachers. One question immediately comes to mind—how could homers be hit both into the (uncovered) RF bleachers and onto the roof in RF? SABR's Robert Tiemann supplied the answer.[3] The 1B pavilion

RON SELTER *is an economist who lives in El Segundo, California. A member of SABR since 1989, he has done research on the minor leagues and on ballparks.*

(which, in the correct use of the term pavilion, was roofed) built for the 1912 season faced towards second base and extended about 15 ft. into fair territory near the 1911 RF corner. Thus, while a small portion of the stands in RF was roofed, the RF bleachers, that accounted for the large majority of the RF seating, remained uncovered.

In RF the intrusion of the 1B pavilion reduced the estimated RF distance to 295 feet. The RF bleachers remained at 90° to an extension of the RF foul line. The 295-ft. RF foul line distance is misleading, as the distance in RF at a point 3° from the foul line was a more substantial 325 feet.

During the Deadball Era (1901-19) in both Sportsman's II and Sportsman's III, home runs to RF were nearly all of the over-the-fence type, including bounce home runs. Inside-the-park home runs, while common to CF, were rare to RF.

The next change to the RF stands occurred in the off-season of 1925-26. The principal change to the park was the extension of the double-deck grandstand down the LF and RF lines to the foul poles, replacing the 1B and 3B pavilions. In the same 1925-26 expansion, the outfield wooden bleachers in LF, CF, and RF were rebuilt in steel and concrete. The RF stands, now 40 feet deep and 300 feet wide at the back with a seating capacity of 3,290, were completely covered. The RF stands, now being completely roofed, were thereafter referred to as the RF pavilion, and no longer as bleachers. This 1925-26 renovation also resulted in the RF foul line having a marked distance of 310 feet (actually 309½ ft.).[4]

During the 1929 season, the visiting Detroit Tigers hit eight home runs in a four-game series (July 2-4). Surprisingly, the Browns won three of the four games. Nonetheless, in an effort to help the Browns shell-shocked pitching staff, the Browns management used the off-day of July 5 to install a 21½-ft. screen in RF, placed above the existing wall. The screen was in place for the game of July 6 against the Yankees.

The screen was in play and raised the barrier to RF homers from 11.5 to 33 feet. The screen ran from the foul line to about right center—near the 354-ft. mark, and covered nearly all of right field. The Browns' calculation of home park advantage was apparently simple—the team had few LHB (only two regulars—Manush and McGowan) and no power hitters (LH or RH). The other AL teams had both more and better LHB and power hitters. The NL Cardinals, as tenants of the Browns', were not a party to the decision to install the screen.

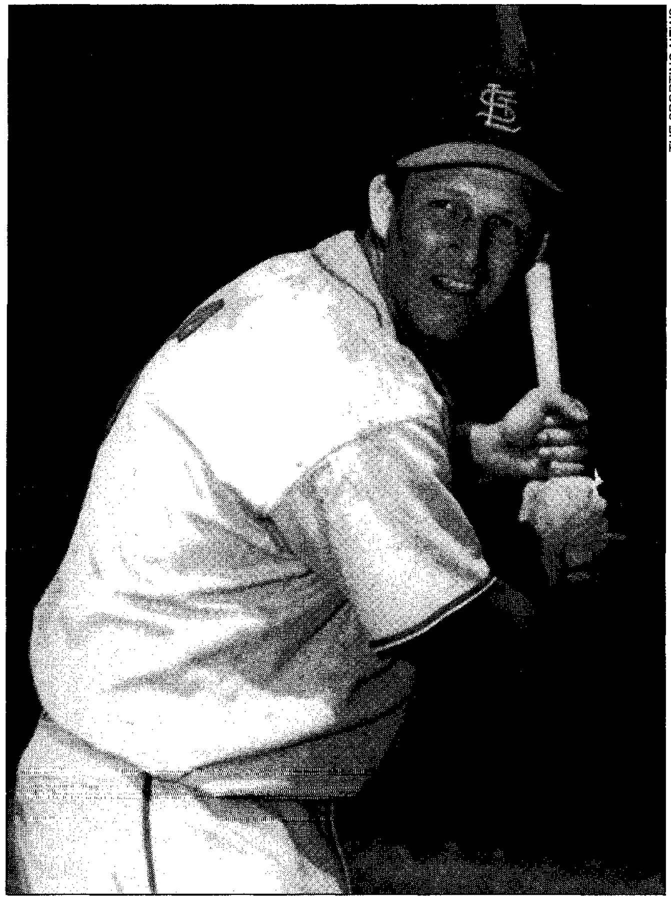
The effect of the screen on the games was quickly apparent. In the game of July 7 between the Browns and the Yankees, five drives landed against the RF screen. As intended by the Browns, the screen cost the visiting Yankees a homer (a line drive by Ruth)—but unfortunately, from the point of view of Browns fans; it also cost the Browns four homers—one by Fred Schulte and three by Heinie Manush. Instead, Manush had two doubles and a single.[5] The Browns won the game, 7–3. The general impression of the effect of the RF screen was reported in *The Sporting News*: "A week of the new screen in front of the RF pavilion has cheated the home guard out of four baggers on several occasions. . . also stopped the enemy. . . many of them who hit the screen are held to a single."[6]

I made a comparison of the Browns' and opponents' combined 1929 offensive performance before and after the RF screen was installed. The Browns' offense put up numbers (BA/Slugging Pct. adjusted for the opponents' pitching) of .291/.412 in the 35 games before the RF screen was installed, and .259/.346 after the screen. The apparent impact of the screen on BA is likely to be just random intra-season variation. The effect on homers was more evident and quite interesting. In the 35 games without the RF screen, the Browns hit 18 homers and the visitors 32. Despite more games, with the screen (42), the Browns' homers dropped to four, while the visitors had a more modest decline of six to 26. What had happened to the Browns' hitters? While the Browns had only two regular LHB (Manush and McGowan), the team did have two switch-hitters in their lineup—Lu Blue and Wally Schang. Before the erection of the screen, the two Browns' switch-hitters combined for nine homers at home (every one as a LHB); after the installation of the screen, they hit none! At home that season the Browns' LHB (including switch-hitters batting left) in 35 pre-screen games hit 15 homers and in 42 post-screen games hit only one (by Manush). In summary, the visitors' homers by LHB dropped from 21 to 13 while the Browns dropped from 15 to one. Clearly, the Browns miscalculated the impact of the screen!

Meanwhile, back in the NL, the Cardinals, as tenants in Sportsman's Park, were also affected. The combined offensive marks for the Cardinals and their opponents were: 294/.361/.442 (BA/OBP/Slugging) before the RF screen and .286/.337/.433 after the installation of the screen.[7] This suggests the screen may have had a modest negative effect on batting—or as is more likely, the drop in batting is due to random variation, plus changes in the mix of visiting teams (with better than average batters and/or pitchers) before and after the installation of the screen. The data for extra-base hits are more striking—in absolute terms, the Cardinals' homers dropped from 29 to 19, while the visitors experienced a similar decline—29 to 22. On a per AB basis, the combined results for the Cardinals and their opponents were an 18% increase in doubles, an 8% increase in triples, while home runs declined by 33%.

Given that the screen was installed only in RF, one would expect the greatest impact to be on extra-base hits by LHB. As the Cardinals' lineup included the same LHB both before and after the RF screen was installed, the change in extra base hits by LHB should be due to the screen. The impact of the screen on the visitors' LHB is more difficult to assess. In the 1929 season, the other seven NL teams played quite different numbers of games in Sportsman's Park before and after the installation of the screen. Likely of greater importance, the visiting NL teams had greatly varying mixes of LHB and RHB, not to mention power-hitting LHB. To correct for this factor, the visitors' batting data was adjusted for equal weighting per team in the pre-screen (37 games) and post-screen (40 games) time periods. The visiting teams' LHB (adjusted for equal weighting per team) in combination with the Cardinals' LHB produced offensive marks of .359/.425/.631 (BA/OBP/Slugging) pre-screen and .352/.402/.556 post-screen. On a per-AB basis, extra base hits by LHB (Cardinals and opponents) were affected by the RF screen as follows: doubles increased 27%, triples increased 33%, while homers dropped 35%. For each category of extra-base hits, the impact of the screen on LHB was greater than the impact on the teams as a whole, as would be expected.

The RF pavilion in Sportsman's Park had two noteworthy distinctions. One, as the screen extended up to the roof, the RF pavilion became the only major league park with extended outfield seating where it was impossible to catch a home run ball. Two, during this same interwar time period, the RF pavilion earned another significant albeit dubious distinction—the RF pavilion was the only part of the ballpark open to black fans during the era of segregated sports facilities.[8] Sportsman's Park was the last major league ballpark to end segregated seating, and it was not until 1944 that black fans were allowed to purchase tickets in other parts of the ballpark.



Stan Musial hit 475 home runs in his career, but never hit 40 or more in any season.

The screen remained in place until the end of the 1954 season. By this time, the St. Louis Browns had sold Sportsman's Park to the Cardinals, and the AL franchise had moved to Baltimore. The Cardinals were then in a position to vary the configuration of the ballpark as they chose. The Cardinals had kept track of drives hit off the screen during the 1954 season—the Redbirds had 35, the visitors only 18. The data for the 1954 season shows Musial would have had 10 more homers without the screen, while Red Schoendiest and Solly Hemus would have had five more apiece.[9] Because the Cardinals had a predominantly left-handed lineup, general manager Dick Meyer, with the concurrence of Manager Eddie Stanky, had the screen removed for the 1955 season. A new and slightly higher screen (overall height of 37 feet versus 33 feet before 1955) was installed for the 1956 season and remained for all subsequent seasons at Sportsman's Park.

A quick measure of the success of the removal of the screen is the Home/Road distribution of the 1955 season's homers by the Cardinals' LHB. The home/road splits were: Stan Musial 22/11, Wally Moon 15/4, Bill Virdon 13/4, and Red Schoendienst (a switch-hitter) 9/2. All nine of Schoendienst's homers at home were hit while batting left-handed.

The impact of the removal of the screen on hitting has been estimated by comparing the 1955 batting statistics of the Cardinals and their opponents with the batting statistics for the prior and subsequent seasons (1954, 1956). The offensive marks (BA/OBP/Slugging) were 1954: .295/.364/.445, 1955: .267/.337/.409, and 1956: .267/.337/.409. Of course, there were two distorting factors—(1) the Cardinals did not have entirely the same hitters in each of the three seasons, and (2) the overall league BA dropped each successive season 1954-56 (.265, .259, .256). Thus, no firm conclusion can be drawn from the St. Louis and opponents batting marks. A better comparison is extra-base hits by LHB in 1955 vs. the average of 1954 and 1956. The combined St. Louis and opponents 1955 extra-base hits (expressed per 100 AB) were doubles 3.38, triples 0.90, and home runs 5.38. The comparable data for 1954/56 were doubles 4.97, triples 1.32, and home runs 3.22. In simple layman's terms, in 1955 with the RF screen removed, extra-base hits (per 100 AB) by LHB at Sportsman's Park were doubles down 32%, triples down 32%, and home runs up 67%.

An interesting question arises: What would have been the effect on Hall-of-Famer Stan Musial if the screen had been removed for the entirety of his 1941-63 career? Only Musial's home batting data for all seasons except 1955 (no screen in place that season) were revised. Only extra base hits were revised since the removal of the screen would not have affected the total number of hits, as any ball hitting the screen was already a hit. Doubles and triples were adjusted by the average ratio of without-screen to with-screen for LHB, for NL 1929 and 1955 versus 1954/56[10] home runs by LHB were adjusted, by the average ratio of without-screen to with-screen for AL 1929, NL 1929, and NL 1955 versus 1954/56. The average without-screen adjustment factors were:

Doubles	.745
Triples	.743
Home Runs	1.873

The career extra-base hits for Musial and without-screen adjusted career extra base hits are:

CATEGORY	ACTUAL	ADJUSTED
Doubles	725	618
Triples	177	155
Home Runs	475	676

In this hypothetical scenario of playing at a Sportsman's Park with no RF screen, Stan Musial would have had a career home run total of 676, good enough for third place all-time. His career slugging percentage would be .610 (versus actual .559)—good enough for fourth all-time.

NOTES

1. Author's estimates based on land plat dimensions and diagram of later Sportsman's Park III.
2. *St. Louis Globe Democrat*, September 28, 29, 1912
3. Interview with Robert Tiemann, July 11, 2003
4. *Green Cathedrals*, Philip J. Lowry, SABR, 1986
5. *Cleveland Plain Dealer*, July 8, 1929
6. *The Sporting News*, July 18, 1929
7. Author's compilation from Official NL Day-by-Day batting data
8. *Even the Browns*, William B. Mead, Contemporary Books, Chicago, 1978, p. 133
9. *Take Me Out to the Ballpark*, Lowell Reidenbach, The Sporting News, St. Louis, 1983, p. 235

10. AL data for 1929 limited to homers for LHB (including switch-hitters as LHB), thus 1929 AL HR ratio is not per AB. In this hypothetical scenario of playing at a Sportsman's Park with no RF screen, Stan Musial would have had a career home run total of 676, good enough for third place all-time. His career slugging percentage would be .610 (versus actual .559)—good enough for fourth all time.

DARREN GLASS

Fair-Weather Fans

In Rob Neyer's chapter on San Francisco in his *Big Book of Baseball Lineups*, he speculates that there aren't really good baseball cities, and that attendance more closely correlates with winning percentage than with any other factor. He also suggests that a statistically minded person look at this. I took the challenge and have been playing with a lot of data.

METHODOLOGY

I looked at all seasons from 1973 until 2002. In particular, I looked at the correlation coefficients between the following variables:

- Average home attendance per game (ATT)
- Home attendance per game divided by average
- Home attendance over all teams (to normalize for nationwide trends) (ATT/AVG)
- Final place in divisional standings (PLACE)
- Winning Percentage (WIN)

There are a few basic properties of correlation coefficients (CC's). If a CC is equal to zero, then the two variables are uncorrelated, if it is close to one they are close to linearly correlated in a positive way, and if it is close to -1, then there is a strong negative relationship between them.

CORRELATION WITH WINNING PERCENTAGE

To begin with, let us look at the most naïve study: the correlation between winning percentage and home attendance. Over the 30 years between 1973 and 2002, the baseball-wide CC was .464. Table 1 lists teams that can be described as having fair-weather fans—their correlation between winning and attendance is more than 0.2 greater than the baseball-wide average.

Table 1. Teams with correlation coefficients between ATT and WIN greater than 0.2 above baseball average

Atlanta	0.884
Seattle	0.815
New York N	0.786
Cleveland	0.755
Montreal	0.753
Chicago A	0.752
San Francisco	0.673

On the other side of the spectrum are those teams that have correlation coefficients significantly lower than the baseball-wide average. An optimistic interpretation of this would be that the fans stick with the team no matter how badly they are doing (the case of the Red Sox and the Cubs), while a pessimistic interpretation might be that the fans refuse to support the team no matter how good they are. Table 2 lists cities that have correlation coefficients between ATT and WIN more than 0.1 below baseball average.

Table 2. Cities with correlation coefficients between ATT and WIN more than 0.1 below baseball average

St Louis	0.345
Chicago N	0.321
Texas	0.304
Tampa Bay	0.266
Milwaukee	0.234
Arizona	0.142
Pittsburgh	0.131
Los Angeles	0.117
Boston	0.004
Colorado	-0.087
Florida	-0.118
Baltimore	-0.246

The presence of all four of the expansion teams of the 1990s on this list makes sense, as the small sample size is distorted by the first few years in which novelty value runs high and the teams are not likely to be very good.

The most interesting data point on this list to the author is the Orioles, where the fans of Baltimore over the past 30 years actually supported the team significantly more the worse they have been. This is likely due in large part to the draw of the new ballpark at Camden

DARREN GLASS *experienced the phenomenon of fair-weather fans first hand when he was one of the dozen people to go to Atlanta Braves games in the mid-1980s. He is currently an assistant professor of mathematics at Columbia University.*

Yards, and that it has been successful in bringing in fans despite the fact that the Orioles have had losing records in six of the 11 years since it opened.

A slightly less naïve study would try to normalize for the effects on attendance of baseball as a whole. The average attendance at baseball games has nearly doubled over the last 30 years, and all of baseball took a hit in 1995, when the average attendance dropped nearly 6,000 fans per game. Thus, I also computed the CC's between ATT/AVG, a given team's average home attendance divided by the average attendance of baseball games league-wide, and winning percentage. The data did not qualitatively change significantly. The league-wide CC went up to .55.

Table 3. Correlation coefficients between ATT/AVG and WIN

Atlanta	0.925
Cleveland	0.832
Seattle	0.786
Philadelphia	0.753
New York N	0.752
Cincinnati	0.724
San Francisco	0.713
Oakland	0.692
Detroit	0.691
Kansas City	0.677
Minnesota	0.667
New York A	0.598
Tampa Bay	0.596
San Diego	0.573
Los Angeles	0.563
Montreal	0.557
Chicago A	0.541
Pittsburgh	0.539
Boston	0.532
Chicago N	0.520
Houston	0.505
Texas	0.489
St Louis	0.485
Toronto	0.478
Milwaukee	0.433
Anaheim	0.387
Colorado	0.303
Arizona	0.079
Florida	-0.035
Baltimore	-0.092

Statisticians say that a correlation coefficient is statistically significant if it is greater than the value of a certain T-test. While I will not go into the details of this calculation, I will point out that for our sample size of 802 team-seasons, any CC over .116 is statistically significant with probability 99.9%. In particular, our league-wide CC of .55 is extremely significant.

For the individual teams, sample sizes are much smaller. In particular, non-expansion teams have 30 data points, and thus a CC over .570 will be statistically significant 99.9% of the time, a CC over .463 is significant 99% of the time, and a CC over .361 is significant 95% of the time. When expansion teams with even smaller sample sizes are included, the CC's are significant at the 99% level for every team except Milwaukee, Anaheim, Baltimore, Toronto, Tampa Bay, Arizona, Colorado, and Florida.

Of course, the CC is not enough to capture what we are interested in. In particular, if a city's ATT/AVE and WIN were strongly correlated to a line with slope zero, we would view it as much less of a "fair-weather fan" city than a city with a weaker correlation to a line and a very large slope. I also computed the slope of the line given by various linear regressions baseball-wide—the results of a linear regression on ATT/AVG and WIN are ATT/AVG = 2.7525 × (WIN) − .3769. While ATT/AVG is a more meaningful statistic, it is also harder to get a feel for. For this reason we will note that the linear regression between ATT and WIN gives ATT = 63,476 × WIN − 7,740. In other words, by increasing winning percentage by .100 (an improvement of roughly 16 wins per season), a team can expect to boost home attendance by an average of 6,347 fans per game.

Table 4. Slopes from linear regressions between ATT/AVG and WIN

Cleveland	4.543672
Philadelphia	4.290944
Atlanta	3.850382
Cincinnati	3.735552
Los Angeles	3.431718
Seattle	3.328853
San Francisco	3.206009
New York N	3.134074
Kansas City	3.067628
Minnesota	2.862508
Montreal	2.772461
Oakland	2.403002
Chicago A	2.214931
New York A	2.202218
Detroit	2.186652
Houston	2.157452
Toronto	2.114608
Boston	1.920404
Anaheim	1.917665
Colorado	1.888440
San Diego	1.858157
Texas	1.775284
St Louis	1.746337
Chicago N	1.634861
Tampa Bay	1.578699
Pittsburgh	1.374932
Milwaukee	1.304664
Florida	-0.15230
Baltimore	-0.37538
Arizona	-0.99382

A natural question to ask, and one that more than a few people are looking at due to its various political implications, is how new stadiums affect attendance. While I did not investigate this phenomenon in any depth, I will note that if you remove all data points in the data set corresponding to the first two years that a team is in a new city or a new stadium, the baseball-wide CC actually raises by .05.

CORRELATION WITH PLACE FINISHED

It is also natural to wonder if it is not the winning percentage that brings in the fans but being in the hunt of a pennant race. I decided to test this hypothesis by calculating the correlation coefficients between our attendance variables and the place in which a team finished within their division, as well as how many games back they finished. Because the nature of both of these variables changed significantly with the realignment in 1994, I ran the study first looking only at the data from the years 1973-1993. In particular, it was not clear how to best handle the situation with the wild card, and teams that might be in the hunt for the wild card despite being many games out of the division lead (see 2003 Phillies and Marlins, for example). It came as a surprise to the author that including the last decade did not significantly change the results, as seen by the following charts:

1973 to 1993	CC	SLOPE
ATT/AVE and PLACE	-0.5590	-0.1050
ATT and PLACE	-0.4632	-2136.5000
ATT/AVE and GB	-0.5300	-0.0164
ATT and GB	-0.4535	-343.1290

1973 to 2002	CC	SLOPE
ATT/AVE and PLACE	-0.5590	-0.0978
ATT and PLACE	-0.5016	-2491.0100
ATT/AVE and GB	-0.4906	-0.0145
ATT and GB	-0.4131	-334.6898

In all of these examples, CC is negative. This is what we would expect as the "higher" your value of PLACE and GB, the less attendance we might expect to see.

I have not included the team-by-team data, but it is qualitatively very similar to the above team-by-team data, with the teams falling in roughly the same order and with the same significance results. Anyone who is interested in the full data should feel encouraged to email me.

CORRELATION WITH PAST PERFORMANCE

Another question that comes up is how correlated attendance is with past performance. In particular, looking at the correlation between winning percentage (or standings) in year x and attendance in year (x+1). The idea being that the rush of winning the World Series creates new fans (and season ticket holders) no matter how badly the team performs the following year.

However, when one runs the numbers, they are not particularly illuminating. In fact, the CC's one gets from comparing last year's winning percentage and this years ATT/AVG is .492, slightly less than when you compare this year's record with this year's attendance, .551. (See below for the full chart of CC's.) Furthermore, the only teams for which there is a substantial difference in the CC's when you run the study the two ways are Colorado (which can be partially explained by the fact that you had a small data set to begin with and are reducing it even further), Minnesota, Montreal, Pittsburgh, and St Louis. Furthermore, in each of these cases there is a weaker correlation. So while my instincts agreed with what many of you suggested might be an interesting effect, the numbers don't seem to bear it out.

	WIN	PREV WIN	PLACE	PREV PLACE
ATT/AVG	0.5505	0.4926	-0.5016	-0.4651
ATT	0.464	0.4293	-0.4669	-0.4329

One problem in trying to do such a study is that there is a relatively strong correlation between how a team does in year X and how it does in year x+1 (CC = .5 for my data set). Isolating that factor would be hard but not impossible.

CONCLUSIONS

Every one of the tests which I ran seems to indicate that Rob Neyer's hypothesis is correct: attendance at ball games is highly correlated with the winning percentage of the home team. This is certainly true baseball-wide, and is also true for almost every team individually. The exceptions by and large are the expansion teams of the 1990s and the Baltimore Orioles. Furthermore, in almost every permutation of the data, it seems that the fans of Cleveland, Atlanta, and Seattle are especially prone to support their teams more the better they do. We do note, however, that all three of these teams got new stadiums while the teams were doing especially well—and in the case of the Braves and the Indians this was also at a time when baseball was seeing a drop in attendance nationwide—which likely skews the data somewhat.

FURTHER EXPLORATIONS

I think it would be very interesting to look at attendance in smaller units than seasons. This could take away some of this effect by looking at when in (for example) the 1991 season the fans stopped punishing the Braves and Twins for previous subpar performance and rewarded them for being good.

However, to do this one would have to control for factors such as weekend games (which generally have higher attendance) or superstar players coming through town (which certainly boosts attendance) or the like, factors which one can ignore over the course of a season but which could significantly affect the data when looking at units of individual games or weeks or even months.

Another thing that I would like to do is to try to adjust for ballpark size. The only way I could think of to do this would be to use "percentage of seats filled" as my attendance variable, but this seems to pose more problems than it solves. I certainly like the idea of "rewarding" the Cubs and Red Sox and other teams which could sell more seats if they had the capacity, but I'm not sure if it makes sense to "punish" cities for having large stadia in this way. For example, if Stadium One holds 50,000 people and Stadium Two holds 60,000, I do not think that it makes sense to treat the fact that they both draw 30,000 fans differently. It also seems like a bit of opening Pandora's box as we really don't know how many fans the Red Sox would average if they had an infinitely big stadium. It could be that their attendance would stay the same or it could be that it would quadruple—we have no real way of knowing.

REFERENCES

All data came from www.baseball-reference.com

CHICAGO CUBS NEWS, AUGUST 6, 1937

Cubs fans who visited Wrigley Field on July 4 and 5, really went to town with the ol' feed bag.

During the two days, on which doubleheaders were played, the Wrigley Field concession stands broke all existing records in the dispensing of food and beverage.

Three tons of hot dogs were consumed; 11,000 ham, cheese and hot roast beef sandwiches; 19,200 bars of candy; 16,000 sacks of peanuts; 11,280 bags of popcorn; 600 loaves of bread (two pounds to a loaf); 18,000 packages of ice cream; 19,200 cans of beer; 18 barrels of beer; 6,000 lemonades and 41,000 bottles of pop.

Wrigley Field is becoming one of Chicago's most popular eating places when the Cubs are in town. Many fans attending the games plan to have their lunch at the park, where there is a wide variety of food selections."

ERIC MARSHALL WHITE

Strong Down the Stretch

Warren Spahn's Fantastic Finishes

It seems unlikely that anything new could be said about the storied pitching exploits of Warren Spahn, who put together the greatest career by any left-handed pitcher during the postwar era. He is remembered not only for his excellence, but also for his consistency, having won 20 or more games in a season 13 times, as well as for his longevity, having pitched no-hitters at the ages of 39 and 40 and becoming the oldest 20-game winner at 42. He led his league in victories eight times, won the ERA title in three different decades, and took the Cy Young Award in 1957 (he would have won three or four more had the NL award been around as long as he was). Despite devoting three years to military service in World War II, he finished with a lifetime total of 363 wins. Clearly, Spahn belongs on the all-time pitching staff—but this is hardly news.

However, one particularly amazing aspect of his performance seems to have escaped notice entirely: *Warren Spahn may have been the greatest pitcher of all time when it came to improving his performance down the stretch run in August and September.* I've read countless anecdotes and analyses concerning Spahn during the past 30 years, but not once have I seen one mention of his incredible career-long pattern of getting hot down the stretch. In the zone, turning it up a notch, huge in the clutch, saving his best for last—whatever you want to call it, Spahn became a different pitcher after August began, raising the relative level of his performance to a degree that is probably unparalleled in the history of the sport. His career numbers before and after the first of August are a revelation.

Mark those career winning percentages down in your mind: Spahn was 130 percentage points higher in the late going. Certainly no clutch hitter so consistently raised his stretch-run numbers to such a degree over two decades. For a team, Spahn's sudden improvement is like starting 53-44 and going on a 44-21 tear to finish with 97 wins. A team playing like Spahn's first four months might sneak into the wild-card slot once in a while, but a team keeping pace with Spahn's final months all year would rank among the greatest.

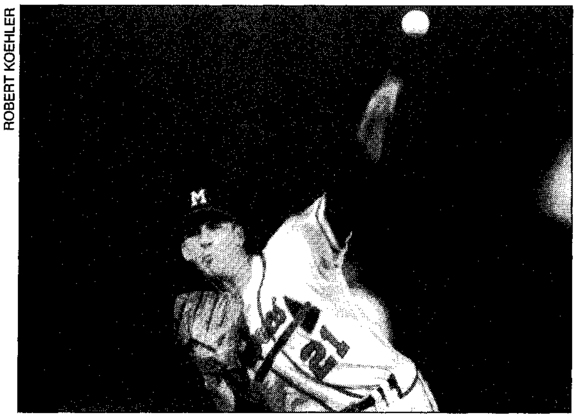
Table 1. Warren Spahn's career record, before and after August 1

YEAR	TOTAL W-L	APR.-JULY	%	AUG.-OCT.	%	DIFF.
1946	8-5	3-1	.750	5-4	.556	-.194
1947	21-10	13-5	.722	8-5	.615	-.107
1948	15-12	8-7	.533	7-5	.583	+.005
1949	21-14	12-8	.600	9-6	.600	+.000
1950	21-17	13-11	.542	8-6	.571	+.029
1951	22-14	11-9	.550	11-5	.688	+.138
1952	14-19	9-11	.450	5-8	.385	-.065
1953	23-7	13-4	.765	10-3	.769	+.004
1954	21-12	10-10	.500	11-2	.846	+.346
1955	17-14	9-11	.450	8-3	.727	+.277
1956	20-11	10-8	.555	10-3	.769	+.214
1957	21-11	10-8	.555	11-3	.786	+.231
1958	22-11	13-7	.650	9-4	.692	+.042
1959	21-15	14-10	.583	7-5	.583	+.000
1960	21-10	11-7	.611	10-3	.769	+.158
1961	21-13	10-12	.455	11-1	.917	+.462
1962	18-14	10-11	.476	8-3	.727	+.251
1963	23-7	13-5	.722	10-2	.833	+.111
1964	6-13	6-11	.353	0-2	.000	-.353
1965	7-16	4-12	.250	3-4	.429	+.179
TOTAL	363-245	202-168	.546	161-77	.676	+.130

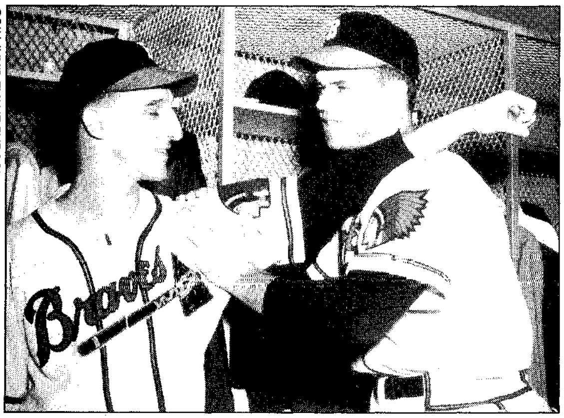
Did Spahn's late-season dominance make him more valuable than his overall 363-245 record would suggest? I'm not convinced that performing well down the stretch is particularly important. It certainly sounds good to say, "He gets hot when it counts," but when exactly *doesn't* it count? Mr. Steinbrenner can call an expensive player "Mr. May" all he wants, but those winning hits in May count just as much toward the standings as those that come in September. As Rocky Bridges rightly said of the Brooklyn Dodgers after Bobby Thomson's home run won the 1951 playoff for the Giants, "We lost the pennant on Opening Day." Still, there is an emotional difference that intensifies those late summer games as the players become more aware of the standings and begin to feel more clarity

ERIC MARSHALL WHITE, *Ph.D. is curator of rare books at Bridwell Library, Southern Methodist University, in Dallas. His baseball card collection includes all the World Series starting lineups since 1951.*

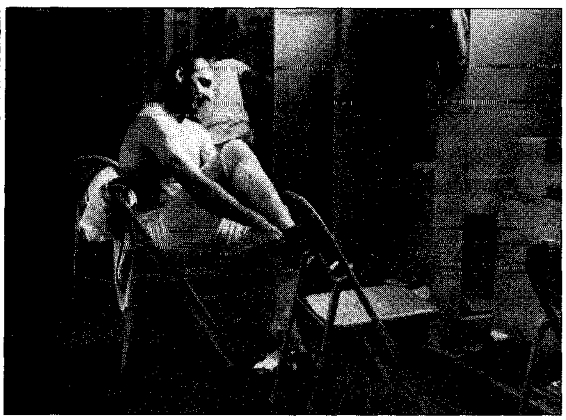
ROBERT KOEHLER



TRANSCENDENTAL GRAPHICS



ROBERT KOEHLER



Warren Spahn is pictured in all three photographs, and he is joined in the middle photo by teammate Johnny Sain. In September 1948 the two were immortalized by Boston Post writer Gerry Hern in a poem: *First we'll use Spahn, then we'll use Sain/Then an off day, followed by rain/Back will come Spahn, followed by Sain/And followed, we hope, by two days of rain.* The poem was shortened in the public memory to *Spahn and Sain and pray for rain.*

of purpose. It may not be more important to win late in the season, but perhaps it is more difficult. I would venture to guess that a lot more pennants are won with hot Septembers than with cold ones. To put it another way, a less worthy team is more likely to collapse late while a true winner is able to play well under the emotional pressure as time runs out. Spahn's teams just missed a couple of pennants they could have won, but that wasn't his fault. Let's look more closely at how he did down the stretch through the years:

1946 Nothing to attract notice, just a solid 5-4 finish.

1947 His 8-5 finish would have been even better if not for 1-0 and 2-0 losses to Brooklyn and Cincinnati. Still, he won three shutouts (including a 1-0 beauty against the Cubs September 14) while allowing only five runs in his six starts during September.

1948 The famous refrain "Spahn and Sain and pray for rain" has drawn many critics, who cite the fact that at 15-10, Spahn's winning percentage was lower than that of his teammates. However, *Boston Post* writer Gerry Hern had ample reason to wax poetic that September. While the other Braves pitched nearly as well as Sain and Spahn, it was not nearly as often. Sain and Spahn dominated Boston's run of 14 wins in 15 games that month, at one point winning 10 straight decisions (four by Spahn). Between September 6 and 18, the duo pitched eight of Boston's 10 games, winning all of them (four wins each). That streak started with Spahn's 14-inning, 2-1 masterpiece over Brooklyn, in which he picked Jackie Robinson off first base *twice*. It was one of five late summer games that he won with only two runs of support. He lost his last two starts, but the pennant had been won.

1949 Although Brooklyn came back to beat him twice in their September run for the pennant, Spahn's 9-6 finish was highlighted by a 4-0 win over Brooklyn's Preacher Roe on August 20 and a 1-0 masterpiece over Philadelphia's Robin Roberts on September 10.

1950 Spahn's 8-6 finish was only a slight improvement on his 13-11 start, but his record rose to 5-1 in September.

1951 With a ledger of 11-5, this stretch run marked Spahn's first truly outstanding finish. After he lost 1-0 to the Phillies on August 7, the Braves scored a grand total of three runs for him in three close losses to the Giants, who were busy winning 37 of their last 44 games. The final defeat was a critical 3-0 win for Sal Maglie on the season's next-to-last day that helped set up the Miracle of Coogan's Bluff.

1952 The dismal final season in Boston had few bright spots. Although he lost 1-0 to Brooklyn's Carl Erskine on September 20, some consolation may have come from the fact that his final four victories were all shutouts, including his own 1-0 whitewashing of Cincinnati on September 9. To this point, Spahn's lifetime record from August 1 stood at 53-39 (.576), very good, but just a shade above his overall record. His truly remarkable run was about to begin.

1953 As the Milwaukee Braves surprised everyone with their success, Spahn had the best start of his career (13-4) and his best finish to that point (10-3). He allowed only 11 runs during a 6-1 month of August, and following a 2-0 loss to Philadelphia, he won his last four decisions.

1954 After a lackluster 10-10 start, Spahn sizzled to a 11-2 record down the stretch. He won 11 straight games between July 18 and September 8, and won his last two starts of the year.

1955 Spahn followed up another slow start with a stretch run of 8-3. His last two losses were both by the score of 2-0.

1956 The southpaw carried the Braves during their pennant bid with another 10-3 finish. In September, while his teammates struggled to a 9-11 finish in games in which he did not work, Spahn held on for a 5-2 record—his final loss coming against the Cardinals in a 12-inning, 2-1 game on September 29.

1957 The pennant was Milwaukee's in 1957, thanks in large part to the Cy Young Award winner's 11-3 finish. He won nine straight decision between August 6 and September 7, and his final defeat, another loss by shutout, came five days after Hank Aaron's homer had clinched the pennant.

1958 Spahn got off to a 6-0 start, but cooled off to .500 over his next 14 decisions (including three shutout losses). He finished 9-4 during the last two months, including a typical 5-1 record in September as the Braves took their second straight pennant.

1959 Leading up to Milwaukee's heartbreaking playoff loss to Los Angeles, Spahn contributed a 7-5 mark with a 4-2 record in September, including a victory over Robin Roberts in his season finale. Still, any improvement upon his season record of 0-5 versus the Dodgers could have made all the difference in the pennant race.

1960 With a 6-0 record in August, Spahn finished the season with another 10-3 run that included his first career no-hitter, a 4-0 masterpiece against the Phillies on September 16. His final three defeats all came at the hands of the world champions to be, Pittsburgh.

1961 On August 11, the 40-year-old won his 300th ball game, leveling his record at a respectable 12-12. He rallied to his greatest finish of all, an amazing 11-1 roll that included three September shutouts, highlighted by a 1-0 affair versus the Phillies on September 6. Milwaukee's opponents scored 11 runs in the 11 wins. His only loss came after 10 straight wins, in a September 15 blowout against his future successor atop the pitching world, Mr. Koufax.

1962 Somehow, he did it again, finishing 8-3.

1963 The oldest 20-game winner in history, Spahn started off great (despite his famous 1-0 loss to Juan Marichal on July 2, when Willie Mays homered in the bottom of the 16th inning), and he finished even better at 10-2, with yet another 1-0 masterpiece versus Pittsburgh's Bob Friend among his three September shutouts.

1964 With an 0-2 finish, Spahn's great run was over.

1965 A 3-4 finish isn't bad—for a 44-year-old!

Thus, we see that Spahn was a fantastic finisher, and not just because he finished 449 contests in his career and won 77 games after turning forty. For whatever reason, Spahn started his seasons relatively slowly but had a formidable late kick. I hesitate to speculate here, but was it because he was unusually fit for his time, a marathon man able to outlast his tiring opponents late in the year? Even though Spahn was a perfect 7-0 in his late-season duels with the famously durable Robin Roberts, that explanation doesn't sound quite right. Isn't it more likely that this pervasive trend of improvement related to Spahn's well-documented pitching intelligence, which made him especially adept at solving problems over the course of the season with craft and guile? I will point out that the effect was barely perceptible during Spahn's early years, but it intensified throughout his maturity. Run support? I have the game scores for Spahn's first 504 decisions (from 1946 to 1961), and the Braves averaged a composite 4.50 runs per game through July and 4.29 runs per game from August to the season's end hardly the offensive push one might expect, given Spahn's record. Milwaukee's opponents scored 3.56 runs per game in Spahn's starts through July, and only 3.03 runs thereafter. Spahn was magnificent down the stretch.

I imagine that a small handful of great pitchers may have compiled late-season winning percentages that were greater than Spahn's, but I have yet to find anyone with a lengthy career that had such a large before-and-after differential. At first glance, the candidates most likely to have exceeded his .676 late summer record were Whitey Ford, with his .690 overall percentage, and Lefty Grove, at .680 lifetime. Ford, though, was "only" 6-2 and 7-3 in the closing months of his 20-win seasons, so it's hard to see where any truly gigantic finishes would be hiding. Grove had several astounding finishes, and most likely exceeded Spahn's .676 mark in the late going, but he could pitch like that all summer long, so I would not expect much of an early-late split. Indeed, the higher a pitcher's lifetime winning percentage, the less likely he is to have run up a large differential: if Grove had compiled such a Spahnian split, his percentages would have been .629 before and .759 after August 1, and it's hard to picture that happening over a couple of decades, given his starts of 17-2 in 1929, 21-2 in 1931, and 14-3 in 1938. Thanks to Ronald A. Mayer's *Christy Mathewson: A Game-*

by-Game Profile, I was able to add up Matty's totals: 233-105 before August 1 and 140-83 thereafter, a drop from .689 to .629, despite his huge 18-4 finish in the 1908 race.

Actually, the candidates most likely to match Spahn's 130-point career differential are those with less stratospheric winning percentages, like Tom Seaver at .603 and Bob Gibson at .591 (although everyone remembers Gibby as better than that). However, a quick check shows that although Seaver put in a 10-1 finish for the Miracle Mets in '69, he was lackluster in his last few seasons. Gibson, at his best in October classics, missed six weeks late in the 1967 race and finished only 7-4 in his famous '68 campaign.

Walter Johnson? He, too, is unlikely to have approached Spahn's feats. Although he finished his amazing 1913 season on a 13-2 run, he had started 23-5, making it a modest improvement only. We may read of his strong finishes in 1915 and 1924, but I suspect these were the exceptions, as his great 1912 season ended with a 4-5 slide, and the Big Train's surprisingly mundane career road record of 184-162 (.532) probably prevented him from compiling too many two-month winning binges. His famous winning streaks of 1912, 1913, and 1924 started in late June and early July.

What sets Spahn's record apart is that he won more than he lost from August onward for 11 straight seasons, and 17 times overall, most often by healthy margins. His winning percentage improved (or held even) late in the season 16 times, and he enjoyed five perfect Augusts, winning six without a loss in 1954, seven in 1957, six in both 1960 and 1961, and five straight in 1963. Perhaps most incredibly, his peak decade came when he was 32 to 42 years young!

Among today's greats, Roger Clemens has been the most consistent, Greg Maddux has faded slightly in the late going, Randy Johnson owns the largest positive differential so far (albeit in half as many stretch-run decisions as Spahn compiled), and Pedro Martínez has come back to earth from incredible heights (statistics through the 2003 season):

Table 2. A comparison with today's aces

PITCHER	APR.-JULY	%	AUG.-OCT.	%	TOTAL	%
Clemens	206-108	.656	104-53	.662	310-160	.660
Maddux	191-104	.647	98-59	.624	289-163	.639
Johnson	145-82	.639	85-32	.726	230-114	.669
Martínez	115-39	.747	51-28	.646	166-67	.712

Perhaps the best way to put what Spahn did in the most meaningful context is to compare him to two of his contemporaries, Sandy Koufax (165-87) and Don Drysdale (209-166). Combined, the pair of Dodgers had a lifetime record of 374-253 (.596), which is directly comparable to Spahn's 363-245 (.597) mark. However, whereas the two Dodgers got off to a better start than Spahn did, they fell way off the pace down the stretch:

Table 3. A comparison with Dodger greats

PITCHER	APR.-JULY	PCT.	AUG.-OCT.	PCT.	TOTAL	PCT.
Koufax	115-49	.701	50-38	.568	165-87	.655
Drysdale	140-107	.567	69-59	.539	209-166	.557
Kfx + Dry	255-156	.620	119-97	.551	374-253	.596
Spahn	202-168	.546	161-77	.676	363-245	.597

What is amazing is that the careers of Koufax and Drysdale added together look a lot like Spahn's splits:

PITCHER	W-L	PCT.	PITCHER	W-L	PCT.
Drysdale	209-166	.557	Spahn(Apr.-July)	202-168	.546
Koufax	165-87	.655	Spahn(Aug.-Oct.)	161-77	.676
Kfx + Dry	374-253	.596	Spahn(Total)	363-245	.597

Thus, in the early going Spahn was *worse* than Don Drysdale, a marginal selection for the Hall of Fame, while down the stretch he was *better* than Sandy Koufax, whom many consider the most dominant pitcher of their lifetime.

Although Spahn's perennial surges seem to have evaded the notice of baseball writers and fans, they certainly merit close attention. Moreover, this data should encourage us to take another look at the career profiles of the other great pitchers, not only to see if anyone else had it in him to raise his performance to such dramatic heights down the stretch, but to better understand exactly how the Braves' crafty lefty did it so often. For now, however, we may appreciate Warren Spahn's greatness in a revealing new light—as the most fantastic finisher of them all.

WARREN WILBERT

1-0 Ball Games

Oh, Those 1-0 Ball Games!

If you're the loser it's a trip down Heartbreak Boulevard, but if you're the winner you're on cloud nine. Missed opportunities, errors, home run pitches, fluke hits, walk-off homers, and those unintentional bases on balls with the bases loaded all have a way of haunting both ballplayers and fans long after the game goes into the record books. It leaves them limp, ecstatic, grousing or high-fiving, depending who got the win or who took it on the chin. Here's a sampling.

It's the bottom of the 13th at Qualcomm Stadium, San Diego, September 2, 2001. 23,475 are on hand. There is nothing but goose eggs on the board. Padres manager Bruce Bochy and Bob Brenly, the Arizona Diamondbacks skipper, have successfully maneuvered their clubs through 12-plus innings, having made 11 pitching changes, a number of batting order and lineup changes, and have otherwise made the moves that have kept their clubs even-up. One more change was needed. It was time for Brenly to bring on his closer, Byung-Hyun Kim. Kim would face lefty-hitting Ryan Klesko, brought into the game as the Padres' third first baseman, and slated to be the leadoff hitter in the bottom of the 13th.

The inning didn't take long. Klesko pulverized a Kim offering for a walk-off homer that beat the Diamondbacks. Klesko's blow made starter Randy Johnson's 14 K's, the brilliant relief work of four relievers, and everything else that Brenly & Co. did to forestall defeat irrelevant.

Between May 11, 1875, the date of the first professional 1-0 baseball game, and the final game of the 2000 World Series, 174,901 professional games were played in seven different leagues and in regular season, All-Star and championship play. Over the course of those years there have been 4,001 games played where the final score was 1-0. That amounts to a miniscule .0229%.

Fort Wayne sabermetrician **WARREN WILBERT**, *a bleeding-heart White Sox fan, has published seven books about baseball, the latest on—you guessed it—the White Sox. His next book will feature the 26 best games played in the history of the World Series.*

But that is precisely what makes the one-run game the thriller it is, a rarity in baseball. Granted, 2-1 and 2-0 games can be just as exciting. But there is something about this minimalist, one-score-takes-it-all kind of contest that sets it apart as something special. No-hitters, one-hitters, games with brilliant fielding gems that have prevented scores, heads-up defensive play and audacity on the base paths are just some of the many factors that come into play as these spine-tingling nail-biters unfold. If nip-and-tuck baseball is your thing, this is the kind of game you want to see.

Table 1. A summary of 1-to-0 games

LEAGUE/EVENT	YEARS	GAMES	1-0 GAMES	%
NABBL	1871-75	1086	4	.0037
American Assoc.	1882-91	5039	45	.0089
Union Assoc.	1884	428	5	.0117
Players League	1890	529	2	.0004
Federal League	1914-15	1243	45	.0362
National League	1876-2000	88765	2075	.0234
American League	1900-2000*	76668	1788	.0233
All-Star Game	1933-2000	71	1	.0140
Championships	1884-1900	85	0	--
Tiebreakers	1946-2000	16	0	--
NLCS-ALCS	1969-2000	413	13	.0314
World Series	1903-2000	558	23	.0412
TOTALS		174,901	4,001	.0229

*The Western League was renamed the American League for the 1900 season, and its 1-0 games were included in the AL's grand total.

The chronological listing in this article presents some noteworthy 1-to-0 ball games. These only scratch the surface, of course, but their historical or individual significance should not go by without mention. I expect the reader could add a few more.

Because pitchers play such a dominating role in most of these contests it is interesting to note who among the many moundsmen in the game's history are most successful. Two lists are presented, one featuring Hall of Famers in Table 2 and another listing the lesser lights in Table 3, some of whose names may just surprise you.

Table 2. Hall of Fame 1-0 pitchers (min. no. of 1-0 wins, 8)

PITCHER (ML YEARS)	1-0 WINS	CAREER W-L	ShO	%*
Johnson, W. (21)	38	417-279	110	.35
Alexander, G. (20)	16	373-208	90	.18
Mathewson, C. (17)	14	373-188	79	.18
Ryan, N. (27)	14	324-292	61	.23
Plank, E. (17)	13**	305-181	69	.19
Palmer, J. (19)	13	268-152	53	.25
Jenkins, F. (19)	13	284-226	49	.27
Young, Cy (22)	12	511-316	76	.22
Brown, M. (14)	12**	208-111	55	.22
Walsh, E. (14)	12	195-125	57	.21
Perry, G. (22)	12	314-265	53	.23
Seaver, T. (20)	12	311-205	61	.20
Willis, V. (13)	11	249-205	50	.22
Koufax, S. (12)	11	165-87	40	.28
Gibson, R. (17)	11	251-174	56	.20
Carlton, S. (24)	11	329-244	55	.20
Nichols, C. (Kid) (15)	10	361-208	48	.21
Joss, A. (9)	10	160-97	45	.22
Coveleski, S. (14)	10	215-142	38	.26
Waddell, G. (13)	9#	193-143	50	.18
Spahn, W. (21)	9	363-245	63	.14
Ford, E. (16)	9	236-106	45	.20
Sutton, D. (23)	9	324-256	58	.16
Grove, R. (17)	8	300-141	35	.23
Bunning, J. (17)	8	224-184	40	.20

*Column lists % of 1-0 games to shutouts
**Includes Federal League wins (Plank, 3; Brown, 2)
#Waddell also had one 1-0 game in the 1900 American League.

Table 3. ML non-HOF pitchers (min. no. of 1-0 wins, 8)

PITCHER (ML YEARS)	1-0 WINS	CAREER W-L	ShO	%*
Chance, D. (11)	15	128-115	23	.65
Blyleven, B. (22)	13	287-250	60	.22
White, G. (Doc) (13)	11	189-158	45	.24
Rucker, G. (Nap) (10)	11	134-134	38	.29
Koosman, J. (19)	11	222-209	33	.33
Tyler, G. (12)	10	127-116	30	.30
Doak, W. (16)	10	169-157	34	.29
Lee, W. (Big Bill) (14)	10	169-157	29	.35
Horlen, J. (12)	10	116-117	18	.56
John, T.J. (26)	10	288-231	46	.22
Donovan. W. (18)	9	185-139	35	.26
Vaughn, J. (13)	9	178-137	41	.22
Russell, E. (Reb) (7)**	9	80-59	24	.38
Nehf, A. (15)	9	184-120	27	.33
Derringer, P. (15)	9	223-212	32	.28
Trucks, V. (17)	9	177-135	33	.27
Perry, J. (17)	9	215-174	32	.28
Finley, C. (17)	9	200-173	15	.60
Cicotte, E. (14)	8	209-148	35	.23
Adams, C. (Babe) (19)	8	194-140	44	.18
Cooper, W. (15)	8	216-178	35	.23
Bush, Joe (17)	8	196-184	35	.23
Shawkey, R. (15)	8	195-150	33	.24
Walters, W. (16)	8	198-160	42	.19
Vander Meer, J. (13)	8	119-121	29	.28
O'Toole, J. (10)	8	98-84	18	.44
Pappas, M. (17)	8	209-164	43	.19
Lolich, M. (16)	8	217-191	41	.20

* Column lists % of 1-0 games to shutouts
** Career years as pitcher only

The Peerless One, Walter Johnson, is also peerless when it comes to getting involved in 1-to-0 ball games. The Big Train piled up a staggering 64 1-0 decisions, winning 38 of them, or, about three out of every five. Considering the Washington teams usually behind him, that 60% reading overall is next to extraordinary. His record of 38 career 1-0 shutout victories is one of the untouchables among baseball records, and is more than twice those of his closest rival, Pete Alexander, who accomplished the feat 16 times. Walter Johnson, Alexander, Mathewson, and Nolan Ryan form an olympian quartet atop this distinguished list.

Among non-Hall of Famers Dean Chance deserves special mention, having won 15 games by 1-0 scores. In 1964 he was victorious at that barest of winning margins no less than six times, and dressed up his shutout list that season with additional 2-0, 3-0, 4-0 (twice) and two 7-0 outings. Bert Blyleven with 13 for several clubs and three southpaws, Guy "Doc" White, of the champion 1906 Chisox, Nap Rucker of the old Brooklyn Robins, and the more recent Jerry Koosman follow with 11 each.

An imposing array of hurlers, 53 all told (25 Famers and 28 non-Famers), make up the two lists. The listing's arbitrary cutoff number of eight leaves behind quite a number who did it seven times, many of them Hall of Fame hurlers.

If your hunch is that the World Series might be a rich source of exciting 1-0 games, it's right on the money.

A few examples follow:

- In that famous 1906 Series won by Chicago's Hitless Wonders, Mordecai Brown fired a two-hitter in Game Four to keep the Cubs in contention.

- Little Artie Nehf beat Waite Hoyt in the deciding game of the 1921 World Series. The game's outcome was decided in the first inning on an unearned run, and Nehf made it stand up as the Giants beat Miller Huggins' Yankees.

- Between 1948 and 1950 three consecutive World Series featured a first game ending in a 1-0 score. In Game One of the 1949 fall classic, the Yankees' Tommy Henrich hit the first Series walk-off blast to beat the Dodgers.

- In 1959 at mammoth Los Angeles Coliseum the largest crowd to see a World Series game, 92,706, saw the White Sox threesome of Bob Shaw (WP), Billy Pierce, and Dick Donovan combine to edge Sandy Koufax and the Dodgers.

- The 1966 four-game World Series showcased three shutouts, two of them by 1-0 scores. Both were won by Orioles; Wally Bunker beat Claude Osteen and Dave McNally defeated Don Drysdale. Jim Palmer threw the third goose-egg special. After scoring a run in the third inning of Game One, Baltimore pitching held the Dodgers scoreless the rest of the way, running up 33 straight scoreless innings on Walter Alston's club. Incredible!

- In one of the greatest classics in World Series history, Minnesota, behind the determined pitching of Jack Morris, beat Atlanta's Braves, winning the Series-deciding game in the 10th inning at the Metrodome in 1991.

- David Justice provided the game's only run as Atlanta beat Cleveland in the final game of the 1995 World Series. His dinger in the sixth inning was all Tom Glavine needed as the Braves' lean lefty stifled the Indians on one hit. Mark Wohlers came on in the ninth to preserve the one-hitter and get the save.

The seven games above are among the best of the 23 World Series 1-0 games played between 1903 and 2000. There'll be more, rest assured. The following is by no means inclusive, but is a timeline of outstanding 1-0 contests played over the years:

May 11, 1875
Professional baseball's first 1-0 game
The Chicago White Stockings won at St. Louis against the Red Stockings in a game that was played in a windstorm. Each team had six hits. George Zettlein was the winning pitcher.

June 12, 1880
The first ML perfect game
Host Worcester's Ruby Legs, behind left hander John Lee Richmond, won MLB's first perfect 1-0 game over the Cleveland Spiders. Jim McCormick was the losing pitcher.

June 13, 1882
American Association and Union Association 1-0 games
The AA (1882-1891) played a modest 82-game schedule in its inaugural season. The league's first champion was Cincinnati's Reds, who also won the league's first 1-0 game at Baltimore. The Union Association (five 1-0 games in 1884, its only season) and the Players League (1890) vied unsuccessfully with the NL as a major baseball league.

August 17, 1882
The first 1-0 game won by a walk-off homer
Hall of Famer Charley "Old Hoss" Radbourn, better known for his pitching, played the outfield in this 18-inning game and won it with his walk-off for Providence. Hall of Famer John Montgomery Ward was the WP against Detroit's Wolverines that day.

June 21, 1890
The first 1-0 game in the Players League
Only two 1-0 games were played in the Players League's one season of play. One of them turned out to be a heartbreaker for the Brooklyn Wonders, whose "Silver" King no-hit Chicago's Pirates but lost when the host Pirates scored a run in the bottom of the seventh. A rainstorm after the eighth halted play.

July 25, 1897
First 1-0 game won on a steal of home
After singling, Bill Dahlen of Chicago's Colts moved around to third on a sacrifice and a throwing error. Dahlen stole home while losing pitcher Cunningham was recovering the ball at the mound.

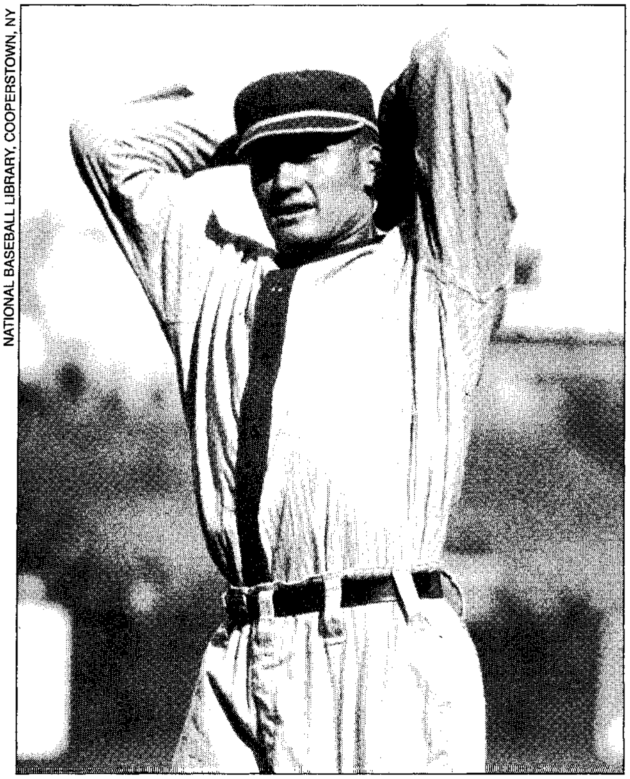
October 13, 1905
McGinnity beats Plank in the first World Series 1-0 game
The Giants' Joe McGinnity shut down Eddie Plank and the Athletics in Game Four, won on an unearned run in the fourth inning. The Giants went on to win the World Series over Connie Mack's A's.

July 24, 1909
Nap Rucker fans 16 in Brooklyn win over St. Louis
Lefty Nap Rucker, who lost 10 1-0 games in a 134-134, 10-year Brooklyn career, also won 11 of them. In this one he fanned 16.

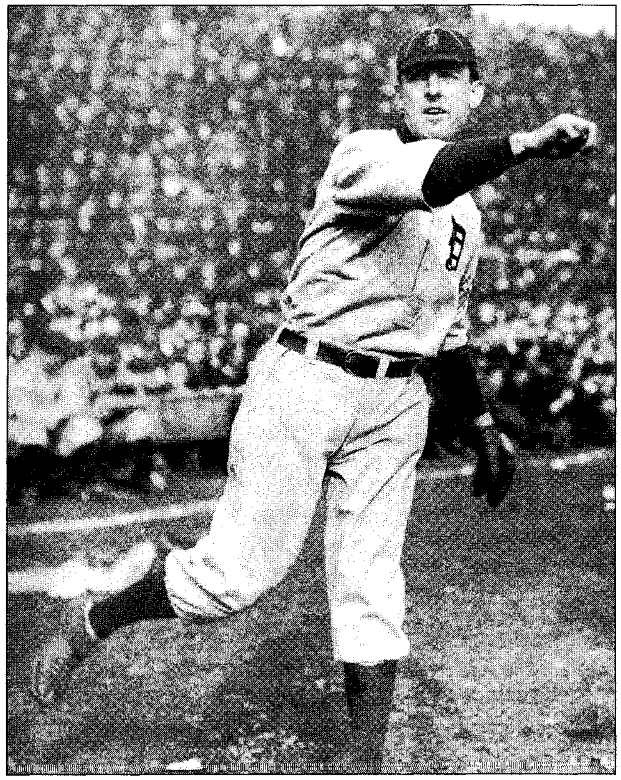
September 22, 1911
Cy Young's final win, number 511
Cy Young's name will dress up any baseball listing, but it is especially noteworthy that the great one's last conquest, number 511, for the Boston Nationals, was a 1-0 gem.

April 14, 1914
The first Federal League 1-0 game
A total of 45 1-0 games were played during the two-year history of the Federal League. The very first of these was Pittsburgh's 1914 season opener. The visiting Brooklyn Tip-Tops, behind Tom Seaton, won. The losing Pittsburgh hurler was Elmer Knetzer. Both were back in the NL for the 1916 season after the Feds called it a day.

May 2, 1918
A double no-no winds up 1-0
Reds lefty Fred Toney and the Cubs' Hippo Vaughn toiled relentlessly through nine innings, neither giving up a bingo. But in the 10th the Reds tallied the unearned run that won it, two hits in that frame helping along. Toney got his no-hitter and big Jim Vaughn a two-hit loss.



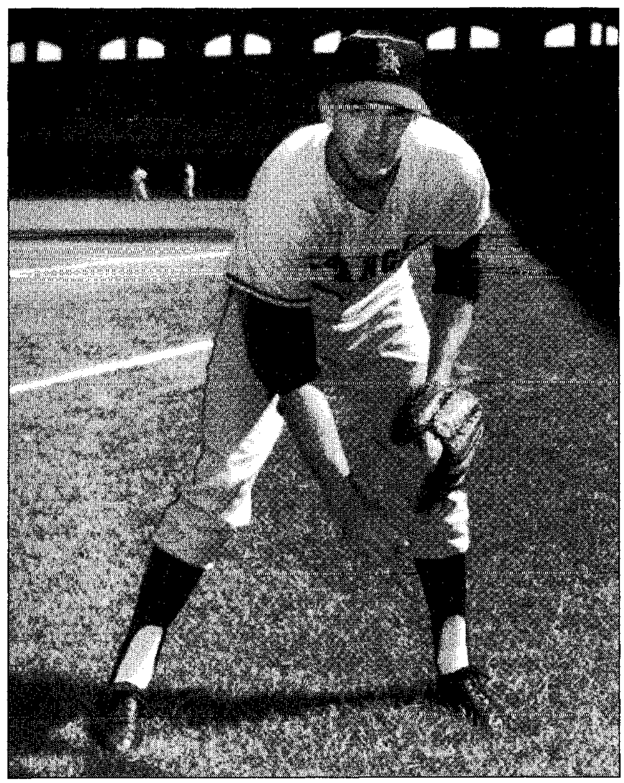
NATIONAL BASEBALL LIBRARY, COOPERSTOWN, NY



NATIONAL BASEBALL LIBRARY, COOPERSTOWN, NY

From top to bottom: Walter Johnson, Bill Donovan, Harvey Haddix, and Dean Chance. Walter Johnson is king of the 1-0 complete games: he holds the record for most wins (38), most losses (26), and most years leading the league in 1–0 wins (8). He set the since-tied mark for most 1–0 wins in a year (5). While pitching for Detroit in 1903, Wild Bill Donovan set the record for most 1–0 losses in a season with 5. It has been tied but never topped. Harvey Haddix lost one of the most famous 1–0 games in history when, in 1959, he pitched perfect ball for 12 innings only to lose in the 13th. Dean Chance is the last pitcher to win five 1–0 complete games in a season, accomplishing that feat in 1964.

NATIONAL BASEBALL LIBRARY, COOPERSTOWN, NY



BRACE PHOTO

May 15, 1918

Walter Johnson and Lefty Williams dueled 18 innings

Two weeks after the Toney-Vaughn spectacular, another Chicago team, the White Sox, were involved in a another tense 1-0 game. This time the crafty little southpaw Claude "Lefty" Williams matched zeroes with Walter Johnson through 17 innings, then gave up the winning run in the 18th on a wild pitch.

July 2, 1933

A 1-0 DH: Hubbell over the Cards' Carleton in 18 innings

In a doubleheader of 1-0 games at the Polo Grounds, Carl Hubbell bested "Tex" Carleton in the first game. The "Meal Ticket" went the route, and Carleton logged 17 innings. St. Louis reliever Jess Haines lost the game in the 18th. In the nightcap Roy Parmalee beat Dizzy Dean 1-0.

April 16, 1940

Bob Feller's Opening Day no-hitter

The only ML no-hitter on an opening day was thrown by "Rapid Robert" Feller at Comiskey Park. Eddie Smith, the Chicago hurler who was victimized by Feller, tossed a six-hitter in the 1-0 loss on a chilly 47° day in the Windy City.

August 6, 1952

46-year-old Satchel Paige beats the Tigers, 1-0

Ancient Leroy Paige, the oldest pitcher to hurl a 1-0 12-inning masterpiece, beat Virgil Trucks, who that same summer authored a pair of 1-0 no-hitters.

September 20, 1958

Wilhelm beats the Yankees

The first pitcher to enter Cooperstown as a reliever, Hoyt Wilhelm no-hit the Yanks in a 1-0 classic. The pitcher he beat, Don Larsen, knew something about no-nos, having tossed the only perfect game in World Series history in 1956.

May 26, 1959

Haddix nearly no-hits the Braves

Harvey Haddix of the Pirates pitched a perfect game against Milwaukee for 12 innings, only to lose in the 13th. An error, an intentional walk to Hank Aaron, and an apparent three-run homer by Joe Adcock ended the game. But Adcock passed Aaron on the basepaths, and both were called out as Mantilla scored. Lew Burdette went all 12 innings, scattering 12 hits. Making Haddix's effort even more remarkable was the fact that the Braves hitter knew what was coming. In 1993, Bob Buhl admitted that the Braves pitchers were stealing the signs from Smoky Burgess, who could not crouch down all the way. They would place a towel on the bullpen fence in such a way to signal fastball or breaking ball.

July 2, 1963

Marichal and Spahn go 16

When asked by his manager if he could go another inning, Juan Marichal said, "If that old guy in the Braves dugout can do it, so can I," and he went on to beat Warren Spahn in a 16-inning thriller on Willie Mays' homer.

September 9, 1965

Koufax wins as only one hit is recorded by both teams

A number of no-hitters have been 1–0 scores, but none featured fewer hits in a game than this gem. Sandy Koufax's record fourth no-hitter was a heartbreaker for Cubs pitcher Bob Hendley, who allowed just one Dodger hit and one unearned run.

July 9, 1968

The NL beats the AL in All-Star play

To date, the only 1-0 game in the history of the All-Star game was played in Houston at the Astrodome. Don Drysdale got credit for the win.

October 5, 1969

The first 1-0 playoff win

MLB moved to its playoff system in 1969. That year Baltimore became the first 1-0 playoff game winner in game two of the best two-of-three series. Dave McNally's complete-game victory in 11 innings beat Minnesota's Twins, and the O's ultimately moved on to the World Series, where the Miracle Mets stunned the baseball world with its victory.

July 1, 1975

Nolan's fourth no-no

Nolan Ryan did his no-hit thing seven times. One of the seven was his 1-0 victory over the Orioles at Anaheim. It was Ryan's 100th career win.

September 21, 1981

Ray Burris beats Steve Carlton in 17 stanzas

North of the border Ol' Lefty and Ray Burris (and several of his stable mates) put on a 17-inning show that netted the Montreal Expos a win despite Carlton's 12 K's and 17 innings of frustration. In that 1-0 loss Lefty did, however, take over the fourth spot on the all-time strikeout list, moving Bob Gibson down to number five.

September 11, 1991

Mercker-Wohlers-Pena combine for a 1-0 no-hitter

In a "first-ever" three Atlanta pitchers combined to no-hit San Diego. The Padres' Greg Harris was the unfortunate victim, and though he pitched well in a complete-game effort, he gave up the winning run in the fifth inning. Mercker got the win, Wohlers, a hold, and Pena, the save.

July 25, 1998

Toronto beats Montreal

In the first 1-0 game in interleague play, the Blue Jays beat the Expos in Montreal on a Woody Williams five-hitter at Olympic Stadium. It was Williams' only complete game of the season.

May 29, 2001

18 innings, 5 hours and 63 minutes, and a 1-0 game

39,709 fans went home bleary-eyed in San Francisco as the Giants and Diamondbacks battled 18 innings and almost six hours before Arizona nudged home the winning run. Each club used seven pitchers, with the win going to Miguel Batista. Ryan Vogelsong, who went the last three innings, took the loss.

BOBBY FONG

Hall of Famers Who Never Played in the World Series

The Chicago Cubs' latest pennant near-miss continues to deny Sammy Sosa, a certain Hall of Famer, an appearance in the World Series. Sammy may yet share the dubious distinction of fellow Cub Ernie Banks, the best-known example of a Hall of Fame player who never played in the World Series. Actually, there have been 31 Cooperstown honorees with playing time since 1903, when the modern World Series was inaugurated, who never participated in the fall classic.

Seventeen of these players had major league experience prior to 1903. The following chart lists them and the number of years they played from 1903 on:

Table 1. Years played, from 1903 on

P	Jack Chesbro	7*
	Addie Joss	8*
	Kid Nichols	3
	Rube Waddell	8*
1B	Jake Beckley	5
	Dan Brouthers	1†
2B	Nap Lajoie	14*
SS	Hughie Jennings	5†
	Bobby Wallace	16*
OF	Jesse Burkett	3
	Ed Delahanty	1
	Hugh Duffy	3†
	Elmer Flick	8*
	Willie Keeler	8
	Joe Kelley	5
	Jim O'Rourke	1†
	Sam Thompson	1†

* Played at least half of major league career from 1903 on
† Made only occasional appearances from 1903 on

Five players were 19th-century stars whose appearances from 1903 on were cursory. Jim O'Rourke and Dan Brouthers played for the 1904 New York Giants with the encouragement of John McGraw, O'Rourke suiting up for one game, Brouthers for two. Similarly, Sam Thompson appeared in eight games for the 1906 Detroit Tigers. Hugh Duffy and Hughie Jennings played occasionally after each had become a coach or manager, Duffy in 34 games over three seasons, and Jennings in 11 games between 1903 and 1918. Jennings did manage the 1907 to 1909 Tigers, who played in and lost three consecutive World Series.

A member of SABR since 1984, **BOBBY FONG** *is president of Butler University, Indianapolis, Indiana.*

Ed Delahanty was a regular with the 1903 Washington club, but his career was cut short by his mysterious death at Niagara Falls in the midst of the season. Kid Nichols, Jake Beckley, Jesse Burkett, and Joe Kelley played three to five seasons from 1903 on, but they were on the downside of their careers and on teams that did not win pennants.

By contrast, six players played at least half of their careers after 1903. Three, Addie Joss, Nap Lajoie, and Elmer Flick, were Cleveland teammates from 1902 to 1910. Despite their presence, the closest the club came to winning a pennant was 1908, when it finished a half-game behind the Tigers because Detroit was not required to make up a rainout.

Jack Chesbro suffered from bad timing: he jumped from the Pittsburgh Pirates to the New York Highlanders (later the Yankees) before the 1903 season, the year that the Pirates won the pennant and played in the first World Series. Then in 1904, Chesbro made the infamous wild pitch that cost the Highlanders a chance at the pennant on the last day of the season. His teammate that day was Wee Willie Keeler, who spent seven of his eight post-1902 seasons with the Highlanders, which would not appear in a World Series until 1921. Bobby Wallace played 16 years for the St. Louis Browns and Cardinals from 1903 on. The Cards played in their first World Series in 1926, the Browns in 1944, long after Wallace had retired.

Perhaps the most agonizing near miss, however, happened to Rube Waddell. He was the ace of the 1905 Philadelphia Athletics staff, going 26-5 to lead them to the American League pennant. Unfortunately, late in the season he got into a scuffle with a teammate and hurt his arm. He did not pitch in the Series as the A's lost to the Giants.

The Hall of Fame credentials of Joss, Lajoie, Flick,

BRACE PHOTO

The 1904 Cleveland Naps featured three Hall of Famers. In the middle row, Lajoie is third from the left, Joss is third from the right, and Flick is first from the right. None of the three ever appeared in the World Series.

Chesbro, Waddell, and Wallace were largely compiled after the commencement of the modern World Series, and thus they represent the first wave of players whose careers were not capped by an appearance in the Fall Classic. More were to come.

Fourteen Hall of Famers played their entire careers in the modern era without appearing in the Series:

Table 2. Years played

P	Jim Bunning	17
	Ferguson Jenkins	19
	Ted Lyons	21
	Phil Niekro	24
	Gaylord Perry	22
C	Rick Ferrell	18
1B	George Sisler	15
2B	Rod Carew	19
3B	George Kell	15
SS	Luke Appling	20
	Ernie Banks	19
OF	Harry Heilmann	17
	Ralph Kiner	10
	Billy Williams	18

Ted Lyons and Luke Appling share with Ernie Banks the distinction of HOFers playing their entire careers with a club that never won the pennant. Lyons and Appling were also longtime teammates on the Chicago White Sox, which went 40 years between World Series appearances.

To date, the Chicago Cubs have gone 58 years since their last fall classic appearance in 1945. That lack of fortune affected not only Banks but also Fergie Jenkins and Billy Williams, longtime Cubs and teammates of Banks. Their trades to other clubs never made up for those years of futility with Chicago.

Similarly, Harry Heilmann spent 15 of his 17 years in the majors with the Detroit Tigers during a period when the club went 25 years between pennants. And Ralph Kiner spent eight seasons of his brief ten-year career with the Pittsburgh Pirates during a period when that franchise went 33 years between pennants.

By contrast, from 1929 to 1947 Rick Ferrell played for the Browns, Red Sox, and Senators. Each club won one pennant during this period, but Ferrell was never on the right team when it cashed in. George Kell played for five teams from 1943 to 1957, but the nearest he ever came to a World Series were three second-place finishes with Detroit, finishing no closer than three games out in 1950.

George Sisler came closest to the Series in 1922, when the St. Louis Browns finished one game behind the New York Yankees. Jim Bunning was a member of the 1964 Philadelphia Phillies, thought to be locks for the pennant until their late-season collapse.

With the introduction of division play in 1969, players like Rod Carew, Phil Niekro, Billy Williams, and Gaylord Perry actually made it into the post-season, but Carew was on the losing side in four League Championship Series, Niekro on two, and Williams and Perry on one apiece.

Gaylord Perry's nearest miss, however, was not in the LCS. In his rookie season of 1962, the San Francisco Giants won the pennant and met the Yankees in the World Series. Gaylord had spent most of the season in the minors before his call-up in September. He contributed three wins to the San Francisco effort that year and played a part in helping the Giants overtake the Dodgers, with whom the Giants finished in a tie at the end of the regular season before beating them in a three-game playoff. But Perry had been called up too late to make the post-season roster! He pitched batting practice during the Series, but he was not eligible to play. Perhaps even Ernie Banks would have preferred not coming *that* close.

A World Series Without Hall-of-Famers?

by Jean-Pierre Caillault

The flip-side of Bobby Fong's article is all of the World Series in which no Hall of Famer participated. The first occurrence of this came in the 1890 Series between the NL champion Brooklyn Bridegrooms and the American Association champion Louisville Colonels. The Players' League champion of 1890, the Boston Reds, were not invited to participate in the Series; if they had, then their Hall of Fame triumvirate of Dan Brouthers, Charlie Radbourn, and King Kelly would have prevented 1890 from having this dubious distinction.

Other World Series with no Hall of Fame players are recent ones in which most players are not yet eligible for election. The Series with the lowest chance of having a participant end up in the Hall was the 1997 edition between the Marlins and Indians. The best candidates from those teams were Gary Sheffield, Manny Ramirez, Jim Thome, and Kevin Brown, none of whom is a certainty.

If we examine World Series Champions only, then the first Championship team not to have anyone in the Hall was the 1981 Dodgers (the 1890 WS ended in a tie, three wins apiece). The 1984 Tigers (Jack Morris?), the 1988 Dodgers (Orel Hershiser?), 1997 Marlins, the 1998 Yankees (Derek Jeter?, Mariano Rivera?), and the 2002 Angels (?) are excellent candidates to join the 1981 Dodgers on this list.

Excluding the most recent Series, there have been 11 Champions with only one player enshrined in Cooperstown – the winners of the very first World Series in 1884, the Providence Grays, with Radbourn as their sole representative; the 1886 St. Louis Browns (Charley Comiskey); the 1919 Reds (Edd Roush); the 1940 Reds (Ernie Lombardi); the 1943 Yankees (Bill Dickey); the 1944 Cardinals (Stan Musial); the 1979 Pirates (Willie Stargell); the 1980 Cardinals (Ozzie Smith); the 1985 Royals (George Brett); the 1986 Mets (Gary Carter); and the 1987 Twins (Kirby Puckett).

The Championship team with the most Hall of Famers was the 1932 Yankees with nine (Earl Combs, Dickey, Lou Gehrig, Lefty Gomez, Tony Lazzeri, Herb Pennock, Red Ruffing, Babe Ruth, and Joe Sewell). There have been seven Champions with six Hall of Famers who participated in the World Series: the 1888 and 1889 New York Giants (Roger Connor, Buck Ewing, Tim Keefe, Jim O'Rourke, John Ward, and Mickey Welch); the 1927 Yankees (Combs, Gehrig, Waite Hoyt, Lazzeri, Pennock, and Ruth), the 1928 Yankees (Combs, Leo Durocher, Gehrig, Hoyt, Lazzeri, and Ruth); the 1934 Cardinals (Dizzy Dean, Durocher, Frankie Frisch, Jesse Haines, Joe Medwick, and Dazzy Vance); and the 1936 and 1937 Yankees (Dickey, Joe DiMaggio, Gehrig, Gomez, Lazzeri, and Ruffing).

The most Hall of Fame players on a World Series losing team was seven, infamously achieved by the 1924 Giants (Frisch, Travis Jackson, George Kelly, Fred Lindstrom, Bill Terry, Hack Wilson, and Ross Youngs).

And the most Cooperstown inductees from both teams in one World Series occurred in 1932 when the Chicago Cubs added four (Kiki Cuyler, Burleigh Grimes, Gabby Hartnett, and Billy Herman) to the Yankees' nine to make a likely never-to-be-broken record of 13.

JEAN-PIERRE CAILLAULT

Hall of Fame Batteries

Mike Piazza of the New York Mets and Ivan Rodriguez of the Florida Marlins, each selected to the All-Star game ten times, are generally acknowledged to be today's best catchers. Both of them seem to be on the road to baseball's Hall of Fame. One question we might ask about Piazza and Rodriguez is: Have either of these two been fortunate enough to catch pitchers who might join them in Cooperstown? Piazza has caught Pedro Martinez, Orel Hershiser, and Tom Glavine, all of whom have reasonably good chances to get inducted; back in the early 1990s Rodriguez caught Nolan Ryan, who already has a plaque in the Hall. So it looks like Piazza and Rodriguez might very well find themselves in familiar company should they end up joining baseball's immortals in Cooperstown. How common is the Hall of Fame battery, though? Do the Piazza and Rodriguez cases stand out as highly unusual or relatively common?

Baseball fans can readily dip into their knowledge of baseball history to name quickly some prominent catcher/pitcher pairs: Johnny Bench and Tom Seaver (Cincinnati); Yogi Berra and Whitey Ford (New York Yankees); Mickey Cochrane and Lefty Grove (Philadelphia Athletics); Bill Dickey and Lefty Gomez (New York Yankees). Those, of course, are the easy ones. But how many other batteries of Hall of Famers can you name?

The Cooperstown inductee who caught the most fellow Hall of Famers was 19th-century standout, Jim "Orator" O'Rourke, who caught seven of them, despite the fact that O'Rourke caught only 209 games in his career (he was primarily an outfielder, playing on the grass in 1,377 games). The Hall of Famers who pitched to O'Rourke were Pud Galvin (Buffalo), Monte Ward (Providence), Tim Keefe, Amos Rusie, Mickey Welch (all three with New York), fellow New York Giants catcher Buck Ewing (for one game in 1885), and "Iron Man" Joe McGinnity, whom O'Rourke caught in the only Giants game of 1904 in which O'Rourke appeared—when he was almost 54 years old!

JEAN-PIERRE CAILLAULT *is a Professor of Astronomy at the University of Georgia. He has published articles in* Baseball Digest *and the* Baseball Research Journal *and is the author of* A Tale of Four Cities *and the forthcoming* New York Clipper Biographies: The Complete Collection..

Three catchers were fortunate enough to be the receivers for five Hall of Fame pitchers. One of those three was Ewing, who, like O'Rourke, also caught Keefe and Ward (while with both Troy and New York) and Rusie and Welch (New York). The fifth Hall of Famer to pitch to Ewing was John Clarkson, for only one Cleveland game in 1893 (in that game, incidentally, Ewing moved to right field after the second inning; Cy Young came in to pitch in the third inning, so Ewing just missed having caught a sixth Hall of Famer!).

Another 19th-century receiver to have caught five Hall of Famers was Mike "King" Kelly, who, like Ewing, also caught Rusie (New York) and Clarkson (Chicago and Boston). The other immortals who pitched to Kelly were Kid Nichols and "Old Hoss" Radbourn (both with Boston), and, amazingly, Hall of Fame first baseman Cap Anson for one game in 1884 while both played for Chicago. The third Hall of Fame catcher to handle five Cooperstown pitchers was Bill Dickey, who caught Lefty Gomez, Burleigh Grimes, Waite Hoyt, Herb Pennock, and Red Ruffing, all while with the Yankees.

There have been five Cooperstown catchers who caught four fellow Hall of Famers: Roger Bresnahan, Ray Schalk, Al Lopez, Ernie Lombardi, and Carlton Fisk. Bresnahan caught McGinnity (Baltimore and New York), Vic Willis (St. Louis), and the great Christy Mathewson (New York), as well as Rube Marquard for one game in the Giants' memorable 1908 season (in which they lost the NL pennant to the Chicago Cubs on Fred Merkle's "boner"). Schalk was behind the plate for White Sox teammates Red Faber, Ted Lyons, and Ed Walsh, and for Carl Hubbell for one game in 1929 when Schalk moved to the Giants for the final year of his career. Schalk just missed adding a fifth Hall of Fame pitcher to his list when he was removed from a game in 1925 before fellow Cooperstown enshrinee


YOGI BERRA MUSEUM AND LEARNING CENTER

Yogi and Whitey

"Chief" Bender came in to pitch the ninth inning (it was Bender's first and only appearance since 1917). The Hall of Fame pitchers who pitched to Lopez were Hoyt and Dazzy Vance (when Lopez was with Brooklyn), and Bob Feller and Bob Lemon, who were Lopez's teammates with the Cleveland Indians in 1947. Lombardi was on the receiving end of pitches from Hubbell (New York), Vance (Brooklyn), Eppa Rixey (Cincinnati), and Warren Spahn, but in Spahn's case, it was for only one inning in a Boston Braves game against the Dodgers, during Spahn's rookie year in 1942. The last Hall of Fame catcher who caught four Hall of Fame pitchers was Fisk, who had the good fortune to catch Ferguson Jenkins and Juan Marichal while with the Red Sox, and Seaver and Carlton while with the White Sox.

One of the catchers who caught three Hall of Fame pitchers was 19th-century Chicago star Anson, normally a first baseman, caught Al Spalding, John Clarkson, and Clark Griffith. Another catcher who made up the receiving half of a Hall of Fame battery for three pitchers was Wilbert Robinson, more famous as the Brooklyn Robins manager than as a turn-of-the-century catcher, but who caught Joe McGinnity (Baltimore), Cy Young (St. Louis), and young Roger Bresnahan, who, while with the Baltimore entry in the new American League of 1901, started out as a pitcher before himself moving behind the plate. Gabby Hartnett was another Cooperstown inductee who was lucky enough to catch three immortals: Grover Cleveland Alexander, Dizzy Dean, and Burleigh Grimes, all while with the Chicago Cubs. Hartnett missed an opportunity for a fourth, Carl Hubbell, when the two played for the Giants in 1941. The two caught and pitched on the same day four times in '41, but in each case it was in different games of a doubleheader. Rick Ferrell caught Lefty Grove and Herb Pennock while with the Red Sox and Early Wynn when Ferrell played for the Senators. The last of the five catchers who caught a trio of Hall of Fame pitchers was Brooklyn Dodger great Roy Campanella, who caught two star pitchers before they became stars (Sandy Koufax and Don Drysdale) and one Hall of Famer who ended up in the Hall because of his managerial skills, not his pitching (Tommy Lasorda).

Two catchers caught two fellow inductees: Chicago Cubs star Frank Chance, who began his career as a catcher before moving to first base, caught Clark Griffith and Rube Waddell, and Mickey Cochrane, who caught Lefty Grove and Waite Hoyt while playing for the Philadelphia Athletics. The four remaining Hall of Fame catchers who caught other Hall of Famers are Connie Mack, much more famous as a manager than as a catcher (he caught Pud Galvin while with Pittsburgh); Jimmie Foxx, who caught Lefty Grove with both the Athletics and the Red Sox; and the two mentioned at the beginning of the article, Yogi Berra of the Yankees (who caught Whitey Ford) and Johnny Bench of the Reds (who caught Tom Seaver).

The accompanying table lists all the batteries in baseball history that featured a Hall of Famer catching pitches thrown by another Hall of Famer, even if it was only for a single inning. The list contains a total of 20 different catchers and 45 different pitchers, who combined for a total of 65 Hall of Fame batteries.

If things had turned out slightly differently, there would have been a few more batteries added to the list on the following page, but instead these pairs will have to remain classified as near-misses. The close calls of Buck Ewing-Cy Young and Ray Schalk-Chief Bender were mentioned earlier, but there were three other such near-misses in baseball history. In the last game of the 1898 season, Hall of Fame slugger Hugh Duffy went behind the plate for a few innings in the middle of the game, but returned to the outfield before another Cooperstown inductee, Kid Nichols, came on in relief in the seventh inning. Hall of Fame shortstop and manager Lou Boudreau took over the catching duties of the Cleveland Indians for the final two innings of a game in the Indians' championship season of 1948, right after pitcher Bob Lemon had been lifted for a pinch-hitter in the top of the eighth inning. The most remarkable near-miss occurred in a game in August 1940, when the "Splendid Splinter," Ted Williams, pitched the final two innings of a game against Detroit, a game in which fellow Hall of Fame member of the 500-home run club, Jimmie Foxx, had earlier caught a few innings. A Foxx-Williams battery, now that would have been something!

Table 1. Hall of Fame batteries

CATCHER	G	PITCHER	TEAM
JIM O'ROURKE	7	Pud Galvin	Buffalo
		Monte Ward	Providence
		Buck Ewing	NY-N
		Tim Keefe	NY-N
		Mickey Welch	NY-N
		Amos Rusie	NY-N
		Joe McGinnity	NY-N
KING KELLY	5	Cap Anson	CHI-N
		John Clarkson	CHI-N/BOS-N
		Hoss Radbourn	BOS-N
		Kid Nichols	BOS-N
		Amos Rusie	NY-N
BUCK EWING	5	Tim Keefe	Troy/NY-N
		Monte Ward	Troy/NY-N
		Mickey Welch	NY-N
		Amos Rusie	NY-N
		John Clarkson	CLE (Spiders)
BILL DICKEY	5	Waite Hoyt	NY-A
		Herb Pennock	NY-A
		Lefty Gomez	NY-A
		Red Ruffing	NY-A
		Burleigh Grimes	NY-A
ROGER BRESNAHAN	4	Joe McGinnity	Balt./NY-N
		Christy Mathewson	NY-N
		Rube Marquard	NY-N
		Vic Willis	STL-N
RAY SCHALK	4	Ed Walsh	CHI-A
		Red Faber	CHI-A
		Ted Lyons	CHI-A
		Carl Hubbell	NY-N
AL LOPEZ	4	Dazzy Vance	BKN
		Waite Hoyt	BKN
		Bob Feller	CLE
		Bob Lemon	CLE
ERNIE LOMBARDI	4	Dazzy Vance	BKN
		Eppa Rixey	CIN
		Warren Spahn	BOS-N
		Carl Hubbell	NY-N
CARLTON FISK	4	Juan Marichal	BOS-A
		Ferguson Jenkins	BOS-A
		Tom Seaver	CHI-A
		Steve Carlton	CHI-A
CAP ANSON	3	Al Spalding	CHI-N
		John Clarkson	CHI-N
		Clark Griffith	CHI-N
WILBERT ROBINSON	3	Joe McGinnity	Balt.
		Cy Young	STL-N
		Roger Bresnahan	Balt.
GABBY HARTNETT	3	Pete Alexander	CHI-N
		Burleigh Grimes	CHI-N
		Dizzy Dean	CHI-N
RICK FERRELL	3	Lefty Grove	BOS-A
		Herb Pennock	BOS-A
		Early Wynn	WAS-A
ROY CAMPANELLA	3	Tommy Lasorda	BKN
		Sandy Koufax	BKN
		Don Drysdale	BKN
FRANK CHANCE	2	Clark Griffith	CHI-N
		Rube Waddell	CHI-N
MICKEY COCHRANE	2	Lefty Grove	PHI-A
		Waite Hoyt	PHI-A
CONNIE MACK	1	Pud Galvin	PIT
JIMMIE FOXX	1	Lefty Grove	PHI-A/BOS-A
YOGI BERRA	1	Whitey Ford	NY-A
JOHNNY BENCH	1	Tom Seaver	CIN

CYRIL MORONG

Historical Trends in Home-Field Advantage

From 1901 to 2002, the average seasonal difference between a team's home winning percentage and its road winning percentage was .082.[1] But has it changed much over the last 100 years and has the change been significant? What teams have enjoyed an especially good home-field advantage?

Table 1. Changes in average home-field advantage, 1901-2002

DECADE	ADVANTAGE	XHW/SEASON
1901-10	0.1042	4.01
1911-20	0.0740	2.85
1921-30	0.0928	3.57
1931-40	0.0976	3.76
1941-50	0.0922	3.55
1951-60	0.0784	3.02
1961-70	0.0796	3.22
1971-80	0.0752	3.05
1981-90	0.0800	3.24
1991-2002	0.0700	2.84

There is a slight downward trend over the century, but quite a bit of year-to-year variance. The average home advantage in 1931 was .164, and then fell rapidly and dramatically afterward. Then there is the huge spike to .146 in 1978, the biggest advantage since 1931. A decade-by-decade summary might be a little more interesting. This is given in Table 1[2] which clearly shows that the home-field advantage has declined over time. Yet the second decade has the second lowest average. So there is a trend, but anomalies as well. Has the change been statistically significant? Yes.[3] But is it significant in a baseball sense? When the average advantage is .103, it means a home winning percentage of about .552. When the advantage is .07, it means a home winning percentage of .535. This difference over 81 games is 1.36 wins. I leave it to each reader to decide if that is significant in a baseball sense. The third column projects how many more games a team would win at home above an even split. For example, in the first decade an even split would give a home team 38 wins. But with a .1042 difference between their home and road percentages (a .552 home winning percentage), they would win 42.51 games at home. Decades from the 1960s on use 81 games.

Homestands are not as long as they used to be, and teams now travel by plane. This might account for the historical trend. But notice that the difference between the 1940s and 1950s is not too great and that the 1980s was higher than the 1970s. The average from 1901 to 1950 was .091, and since it has been .076. This seems like a small difference.

Table 2. The best teams in each decade. A team had to play at least five years in the decade to qualify.*

DECADE	LEADER	ADVANTAGE	XHW/SEASON
1901-10	Phila. (AL)	0.2289	8.81
1911-20	Brooklyn	0.0936	3.60
1921-30	Cincinnati	0.1277	4.92
1931-40	Cleveland	0.1183	4.55
1941-50	Boston (AL)	0.1947	7.50
1951-60	Cincinnati	0.1113	4.29
1961-70	Houston	0.1895	7.67
1971-80	Houston	0.1202	4.87
1981-90	Minnesota	0.1731	7.01
1991-2002	Colorado	0.1658	6.71

Table 3. The best ten teams since 1901 that existed in one city for at least ten years. Changes in ballparks for these teams are not taken into account. Assumes an 81-game home schedule for all teams.

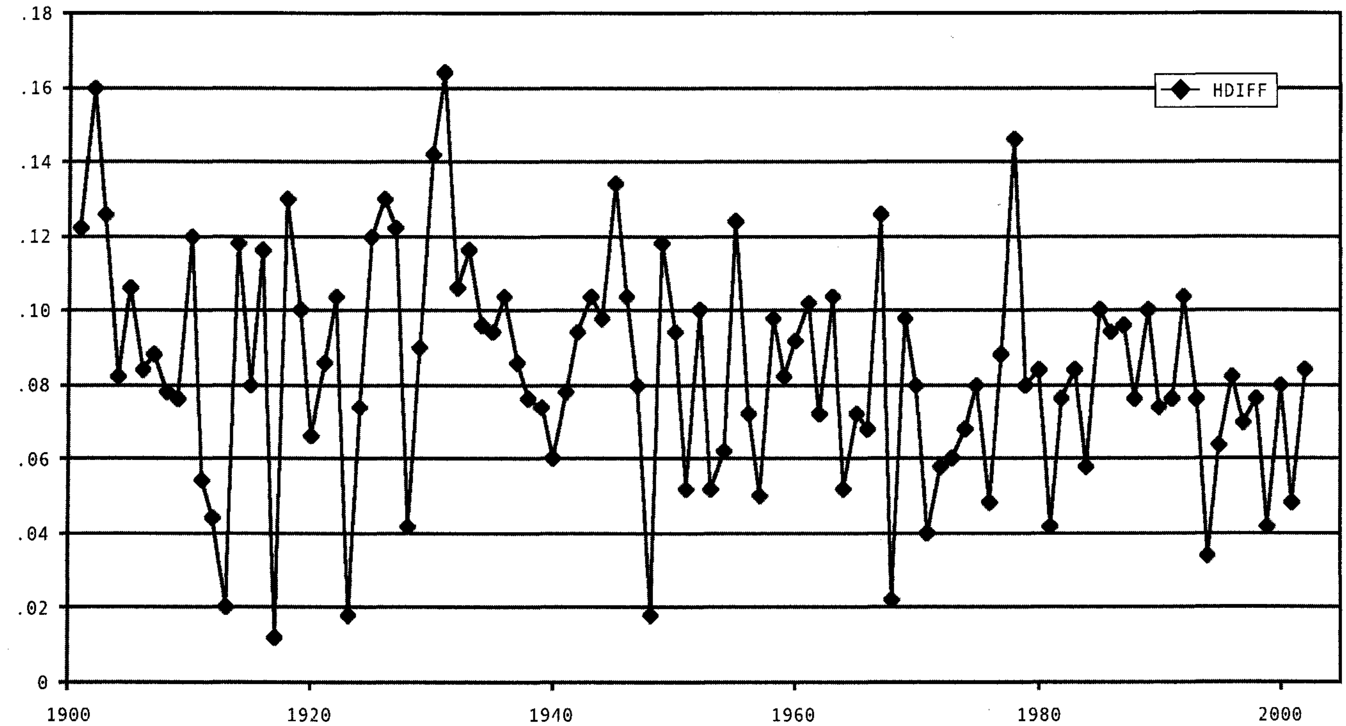
DECADE	ADVANTAGE	XHW/SEASON
Rockies	0.1433	5.52
Marlins	0.1216	4.68
Astros	0.1190	4.58
A's-Phi.	0.1077	4.15
Red Sox	0.1045	4.02
Browns	0.0942	3.63
1st Senators	0.0911	3.69
Braves-Bos.	0.0910	3.69
Giants-SF	0.0900	3.65
Yankees	0.0881	3.57

The lowest home advantage belongs to the Baltimore Orioles, at .051.

Has a big home-field advantage been an aid? Not really. The overall winning percentage of the teams with the 25 highest single-season home advantages is

CY MORONG *is a professor of economics at San Antonio College in San Antonio, TX. He is originally from Chicago and is a life-long White Sox fan.*

Figure 1. Average yearly home advantage.



.492. For the 25 worst teams it is .498. The only one of the best 25 to win a pennant was the 1987 Twins.

Table 4. The best 10 home-field advantages since 1901, using the 77 home game and 81 home game schedules where appropriate.

TEAM	YEAR	ADVANTAGE	XHW/SEASON
Phila. (AL)	1945	0.356	13.71
Phila. (AL)	1902	0.338	13.01
Boston (AL)	1949	0.337	12.97
Colorado	1996	0.333	13.49
Minnesota	1987	0.333	13.49
Houston	1978	0.321	13.00
Phila. (AL)	1908	0.319	12.28
Chicago (AL)	1903	0.315	12.13
Chicago (AL)	1902	0.312	12.01
Boston (AL)	1952	0.311	11.97

Despite the trend toward a lower home field advantage, 2003 saw a fairly big one. The average home field advantage was .099. There were also seven teams that had at least a .160 advantage, which is just about twice the historical average. With 23.33% of the teams being above the .160 mark, this is the highest percentage since 1986, when seven of the 26 teams were above .160. This past season was the 25th highest percentage of teams going above .160.

Table 5. Home advantage for all teams in 2003.

TEAM	ADVANTAGE	HOME W	HOME L	ROAD W	ROAD L
Colorado	0.296	49	32	25	56
Montreal	0.259	52	29	31	50
Oakland	0.222	57	24	39	42
Chi White Sox	0.198	51	30	35	46
Texas	0.185	43	38	28	53
Florida	0.185	53	28	38	43
San Francisco	0.166	57	24	43	37
Anaheim	0.149	45	37	32	48
Philadelphia	0.148	49	32	37	44
Boston	0.136	53	28	42	39
St. Louis	0.136	48	33	37	44
Baltimore	0.122	40	40	31	51
Atlanta	0.111	55	26	46	35
Tampa Bay	0.111	36	45	27	54
Houston	0.111	48	33	39	42
Cleveland	0.099	38	43	30	51
Los Angeles	0.086	46	35	39	42
Seattle	0.086	50	31	43	38
Arizona	0.074	45	36	39	42
Minnesota	0.074	48	33	42	39
San Diego	0.074	35	46	29	52
Detroit	0.037	23	58	20	61
Pittsburgh	0.037	39	42	36	45
NY Mets	0.030	34	46	32	49
Cincinnati	0.012	35	46	34	47
Chi Cubs	0.000	44	37	44	37
Kansas City	-0.024	40	40	43	39
NY Yankees	-0.028	50	32	51	29
Toronto	-0.049	41	40	45	36
Milwaukee	-0.074	31	50	37	44

TREY KINZER, NATIONAL TELEPRINTING, INC.

The Rockies have the greatest home-field advantage of any team, playing in 1993-94 at Mile High Stadium and, since 1995, at Coors Field.

NOTES

1. This is just a simple average, adding up every team's single-season advantage and dividing by the number of team seasons. Season with longer schedules are not given extra weight.
2. Again, this is just a simple average. The overall home field advantage is added up for each of the 10 seasons and then is divided by 10. Seasons with more teams or games don't get extra weight. The low figure for the 1911-20 period is not affected much by the Federal League, which had about a .084 advantage during its two seasons.
3. Using the means test, the difference between the first decade and the last is significant, with a z-score of 3.97. Also, here is the standard deviation for all teams for each decade: 0.0914, 0.0848, 0.0769, 0.0787, 0.0857, 0.0855, 0.0831, 0.0837, 0.0773, 0.0810. This shows that the dispersion in home field advantage across teams has not changed much since 1901. The correlation between year and the average yearly home advantage is -.247. It has a t-value of -2.55, which is statistically significant.
4. The advantage listed for teams that changed parks in their respective decade only includes data from the park they played the most seasons in. Only one of those teams, the Indians of the 1930s, actually did not have the biggest advantage when only their most commonly used stadium was considered. The Reds would then actually be a little higher, at .1188. The following are the teams that changed parks and their home advantage for the entire decade: Philadelphia (01-10), .2021; Cleveland (31-40), .1234; Houston (61-70), .1783; Minnesota (81-90), .1438; Colorado (91-2002), .1433. For all of them, except Cleveland, they would still have the highest yearly average even if all years of the decade are used.

BOB McCONNELL

Career .300 Batting Averages

The magic number for a batting average is .300. When a player hits .300 or better, he has had a good season. A .299 average just doesn't look as good. Many record books list players with a career average of .300, but they usually limit their lists to players with 1,000 or more hits. What about the .300 hitters with less than 1,000 hits? You would think that a player who can maintain a .300 average should be able to stick around long enough to accumulate 1,000 hits. Let us look at these in greater detail.

The following table is broken down into nine groups: from players with 900-999 hits down to players with 100-199 hits. Obviously, a player with 100 hits should not be bracketed with a player with over 900 hits.

The best explanation for not sticking around is death. Three players died during their active major league careers. Austin McHenry died of a brain tumor shortly after the end of the 1922 season. Willard Hershberger, playing for the Cincinnati Reds, committed suicide in August 1940, while the Reds were making a successful run for the National League pennant. Lyman Bostock was killed in 1978 when he was accidentally shot while riding in a car. The shot was meant for one of the other passengers.

There are nine active players on the list; some of these players would go on to bang out 1,000 hits while maintaining a .300 average and thus get into the record books. Others would fall below .300 as their careers wound down. Duke Snider dropped from .300 to .295 during his last two seasons. Mickey Mantle slipped from .302 to .298 during his last season.

Six players—Ross Barnes, George Hall, Dick Higham, Cal McVey, Levi Meyerle, and Lipman Pike—played in the National League in 1876. Many historians consider this as the first major league season. All six of these players had played for five years in the National Association of 1871-75 and had played on independent clubs prior to that. Barnes had an overall average of .390 in the NA, Meyerle .365, McVey .362, and Pike .332. It can be said that the careers of these players were already on the way down. George Hall was banned from baseball after the 1877 season for throwing games.

For the 1887 season, bases on balls were counted as hits, which really inflated batting averages. Myron Allen, Bob Caruthers, Otto Schomberg, and Ed Swartwood all played in 1887, and they greatly benefited from this rule. By recalculating their career averages by taking away their hits as a result of bases on balls, the averages for all four players drop below .300 (Swartwood to .2994).

The Federal League of 1914-15 is listed in most baseball books as a major league. However, it was a notch below the other two major leagues of its time in the quality of play. Benny Kauff led the league in batting for both years of its existence with averages of .370 and .342. Benny was called the "Ty Cobb of the Federal League." Without his Federal League numbers, Kauff's major league average was .287. Ted Easterly and Bill Kenworthy would also slip below .300 without the benefit of their Federal League stats. A fourth Federal Leaguer, Vin Campbell, is an interesting character. He played in the National League for several years prior to jumping to the FL. His overall average in the NL was .306. He had a good rookie year with Pittsburgh in 1910, but then quit to enter the brokerage business in St. Louis. He had a change of heart and rejoined Pittsburgh in the middle of the 1911 season. The Pirates traded him to the Boston Braves, and he had another good season in 1912. He refused to report in 1913, and it is not known how he spent that summer. In 1914 he signed a three-year contract for $25,000 with the Indianapolis club in the Federal League. The club moved to Newark for the 1915 season, and the league folded before the 1916 campaign. As part of the peace agreement Newark owner Harry Sinclair was allowed to sell many of his players to Organized Baseball clubs.

BOB McCONNELL *lives in Wimington, Delaware, and is a founding member of SABR. He was the first recipient of the Bob Davids Award, SABR's highest honor.*

Campbell was sold to the St. Louis Browns in late February, but he never reported. Instead, he went into the auto business in Pittsburgh. He sued Sinclair for his 1916 salary, and collected.

As all fans know, the level of play really dropped during World War II. Two players on the list played during the war. John Bolling hit .351 in 1944 after hitting .289 in 1939, his only other season in the majors. Augie Bergamo hit .286 in 1944 and .316 in 1945 for an overall average of .304 for his two major league seasons. Bob Dillinger had a very good season at Toledo (American Association) in 1942. He then spent the next three years in military service before going up to the St. Louis Browns in 1946. He surely would have garnered the 112 hits that he needed to push him over the 1,000 mark had it not been for the war. Joe Harris, with 963 hits, certainly would have reached the 1,000 level, but for the fact that he spent 1918 in military service.

There are four pitchers on the list. Three of them—Bob Caruthers, Tom Parrott, and Walter Thornton—played during the 19th century. Many pitchers in those days filled in as position players from time to time. This was the case with Caruthers, Parrot, and Thornton. However, they were getting paid to pitch, and their stay in the majors depended on how well they performed on the mound. Caruthers did stick around long enough to win 218 games. Erv Brame, the fourth pitcher, played in the 20th century and was used only as a pitcher.

Catchers do get into many more games than pitchers, but still not as many as other position players. Two catchers on the list—Willard Hershberger and Ted Easterly—have already been mentioned. Four additional catchers—John Bassler, Bubbles Hargrave, Babe Phelps, and Earl Smith—all played for at least nine years in the majors. They just didn't get into enough games to reach the 1,000 mark.

Three all-time minor league greats—Buzz Arlett, Ike Boone, and Smead Jolley—had brief stays in the majors. Arlett had a minor league career batting average of .341 with 2,726 hits, 598 doubles, 432 homers, and 1,786 RBI. Boone hit .370 with 2,521 hits, 477 doubles, 128 triples, 217 homers, and 1,334 RBI. Jolley hit .366 with 3,037 hits, 636 doubles, 336 homers, and 1,593 RBI. These players were stuck with the good-hit, no-field label. It is hard to imagine them being that bad as fielders. After all, Zeke Bonura (who was a notoriously poor fielder) lasted long enough in the majors to collect 1,000 hits. Dick Porter spent eight years with Baltimore of the International League (1921-28) before reaching the majors. During that period, Baltimore players were not subject to the major league draft and the club held back a number of good players. Other players who, for some reason, took a long time to reach the majors were Eddie Brown, John Frederick, Ben Paschal, Lance Richbourg, and Earl Webb.

Jay Kirke was an interesting and, some say, eccentric person. He played for Joe McCarthy in Louisville and Joe loved to tell funny stories about him. Kirke played in the minors for 21+ years in addition to one full season and parts of six others in the majors. An anonymous quote might explain one reason why Kirke didn't stay longer: "He can hit, but as a fielder he can only retrieve."

Emmet Hendrick quit baseball to go to work on the railroad. By coincidence, his brother was the president of the railroad. Could there have been a salary increase involved?

A number of other players on the list had long and successful minor league careers. Were they all bad fielders? Among other long-time minor league players on the list are Jim Bannon, Del Bissonette, Pat Duncan, George Fisher, Bill Keister, Bill Lamar, Cliff Lee, Jack Lelivelt, Fred Nicholson, and Babe Twombly. Was Keister a bad fielder?; the answer is yes. Bill holds the major league record for lowest fielding average for a shortstop in 100 games or more games with a mark of .861. Playing for Baltimore in 1901, he made 88 errors in 114 games.

Maurice Archdeacon is another interesting story. Ty Cobb had scouted him and reported that he would never hit in the majors. However, the White Sox paid a hefty price for him. He went up to the Sox in late 1923 and hit .401 in 22 games. Johnny Mostil beat him out of the center field job the following year, but missed a number of games due to illness and injuries. This gave Archdeacon playing time, and by August 1 he was hitting .386. His career average at that point was .391, possibly the best start for any player in history. But that was his high point. He hit .185 for the balance of his major league career, and the White Sox practically gave him to Baltimore in early 1925. Archdeacon's main asset was speed, although he was not a great base stealer. He beat out many grounders, and infielders probably would have learned how to play him.

Then there are those players who were asked the question, "Yeah, but what have you done recently?" These players started their major league careers strongly, but slumped during their last year or two and earned a one-way ticket back to the minors. Tuck Turner hit .418 in his second season in the majors. He hit .243, .291, and .199 in his last three seasons. Other players who had a poor final season were Charlie Hollocher, Sam Leslie, Dusty Miller, Ed Morgan, Fred Nicholson, Ernie Orsatti, Harlan Poole, and George Stone. Hollocher missed a great deal of time due to various illnesses. He had a reputation of being a hypochondriac and he eventually committed suicide.

Louis Sockalexis (what a great name for a slugger) is a sad case. He was a Penobscot Indian from Old Town, Maine, and he starred in baseball at Holy Cross College. With much fanfare, he went directly to the majors with Cleveland in 1897. He proceeded to hit .338 in 66 games. He soon began to drink heavily. The club put up with his problem for two more years, but finally had to let him go. He died at the early age of 42.

Harry Moore was a mystery player until recently. He played a full schedule for Washington in the Union Association of 1884. Yet the record books have no biographical data on him. Two researchers have been working on him and have uncovered a great deal of information. His correct name is Henry Scott Moore and he was born around 1862 in California, probably in San Francisco, where he spent his early childhood. He started his pro career with Reading of the Interstate Association in 1883. After his stint with Washington in 1884, he played for Washington and Norfolk of the Eastern League in 1885. Other stops in the minors were at Atlanta, Topeka, Sacramento, Stockton, and San Francisco.

Cuckoo Christensen and Glass Arm Brown made the list. Is there any connection between their nicknames and the fact that they didn't last long in the majors?

The following are interesting bits of information about players on the list:

- Tuck Turner played in the great Philadelphia outfield of 1894 (Turner .418, Sam Thompson .415, Ed Delahanty .404, Billy Hamilton .403).

- Red Wingo, the brother of Ivy, played a career-high 130 games in the Detroit Tigers outfield of 1925 (Harry Heilmann .393, Ty Cobb .378, Wingo .370, Bob Fothergill .353).

- Ross Barnes led the NL in batting in 1876, George Stone led the AL in 1906, Benny Kauff led the FL in 1914 and 1915, Bubbles Hargrave led the NL in 1926, and Dale Alexander led the AL in 1932.

- Pat Duncan was one of the stars of the 1919 World Series, and Joe Harris starred in the 1925 Series.

- Oscar Ray Grimes had a twin brother, Roy, who played in the majors. His son Oscar Jr. also played in the majors.

- Earl Webb holds the major league record for most doubles in a season with 67, set in 1931.

There are, no doubt, stories to be told about the other players on the list. Maybe SABR members can dig up some of them.

BRACE PHOTO

Ike Boone played parts of eight seasons in the major leagues, hitting .321, but he did even better over 17 years in the minors, posting a career .370. His best year was 1930 when he hit .448 for Mission in the Pacific Coast League. The three Boone brothers—Ike, Danny, and Bill—combined to hit a career .361 in the minors.

Table 1. Players hitting .300 with fewer than 1,000 at-bats

900–999 HITS	AVG	H	PLAY SPAN
Bill Everitt	.317	902	1895-1901
Joe Harris	.317	963	1914-28
Benny Kauff	.311	961	1912-20
Ed Swartwood	.310	907	1881-92
John Frederick	.308	954	1929-34
John Moore	.307	926	1928-45
JOSE VIDRO	.306	940	1997-2003
Homer Summa	.302	905	1920-30
George Stone	.301	984	1903-10
Emmet Hendrick	.300	914	1898-1908

800–899 HITS	AVG	H	PLAY SPAN
Dale Alexander	.331	811	1929-33
Ed Morgan	.313	879	1928-34
John Hodapp	.311	880	1925-33
Harvey Hendrick	.308	896	1923-34
Lance Richbourg	.308	806	1921-32
Pat Duncan	.307	827	1915-24
Bob Dillinger	.306	888	1946-51
Charlie Hollocher	.304	894	1918-24
Eddie Brown	.303	878	1920-28
Sam Hale	.302	880	1920-30
SEAN CASEY	.300	875	1997-2003

700–799 HITS	AVG	H	PLAY SPAN
Bill Keister	.312	758	1896-1903
Bubbles Hargrave	.310	786	1913-30
Dick Porter	.308	774	1929-34
Sam Leslie	.304	749	1929-38
John Bassler	.304	704	1913-27
Dusty Miller	.301	771	1889-99
Bob Caruthers	.301	761	1884-93
Ray Blades	.301	726	1922-32

600–699 HITS	AVG	H	PLAY SPAN
ICHIRO SUZUKI	.328	662	2001-03
Lyman Bostock	.311	624	1975-78
Babe Phelps	.310	657	1931-42
Vin Campbell	.310	642	1908-15
Bill Lamar	.310	633	1917-27
JUAN PIERRE	.307	638	2000-03
Ernie Orsatti	.306	663	1927-35
Earl Webb	.306	661	1925-33
Del Bissonette	.305	699	1928-33
Earl Smith	.303	686	1919-30
Ted Easterly	.300	607	1909-15
LANCE BERKMAN	.300	642	1999-2003
Reggie Jefferson	.300	637	1991-99

500–599 HITS	AVG	H	PLAY SPAN
ALBERT PUJOLS	.334	591	2001-03
Ray Grimes	.329	505	1920-26
Smead Jolley	.305	521	1930-33
Evar Swanson	.303	573	1929-34
Austin McHenry	.302	592	1918-22

400–499 HITS	AVG	H	PLAY SPAN
Tuck Turner	.320	482	1893-98
Jimmy Bannon	.320	460	1893-96
Red Wingo	.308	409	1919-28
Spud Johnson	.302	400	1889-91
ALEX PIERZYNSKI	.301	430	1998-2003
Sam Dungan	.301	464	1892-1901
Cliff Lee	.300	475	1919-26

300–399 HITS	AVG	H	PLAY SPAN
Cal McVey	.327	393	1876-79
Hack Miller	.322	387	1916-25
Ike Boone	.321	372	1922-32
Ross Barnes	.313	329	1876-81
Bill Kenworthy	.304	301	1912-17
Jay Kirke	.301	346	1910-18
Tom Parrott	.301	301	1893-96
Jack Lelivelt	.301	347	1909-14

200–299 HITS	AVG	H	PLAY SPAN
Otto Schomberg	.331	276	1886-88
Dick Cox	.314	261	1925-26
Fred Nicholson	.311	247	1917-22
Ben Paschal	.309	243	1915-29
James Burns	.305	222	1888-91
Walter French	.303	297	1923-24
Lipman Pike	.301	223	1876-87

100–199 HITS	AVG	H	PLAY SPAN
George Hall	.339	185	1876-77
Harry Moore	.336	155	1884
George Fisher	.335	114	1923-32
M'ce. Archdeacon	.333	128	1923-25
Levi Meyerle	.326	123	1876-84
Dick Higham	.322	193	1876-80
Doc Prothro	.318	191	1920-26
Marv Rackley	.317	151	1947-50
W'd. Hershberger	.316	127	1938-40
Chicken Hawks	.316	124	1921-25
Wall Christensen	.315	162	1926-27
Buzz Arlett	.313	131	1931
Louis Sockalexis	.313	115	1897-99
John Bolling	.313	107	1939-44
Walter Thornton	.312	162	1895-98
Roy Carlyle	.312	157	1925-26
Irv Waldron	.311	186	1901
John Sullivan	.309	153	1920-21
Joe Knight	.309	156	1884-90
SCOTT PODSEDNIK	.308	180	2001-03
Erv Brame	.306	121	1928-32
Babe Twombly	.304	109	1920-21
Augie Bergamo	.304	151	1944-45
Harlin Pool	.303	129	1934-35
Pete Scott	.303	158	1926-28
ALEX CINTRON	.302	160	2001-03
Myron Allen	.301	193	1883-88

Players in CAPS were active during the 2003 season. Where two players have the same batting average, the averages were carried out to four or more places to determine the player's slots on the list. Play Span indicates the first and last season in the major leagues. The player did not necessarily play in the majors during all of the intervening seasons. Statistical Sources: *Total Baseball,* 7th ed., *Baseball Register,* 2003 ed., MLB.com.

STEPHEN BOREN, JAMES SMITH JR., AND HERM KRABBENHOFT

Who Made the Most Triple Plays?

The Detroit Tigers take credit for having made the most triple plays.[1,2] But is this correct? Actually, it is a very complicated question, and the correct answer depends on multiple definitions and positions taken. The complex answer to this seemingly simple question depends on three further questions.

First, do triple plays made only in the 20th century count? Should triple plays in the 19th century count? What about triple plays made in the 21st century?

Second, should franchise shifts count as one team, or should each individual city be a separate team? The answer to this question significantly changes the answer to our original question. There is no simple agreement on this. Most people consider the California Angels to be the same team as the Anaheim Angels or the Los Angeles Angels. However, what about the San Francisco Giants and the New York Giants? Are the three Braves locations (Boston, Milwaukee, and Atlanta) one team or three? What if a team merely changes its name without changing any location such as the Houston Colt .45s to the Houston Astros? Are they the same team?

Finally, if a team switched leagues, do their records in the other league count? Most baseball fans consider the Milwaukee Brewers in the National League to be the same team as the Milwaukee Brewers in the American League. However, what about the American Association teams that jumped to the National League? Should their triple plays count? Also, after the merger of the American Association and the National League following the 1891 season, should their previous statistics be combined?

STEPHEN D. BOREN *MD, Ph.D. is a long-time contributor to SABR publications and is a practicing physician.* **JAMES A. SMITH,** *a retired computer operator, joined SABR in 1983. Besides baseball, he has many dedicated research interests, including submarines, the Civil War, and calendars from various cultures and eras.* **HERM KRABBENHOFT,** *who has not yet seen a ML triple play in person, is presently researching the uniform numbers worn by the players on his favorite team—the Detroit Tigers.*

Of the present American League teams, the Detroit Tigers have made the most triple plays (33). However, the Baltimore Orioles franchise can claim a grand total of 36. They made 12 as the Baltimore Orioles (1954-present), 23 when they were known as the St. Louis Browns (1902-1953), and 1 when they were the Milwaukee Brewers (1901). Note that this does not include the three TPs made in 1901 and 1902 by the Baltimore Orioles who in 1903 became the New York Highlanders/Yankees. The Minnesota Twins franchise is third with 30 triple slaughters. As the original Washington Senators (1901-1960) they made 20, and ten more since moving to Minnesota (1961-present).

Of the present National League teams, the Chicago Cubs have made the most with 40 triple plays. Eleven of these were in the 19th century and 29 from 1901 forward. The Giants with 33 triple plays in New York as the New York Giants and six as the San Francisco Giants have 39.[3] Thirteen of the New York total were in the 19th century and 20 in the 20th century.

The Braves franchise has made four as the Atlanta Braves, 30 as the Boston Braves, and three as the Milwaukee Braves, for a total of 37. Sixteen of the Boston triple plays were performed in the 19th century. Of note, the Braves and the Cubs are the only two continuous franchises from the original eight-team National League of 1876. The 1876 Cincinnati, New York, St. Louis, and Philadelphia teams are unrelated to the current major league teams.

The National League's Pittsburgh Pirates and the St. Louis Cardinals have made 35 and 36 triple plays, respectively. The Cincinnati Reds have made 29 triple plays in the National League. However, they jumped from the American Association after the 1889 season. Cincinnati made five triple plays while in the American Association (1882-1889). Thus, the continuous Reds franchise has made 34.[4]

A similar argument can be made for the Pittsburgh Pirates having made two more triple plays. As a member of the American Association (1882-1886) they made a pair. After the 1886 season, they too jumped to the National League. Thus, the team can claim 37 triple plays.

The merger of the American Association and the National League after the 1891 season resulted in a twelve-team league. The St. Louis Brown Stockings, Baltimore Orioles, Washington, and Louisville teams joined the National league, while the other four American Association teams folded. These Baltimore, Washington, and Louisville teams are unrelated to present-day major league baseball teams. However, should the present-day St. Louis Cardinals get credit for six triple plays they made in the American Association? This would raise their total to 42.[5]

There is also the issue of what to do about the St. Louis Terriers and the Chicago Whales of the Federal League. After the 1915 season, there was a "merger" of the Federal League and the American and National Leagues. The Terriers were merged with the St. Louis Browns, while the Chicago Cubs and Whales combined. Of note, present-day Wrigley Field was the Whales' ballpark. Should the Chicago Cubs receive credit for the one Whales triple play and the present-day Baltimore Orioles receive the two Terriers' triple plays? This does not seem appropriate.

Giving credit for all franchise shifts, but staying in the same league, the top teams for triple plays (excluding Federal League figures) are presented in Table 1.

Table 1. Teams who executed the most triple plays, including franchise shifts within the same league

TEAM* (CITIES)	LEAGUE	YEARS	TPs
Cubs (CHI)	NL	1876-2002	40
Giants (NY/SF)	NL	1883-2002	39**
Braves (BOS/MIL/ATL)	NL	1876-2002	37
Orioles (MIL/STL/BAL)	AL	1901-2002	36
Cardinals (STL)	NL	1892-2002	36
Pirates (PIT)	NL	1887-2002	35
Tigers (DET)	AL	1901-2002	33
Twins (WASH/MN)	AL	1901-2002	30
Phillies (PHI)	NL	1883-2002	30

*The team name listed is the current team name.
**See end note 3.

The top teams—including league transfers—for triple plays are listed in Table 2.

Table 2. Teams who executed the most triple plays, including franchise and league shifts

TEAM* (CITIES)	LEAGUE	YEARS	TPs
Cardinals (STL)	AA-NL	1882-2002	42†
Cubs (CHI)	NL	1876-2002	40
Giants (NY/SF)	NL	1883-2002	39**
Braves (BOS/MIL/ATL)	NL	1876-2002	37
Pirates (PIT)	AA-NL	1882-2002	37
Orioles (MIL/STL/BAL)	AL	1901-2002	36
Reds (CIN)	AA-NL	1882-2002	34
Tigers (DET)	AL	1901-2002	33
Twins (WASH/MIN)	AL	1901-2002	30
Phillies (PHI)	NL	1883-2002	30

*The team name listed is the current team name.
**See end note 3.
†See end note 5.

Excluding 19th century triple plays, the top teams for triple plays from 1901 through the present are given in Table 3.

Table 3. Teams who executed the most triple plays since 1901, including all franchise and league shifts

TEAM* (CITIES)	LEAGUE	YEARS	TPs
Orioles (MIL/STL/BAL)	AL	1901-2002	36
Tigers (DET)	AL	1901-2002	33
Cardinals (STL)	NL	1901-2002	33
Pirates (PIT)	NL	1901-2002	31
Twins (WASH/MIN)	AL	1901-2002	30
Red Sox (BOS)	AL	1901-2002	29
Cubs (CHI)	NL	1901-2002	29
Indians (CLE)	AL	1901-2002	28**

*The team name listed is the current team name.
**The Indians also pulled a triple play in the 1920 World Series against the NL Dodgers; thus, their grand total of triple plays (including post-season play) is 29.

Table 4 presents the top teams for triple plays from 1901 thorough 2003 without franchise shifts.

Table 4. Teams who executed the most triple plays since 1901, excluding all franchise and league shifts

TEAM* (CITIES)	LEAGUE	YEARS	TPs
Tigers (DET)	AL	1901-2002	33
Cardinals (STL)	NL	1901-2002	33
Pirates (PIT)	NL	1901-2002	31
Red Sox (BOS)	AL	1901-2002	29
Cubs (CHI)	NL	1901-2002	29
Indians (CLE)	AL	1901-2002	28**

*The team name listed is the current team name.
**The Indians also pulled a triple play in the 1920 World Series against the NL Dodgers; thus, their grand total of triple plays (including post-season play) is 29.

Finally, to be complete, Table 5 lists the top triple play makers from the other four major leagues.

Table 5. The teams who executed the most triple plays in the other major leagues

LEAGUE	YEARS	TEAM (CITIES)	TPs
American Association	1882-91	Athletics (PHI)	6
Players League	1890	Pirates (CHI)	2
Federal League	1914-15	Terriers (STL)	2
Union Association	1884	--	0

So, who has made the most triple plays?

It could be the Cubs with 40 (Table 1); or the Cardinals with 42 (Table 2); or the Orioles with 36 (Table 3); or the Tigers with 33 (Table 4). It depends on your point of view!

NOTES

1. "Tigers Best at Triple Plays," R.J. Gonzalez, *The Baseball Research Journal* (p. 76, 1972). The article states: "Since 1901 when the American and National Leagues began operating simultaneously, there have been a total of 374 triple plays in the majors, 188 in the AL and 186 in the NL. Detroit leads all teams with 30, followed by the Cards and Pirates with 26." It is also noted that the Detroit Tigers issued a press release (February 5, 1974) in which it was stated, "The Tigers can claim the triple play championship of the major leagues. They have made more triple-killings than any other club in modern baseball history, according to Raymond J. Gonzalez, a baseball statistician from Woodside, NY, who has tracked down every triple play made since 1900 for the Society for American Baseball Research."
2. "Tigers Hold All-Time Triple Play Lead," The *2003 Detroit Tigers Information Guide* (p. 7). The information box states: "The Tigers' all-time total of 33 (triple plays) is more than any other team in major league history. The Pittsburgh Pirates and the St. Louis Cardinals are tied for the National League lead with 31." The Detroit Tigers have had such a statement published in their annual media/information guides every year since 1978 (although for the guides from 1978 through 1982 the qualifying statement "since 1900" was included).
3. It can also be argued that the Troy Trojans, who played in the National League from 1879 through 1882, were the direct precursors to the New York Giants, since four of the 1882 Troy players were regulars in the 1883 New York lineup. And since the Troy team made two triple plays, the grand total number of triple plays pulled by the Troy-New York-San Francisco conglomerate is 41.
4. The original 1876 Cincinnati team was a distinct team from the present Cincinnati one and did not make any triple plays. Similarly, there was a short-lived Cincinnati franchise in the American Association that replaced the original AA Cincinnati team; they also did not execute any triple plays.
5. The American Association St. Louis club also pulled a triple play in the 1887 World Series against the National League Detroit Wolverines. Including this post-season TP gives St. Louis franchise a grand total of 43 three-ply killings.

PLEASED TO MEAT YOU — **Submitted by David MacDonald**

"When the Giants purchased Rube Marquard from Indianapolis for $11,000, his side partner, Bull Durham, also a pitcher, was included in the deal. Last winter in Los Angeles Marquard was hailed by a fat person who wore a butcher's apron and stood in front of a meat shop. 'Hello, Rube,' cried the corpulent one, extending his hand. 'Who are you?' queried the Giants' southpaw. 'Bull Durham,' was the reply, 'I couldn't hold on in the big leagues so I went back to chopping beef.' Bull is there yet." —*Ottawa Citizen*, May 17, 1913

ROBERT H. SCHAEFER

Cap's Bats

The Baseball Bats of Captain Adrian C. Anson

Adrian C. Anson was the venerable captain of the famous Chicago White Stockings when they were the kingpins of the baseball world in the 1880s. Anson is more popularly, although slightly incorrectly, known to today's fans as "Cap." In his own time Anson was identified in the press, variously, as "Ans," "The Big Swede," and "Baby." It wasn't until 1879, when he was appointed captain and manager of the Chicago nine, that his name appeared in print prefaced by "Captain," and sometimes "Capt." but almost never as "Cap." True, players of that day normally addressed the team captain as "Cap," but that was a familiarity not usually taken by the reporters. In the last stages of his 26-year career Anson was known as "Uncle Anson," "Your Uncle," "Pop," "The Old Man," and finally after he retired in 1897, "The Grand Old Man." Incidentally, his wife called him "Pop."

One of Uncle Anson's many claims to fame is as a hitter. He was the first major league hitter to collect more than 3,000 base hits, including those he accumulated in the National Association. This was despite the fact that during his career the schedule never approached 154 games per season. Between the years 1871 to 1897 Anson participated in a low of 29 games (1871) and a high of 134 games (1889) for a grand total of 2,523. Anson played 247 games in the Association and 2,276 in the League.

Players of the 19th century indulged themselves in an abundance of superstitions. Many players named their bats, and some, notably Pete Browning, believed that each bat contained only a given number of hits. Once all the hits had been knocked out of the bat, you might as well throw it away. A hitter knew exactly when he had exhausted the quota of hits in any given bat—because the base hits just stopped coming. It was a sure sign you needed to change your bat. And that is just what many players did. But instead of putting the now hitless bat in the trash, hitters who held such mystical beliefs, such as Browning and Anson, retired it with honor to their basements and preserved it. Anson treated his bats with awe and reverence that bordered on superstitious behavior. A bat was not just an inanimate hunk of wood. No, it was a living entity that had to be treated with respect if it was to do its intended work—lining out base hits on a regular basis.

The size and weight of bats in Anson's time were unique to that period. Following his career as a brilliant hitter George Sisler became a very effective hitting instructor. In addition to performing this service for several major league clubs, Sisler wrote detailed instructional booklets on the art and science of hitting. In 1935 he made the following comparison of bats of that day to bats of the past:

"Have you ever seen specimens of the bats used in the early days of baseball? Veritable wagon tongues they were, not much longer than modern bats but much heavier and much thicker in the handles. The bat used by Pop Anson, whose record, as a leading batsman, almost equals that of Ty Cobb, was 36 inches long and weighed 48 ounces and its handle had a circumference of 4 inches."

John Phillips, in *The Riotous 1896 Cleveland Spiders*, provides this description of Anson's bat:

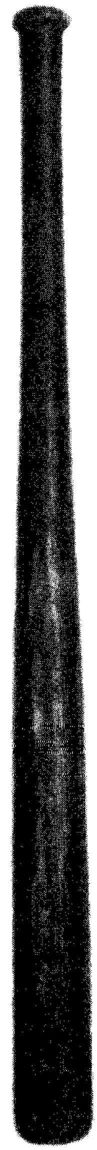
"Tacks Parrott of the Browns uses the longest bat in the League. Uncle Anson, of course, uses the heaviest. Ans says it's easier to place hits with heavy bats than with light ones. He knows what he's talking about. His bat is made of hickory and could be broken only by a rock crusher. Nobody would ever steal it because nobody else could swing it."

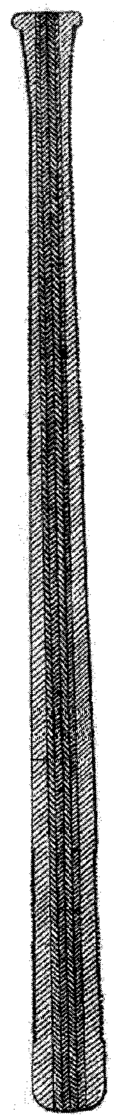
On July 9, 1884, *The Sporting Life* published the following report on what was reputedly a revolutionary technique in bat design and construction:

BOB SCHAEFER *is retired from the aerospace industry. He is a long-time contributor to SABR publications.*

ANSON'S NEW BAT—SOMETHING NEW
A Bat Which is Calculated to Fatten Batting Averages

> Captain Anson in the game of the 23rd made a trial of a new style of bat just made as an experiment. The bat is made of several pieces of ash, jointed and glued together lengthwise, while in the center is inserted a rattan rod about one inch square, and composed of twelve strips of rattan firmly glued together, running from end to end of the bat. The handle is wound with linen cord. This wrapping of the handle, however, is technically a violation of the rule, which requires the bat to be made "wholly of wood," but it is a rule which nobody will object to changing if the wound handle proves to be an improvement. The object of the glue joints and the rattan rod in the center is to make the bat less liable to break and at the same time to give it more spring. That both of these objects are accomplished there can be no doubt. The first ball hit by Anson with the new bat was a terrific liner to left field for two bases, and he used it throughout the game with great success, Captain Morrill having agreed to waive any objection to the wrapping of the handle. Heretofore bats have been made of a single stick, and the improvement adds materially to the expense of manufacture. Players who have tried it say that the ball can be driven 25% farther by the exercise of equal force than the common bat. Anson certainly made a remarkable record in the two games in which he used it. On June 23, Buffington pitcher, in three times at bat he made a single and a double; on June 24, Whitney pitcher, four times at bat, two singles, one double and a home run. The cost of the new bat will be about $5 each.

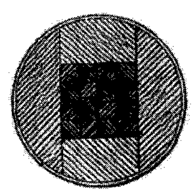




The premium wooden bats of 1884 cost from $1.00 to $1.25 each, so the proposed cost of Anson's new bat, five dollars, was extremely high. The matter of wrapping the handle with linen cord was but a trifle to deal with. In 1885 the rules were amended to allow the bat to be wound with twine for a distance of 18 inches from its end, so that objection disappeared.

An interesting aspect of Anson's "new" bat of 1884 is that it wasn't really a "new" concept. Henry Chadwick, of the *New York Clipper*, reported on May 3, 1874, that he was shown a new type of bat at George Wright's sporting goods store in Boston. This bat was made with a cane fitted through the whole length of the bat. The purpose of this cane was two-fold: to prevent the bat from breaking, and to impart elasticity, which drives the ball farther. It was claimed that the bat would last the entire season. The bat only weighed thirty-two ounces, extremely light for the time. It was also very expensive, costing $4.00. Wright's bat of 1874 sounds almost identical to Anson's bat of 1884.

One of Anson's many bats, ca. 1890



Cutaway View of Anson's Laminated Bat

End View of Anson's Laminated Bat

Specific follow-up reports of George Wright's bat being used in games, or subsequent use of Anson's similar bat of 1884, have not been uncovered at the present time. The rule governing bat design in 1884 simply stated that the bat was to be made wholly of wood so that both Wright's bat and Anson's bat were in compliance. It wasn't until 1940 that the rule governing bat construction was amended to specify, "The bat must be made entirely of hardwood in one piece." Clearly Anson's bat met all the applicable rules of 1884.

A discussion on the use of multi-wood bat construction appeared in the *Spalding Official Base Ball Guide for 1925,* long after Anson's time in baseball. The 1925 guide included an article about Jack Pickett, who had passed away in Chicago during the summer of 1922. Pickett had been the chief bat designer for the Spalding factory, and it was claimed that he had designed more baseball bats than any other man in the world.

"The heaviest bat that Pickett ever put together was one that was used by Anson, and those who recall Anson in his prime will also remember that the bat which he took to the plate with him was left severely alone by the other players. They couldn't swing it. Anson had a core of hickory in the bat and over it was split bamboo, or occasionally ash. The bat was so heavy that players who were fast swingers with light sticks could not get Anson's hardwood war club around in time to meet the ball."

From this account it is reasonable to infer that Anson used bats of laminated construction throughout his career, although not exclusively, as he used many bats of traditional design and fashioned from a single piece of hardwood.

Pop was continually experimenting in an attempt to find the perfect piece of lumber with which to sting the ball. Legend has it that he would drive his horse-drawn rig throughout the countryside, keeping a sharp eye peeled for candidate lumber that was properly seasoned. Sometimes an old fence post caught his fancy; other times he became enamoured of a fallen log or a stump. Anson would then strike a bargain with the farmer on whose land he discovered the raw material for his new bat, load the lumber into his rig, haul it home, and have it converted to a war club. Over the course of his 26-year career Anson constantly changed bats, but he never disposed of a single one. Anson "retired" his bats to his basement.

On March 24, 1906, *The Sporting Life* carried the following report:

ANSON CAN FURNISH A TREAT TO UP-TO-DATE PLAYERS

A Bat Museum at His Chicago Home Well Worth Looking Over

PITTSBURG, MARCH 19.—Sitting in a corner of base ball headquarters, well wrapped up in paper, is a ball bat and the owner guards it carefully. Harry Pelitz owns the stick. Last year the veteran chanced to meet Uncle Anson in Chicago. The chat drifted to bats. The old man brought a gleam of joy to Peitz's face by saying: "Harry, come out to my house this evening after the game and I will show you something to make your eyes water."

Peitz was on hand. The old man took him down to the cellar. There piled in racks were hundreds of bats, the genuine hickory kind, which would be stolen, even if allowed to lay on a crowded field. Anson allowed Peitz to take his pick of the bludgeons. The Premier grabbed the chance. Now there isn't a more cherished stick owned by a Pittsburg player. Peitz has allowed Clymer to have the bat temporarily so that he can make a club just after its mold. 'That's a great piece of wood," said Clymer enviously.

In a private letter now on file in the National Baseball Library in Cooperstown, NY, dated February 20, 1957, Jack Corbett[1] of the Hollywood Star Base, tells this tale of Anson's devotion to his lumber: "Anson had over 400 bats in his cellar when he quit playing. All were oiled and dusted every day except Sunday up until 1907. While we were playing poker at Cap's home one evening [in 1907] Mrs. Anson told me that there would be a present for me outside one of the cellar windows when I left for home. There was—one of the Cap's bats that he had sworn he would never part with.

"I took it to the old Spalding bat factory and Jack Pickett put it in a lathe and turned off just enough to make it look fresh. When I stepped out on the club house porch with it Cap was sitting in the stands at least fifty yards off. He looked at me and then the stick—he got up and roared at me—"Where did you get that bat ?" I told him I had swiped it out of the cellar and after kicking that around a little he said, "Well, now that you have it, use it and be sure you get some hits with it."

Corbett continued with his tale of Cap's bats:

"When we later played Callahan I let Mike Donlin use it. He showed it to Callahan and Jimmy got hold of the Cap and gave him $7.00 per bat for all he had."

This evidently was the fate of the bats Anson had stored in his cellar. The Old Man must have been hard pressed for money in order to sell his most cherished possessions in that manner.

As Anson's career wore on, speculation abounded each winter over when The Old Man would retire from the diamond. By 1893 Anson had achieved the ripe old age of 42, and more than half of the summers of his life had been spent in professional baseball. On February 5 *The Sporting Life* carried the following report:

NO THOUGHT OF RETIRING
Anson Laying Plans for the Next Two Years

One of the sensations of the season was sprung the other day from Chicago. Just as the newspapers began once more to discuss the probable withdrawal of Uncle Adrianapolis Constantinople Anson from active life on the base ball field a report went out of Chicago that caused the presses to stop and every reporter to say, "Where am I ?" The report was innocent in itself, but far reaching in its scope—Anson had bought two dozen bats. That in itself signifies nothing, but when the further information was tacked on that the bats were of green wood and that the old man had carried them to his garret with this comment, "In two years them sticks will be dried out, and then I'll fetch 'em and kill the ball"—that was what paralyzed the pulse of the prophecy. If a man is laying his plans for two years hence, why should he not be looking into the next century?

Although Anson was forced into retirement at the end of the 1897 season, his legendary bats continued to do their work on the diamond well into the 20th century. On May 25, 1911, *The Sporting News* reported the current use of one of Anson's bats:

"John Titus, the Philadelphia right fielder, is using a bat that has some history connected with it. In 1894 Captain Anson discovered a piece of timber that is considered ideal wood for a base ball bat and he proceeded to have it turned into a cudgel. Anson in his day merely had to swing it and the ball would go to the fence. It is so heavy, however, that many an ordinary player would hardly care to handle it. When Pop Anson retired from the game he retained this great stick as a treasure. At last when the former star's belongings went under the hammer Pat Moran purchased this bat, and when Pat was bought from the Cubs he brought it to Philadelphia. Titus coaxed and finally Moran consented to let him have the bat. His first hit was a home run over the fence off Bob Harmon, of St. Louis. Titus has been batting consistently ever since he came into possession of Pop's old lumber."

John Titus played in 76 games in 1911 and posted a .284 batting average, not much over his .282 lifetime mark. However, he achieved career highs for both slugging average and home runs. How much credit for this performance was due to Pop's old bat?

NOTES

1. SABR member Daniel Ginsburg relates that Jack Corbett was a lifetime baseball man. He played in the minor leagues, as well as outlaw leagues, for 14 years, mostly during the Deadball Era. Later Corbett owned teams in Atlanta, Jersey City, Syracuse, and El Paso.

A UNIQUE ACCIDENT

Submitted by Bob Schaefer

A Foul Ball Drove Open Knife into a Scorer's Heart

Bellefontaine, Ohio., Oct. 28. - The most peculiar fatality that has ever occurred on a base ball field was that at Morrisontown. Stanton Walker was keeping score of the game and asked a companion for a knife to sharpen his pencil. As the knife was being handed to Walker the batsman sent a line foul toward the crowd of which Walker was one. The ball struck Walker's hand as he held the knife with the blade pointing toward his heart. Walker fell without a groan in the ground, and when picked up was dead, the knife sticking in his heart and his hand still clasping the handle." — As reported in the November 8, 1902 issue of *Sporting Life*.

HERM KRABBENHOFT

Normalized Winning Percentage (NWP)

Eddie Lopat vs. the Indians, Frank Lary vs. the Yankees

It was recently reported that Eddie Lopat, who pitched for the Chicago White Sox, the New York Yankees, and the Baltimore Orioles, compiled a phenomenal 40-13 W-L record versus the Cleveland Indians during his major league career (1944-1955).[1] For the years that he was a full-time, full-season player (from 1944 through 1954), his W-L record against the Tribe was 40-12 (which yields a .769 winning percentage). How does this exceptional individual performance compare with that achieved by Lopat's mound mates as a group?

To answer that question, one can make use of "Normalized Winning Percentage" (NWP). NWP, devised by Bill Deane in 1983, projects how a pitcher might perform on a .500 team, thus putting all hurlers, past and present, on an even plane of comparison.[2] NWP is defined in Equation 1, where WAT is the pitcher's Wins Above Team, and PD is the Pitcher's Decisions (i.e., wins plus losses). WAT is the number of wins a pitcher garnered beyond those expected of an average pitcher for that team.[3] WAT is defined in Equation 2, where P% is the pitcher's winning percentage and T% is the team's adjusted winning percentage (i.e., the winning percentage obtained after subtracting the pitcher's wins and losses from the team's overall wins and losses).

$$\text{NWP} = .500 + (\text{WAT}/\text{PD}) \qquad (1)$$

$$\text{WAT} = \text{PD} \times (\text{P\%} - \text{T\%}) \div [2 \times (1.000 - \text{T\%})] \qquad (2)$$

My reason for wanting to determine Lopat's NWP versus the Cleveland Indians arose from my interest in Frank Lary's spectacular won-lost record versus the New York Yankees.[4] Lary (known as "The Yankee Killer") twirled for the Detroit Tigers and fashioned an amazing 28-11 W-L ledger (which affords a winning percentage of .718) against the perennial pennant winners during an eight-year stretch (1955-1962).[5]

Based on standard winning percentages, Lopat (.769) seems to have been somewhat better versus the Indians than Lary (.718) was against the Yankees. Accordingly, I was curious how their corresponding Normalized Winning Percentages would compare.

Table 1 collects the pertinent WAT and NWP results (obtained via a Microsoft Works spreadsheet treatment) for Eddie Lopat versus the Cleveland Indians in the 1944-1954 period. Table 2 collects the pertinent WAT and NWP results (obtained via a Microsoft Works spreadsheet treatment) for Frank Lary versus the New York Yankees for the 1955-1962 period.

Inspection of Table 1 reveals that without Eddie Lopat on the mound, his teams (either the White Sox or the Yankees) were better than .500 against the Indians, .514 and .513, respectively. During Eddie's tenure with the Pale Hose (1944-1947), which includes two World War II seasons, his NWP against the Tribe was a glowing .863. And for his full-time stint with the Pinstripers, his NWP was an impressive .723. Overall, he produced a NWP of .763.

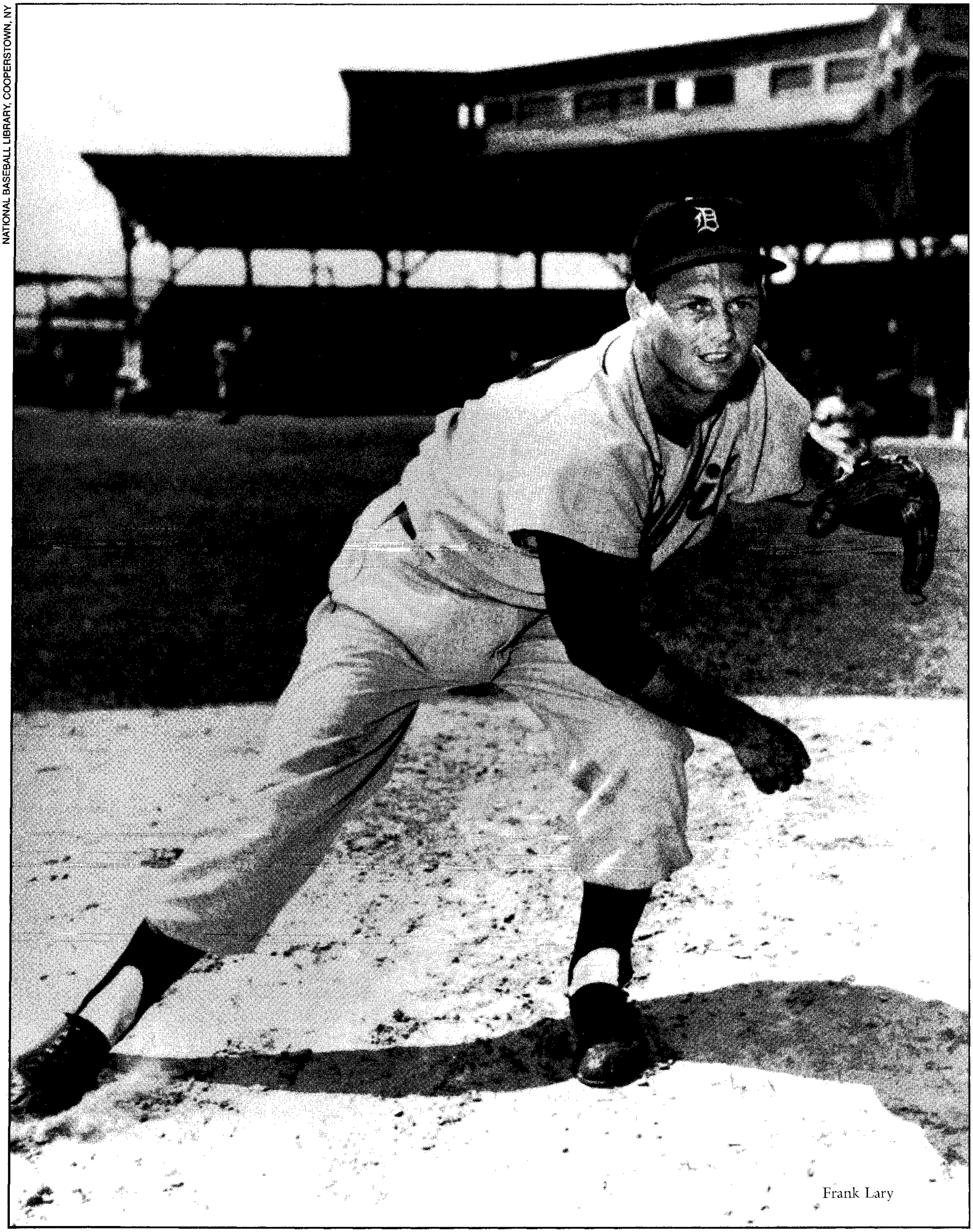
Examination of Table 2 suggests that Frank Lary's NWP performance against the Yankees (who captured the AL pennant in seven of the eight years from 1955 through 1962) was virtually the same as Lopat's versus the Indians. "The Yankee Killer" compiled a fantastic .761 NWP in his confrontations with the Bronx Bombers. The Tigers, on the other hand, without Lary on the hill had a dismal .411 winning percentage against the Yankees.

So, in summary, it is seen that Lopat and Lary each outperformed their collective mound mates—in an enormous way—in their respective battles against the Indians and Yankees, achieving nearly identical NWPs of .763 and .761, respectively.

As a kid, **HERM KRABBENHOFT** *attended and scored four games that Frank Lary pitched against the Yankees (at Briggs Stadium).*

ACKNOWLEDGMENTS

Thanks to Bill Deane for his interest and helpful comments on my NWP research.



Frank Lary

Table 1. WAT and NWP results for Eddie Lopat's pitching vs. the Cleveland Indians (1944-1954)

YEAR	TEAM W-L (W%)	LOPAT W-L W%)	ADJ. TEAM W-L (W%)	LOPAT WAT	LOPAT NWP
1944	14-8 (.636)	3-0 (1.000)	11-8 (.579)	1.50	1.000
1945	11-8 (.579)	3-0 (1.000)	8-8 (.500)	1.50	1.000
1946	13-9 (.591)	3-1 (.750)	10-8 (.556)	0.88	.719
1947	11-11 (.500)	4-1 (.800)	7-10 (.412)	1.65	.830
Σ CHI	49-36 (.576)	13-2 (.867)	36-34 (.514)	5.44*	.863
1948	12-10 (.545)	5-2 (.714)	7-8 (.467)	1.63	.732
1949	12-10 (.545)	4-2 (.667)	8-8 (.500)	1.00	.667
1950	14-8 (.636)	6-0 (1.000)	8-8 (.500)	3.00	1.000
1951	15-7 (.682)	5-2 (.714)	10-5 (.667)	0.50	.571
1952	12-10 (.545)	2-1 (.667)	10-9 (.526)	0.44	.648
1953	11-11 (.500)	2-2 (.500)	9-9 (.500)	0.00	.500
1954	11-11 (.500)	3-1 (.750)	8-10 (.444)	1.10	.775
Σ NY	87-67 (.565)	27-10 (.730)	60-57 (.513)	8.24**	.723
TOTAL	136-103 (.569)	40-12 (.769)	96-91 (.513)	13.67***	.763

* Adding up the WAT for the individual seasons yields a total WAT of 5.53, a value in close agreement with the 5.44 value calculated from the Σ CHI row entries.

** Adding up the WAT for the individual seasons yields a total WAT of 7.67, a value in line with the 8.24 value calculated from the Σ NY row entries.

*** Adding up the WAT for the individual seasons yields a total WAT of 13.20, a value in line with the 13.67 value calculated from the Total row entries.

Table 2. WAT and NWP results for Frank Lary's pitching vs. the New York Yankees (1955-1962)

YEAR	TEAM W-L (W%)	LARY W-L W%)	ADJ. TEAM W-L (W%)	LARY WAT	LARY NWP
1955	10-12 (.455)	2-1 (.667)	8-11 (.421)	0.64	.712
1956	12-10 (.545)	5-1 (.833)	7-9 (.438)	2.11	.852
1957	10-12 (.455)	2-2 (.500)	8-10 (.444)	0.20	.550
1958	12-10 (.545)	7-1 (.875)	5-9 (.357)	3.22	.903
1959	14-8 (.636)	5-1 (.833)	9-7 (.563)	1.86	.810
1960	8-14 (.364)	2-2 (.500)	6-12 (.333)	0.50	.625
1961	8-10 (.444)	4-2 (.667)	4-8 (.333)	1.50	.750
1962	7-11 (.389)	1-1 (.500)	6-10 (.375)	0.20	.600
TOTAL	81-87 (.482)	28-11 (.718)	53-76 (.411)	10.16*	.761

* Adding up the WAT for the individual seasons yields a total WAT of 10.23, a value in close agreement with the 10.16 value calculated from the Total row entries.

REFERENCES AND NOTES

1. Spatz, Lyle. "Ed Lopat and the Cleveland Indians," SABR Committee newsletter, p. 1 (February 2003); the material presented was based on contributions from Steve Gietschier, Pete Palmer, and Zita Carno.
2. Deane, Bill. "Normalized Winning Percentage," *The Baseball Research Journal,* No. 25, pp. 42-44 (1996), and "Normalized Winning Percentage Revisited," *By the Numbers,* vol. 9 (No. 2), pp. 6-7 (May 1999).
3. Thorn, John, Pete Palmer, and M. Gershman, eds., *Total Baseball,* 7th ed., Total Sports Publishing, Kingston, NY, p. 2,501.
4. Krabbenhoft, Herm. "The Phenomenal Achievement of Frank Lary—Premier Yankee Killer," *Baseball Quarterly Reviews,* vol. 1 (No. 3), pp. 65-81 (Fall 1986).

5. In an injury-plagued 1963 season, Lary was 0-2 against the Pinstripers, which gave him a career 28-13 W-L mark versus the Bronx Bombers. In the 1964 season, Lary started with the Tigers (and did not face the Yankees), but was sold to the New York Mets and then traded to the Milwaukee Braves. In 1965 (Lary's last season in the majors), his contract was again purchased by the Mets, who later traded him to the Chicago White Sox (with whom he faced the Yankees three times, each a relief appearance, and had no decisions.

FRANK M. CHIMKIN

Another Look at Runs Created

One of the many things that makes baseball great is the ability to both objectively and subjectively compare which players are the best. These comparisons range anywhere from scholarly research[1] to radio talk show discussions to barroom arguments.

In comparing players, many times researchers have developed new statistics in an attempt to find one all-encompassing number and more objectively assess the value of one player versus other players of their eras. This number can then be adjusted to league averages and for park effects to compare players of all eras. Nowadays, the number most often used in this vein is OPS (On-Base Percentage plus Slugging Average, also known as Production). This measure is popular mainly because it is a "simple but elegant measure of batting prowess, in that the weaknesses of one-half of the formulation, On-Base Percentage, are countered by the strengths of the other, Slugging Average, and vice versa."[2]

Another such statistic, Runs Created (RC) was developed by Bill James, based upon the fact that the "best hitter is the hitter who creates the most runs."[3] Over the years James has introduced several more complicated versions of the RC formula, each adding more statistics not available in all eras (e.g., hit-by-pitch), to more closely associate the value to runs.[4]

However, one of the disadvantages of developing a single number is that you lose the component numbers and traditional statistics, which are, arguably, more fun to compare. At the same time, comparing players of different eras is quite difficult using many of these component statistics, simply because many are next to impossible to adjust due to such large differences in league averages over the years.[5]

In his *New Historical Baseball Abstract,* Bill James created two algorithms for adjusting these component statistics using his RC formula. In his Willie Davis comment (pp. 740-43), James used the first algorithm for adjusting Davis' statistics as if each of the teams he was on had scored 750 runs per year.[6] In his Sam Crawford comment (pp. 795-96), James expands on the first algorithm by including a second algorithm to convert Crawford's Deadball Era statistics as if he had started his career in 1920 instead of 1900.[7] Rather than using a constant 750 runs per year, James used whatever amount the team Crawford was on in a particular year had scored 20 years later.

Using the Sean Lahman Baseball Archive Database (v. 5.0) available online at www.baseball1.com, and Microsoft Access Basic/SQL, I created a hybrid of James's two algorithms to adjust all players statistics (1876–2002) as if their teams had scored 750 runs per year, as well as adjusting for the Deadball Era conversion, and park factors. By adjusting for each of these factors, we can then better compare players' traditional statistics across eras and teams.

FRANK CHIMKIN *is Data Manager/Analyst for the Division of General Pediatrics, Columbia University. He has been a SABR member since 1993. He dedicates this article to his better half, Michele; his father, Stuart; and his late maternal grandfather, Irving Weisman.*

METHODOLOGY

As in James' algorithms, all counting batting statistics rise and fall with hits. Therefore, the RC formula is adjusted as the elements relate to hits. Thus, the formula becomes:

$$\text{RC} = \frac{((\text{H} + (\text{H} \times \text{BB/H})) \times (\text{H} \times \text{TB/H}))}{(\text{Outs} + \text{H}) + (\text{H} \times \text{BB/H})}$$

From there, you solve the equation for H. Without going into the algebra to make the quadratic equation that results, the formula to solve for H is:

$$\text{H} = \frac{-1\times(-1-\text{BB/H})) + \text{SQRRT}((-1-(\text{BB/H}))^2-4\times(\text{TB/H}+(\text{BB/H}\times\text{TB/H}))}{\text{RC}\times(-1\times\text{Outs}))) \div (2\times(\text{TB/H}+(\text{BB/H}\times(\text{TB/H}))/\text{RC})}$$

For each season prior to 1920, the program first converted the appropriate statistics using James' 1920 algorithm. Then, for each season the program con-

Table 1. Adjusted career leaders through 2002

a. Career home runs

POS	PLAYER	ADJ-HR	REAL-HR	DIFF	RP†
1	Hank Aaron	811	755	56	1
2	Babe Ruth	733	714	19	2
3	Willie Mays	699	660	39	3
4	Barry Bonds*	635	613	22	4
5	Frank Robinson	627	586	41	5
6	Harmon Killebrew	614	573	41	7
7	Reggie Jackson	600	563	37	8
8	Mike Schmidt	583	548	35	9
9	Mark McGwire	578	583	-5	6
10	Mickey Mantle	571	536	35	10
20	Sam Crawford	488	97	391	***
29	Honus Wagner	462	101	361	***
44	Ty Cobb	399	117	282	***
T63	Jake Beckley	367	86	281	***
T69	Roger Connor	356	138	218	T401

b. Career batting average (at least 4,000 AB's)

POS	PLAYER	ADJ-BA	REAL-BA	DIFF	RP†
1	Ty Cobb	.3659	.3664	-0.0005	1
2	Rogers Hornsby	.3601	.3585	0.0016	2
3	Joe Jackson	.3503	.3558	-0.0055	3
4	Tony Gwynn	.3495	.3382	0.0113	17
5	Ted Williams	.3442	.3444	-0.0002	6
6	Rod Carew	.3413	.3278	0.0135	28
7	Tris Speaker	.3409	.3447	-0.0038	5
8	Stan Musial	.3378	.3308	0.0070	24
9	Bill Terry	.3377	.3412	-0.0035	12
10	Harry Heilmann	.3346	.3416	-0.0070	10
11	Babe Ruth	.3344	.3421	-0.0077	9
173	Ed Delahanty	.2988	.3459	-0.0471	4
268	Dan Brouthers	.2912	.3424	-0.0512	8
322	Billy Hamilton	.2882	.3443	-0.0561	7

c. Career hits

POS	PLAYER	ADJ-H	REAL-H	DIFF	RP†
1	Pete Rose	4610	4256	354	1
2	Ty Cobb	4181	4189	-8	2
3	Hank Aaron	4044	3771	273	3
4	Stan Musial	3745	3630	115	4
5	Carl Yastrzemski	3614	3419	195	6
6	Willie Mays	3494	3283	211	11
7	Tris Speaker	3456	3514	-58	5
8	Eddie Murray	3346	3255	91	12
9	Paul Molitor	3330	3319	11	9
10	Tony Gwynn	3302	3141	161	18
11	Honus Wagner	3268	3415	-147	8
13	Eddie Collins	3264	3315	-51	10
59	Cap Anson	2637	3418	-781	7

d. Career RBI

POS	PLAYER	ADJ-RBI	REAL-RBI	DIFF	RP†
1	Hank Aaron	2566	2297	269	1
2	Ty Cobb	2141	1937	204	6
3	Babe Ruth	2138	2213	-75	2
4	Willie Mays	2099	1903	196	9
5	Stan Musial	2043	1951	92	5
6	Frank Robinson	2028	1812	216	15
7	Carl Yastrzemski	2013	1844	169	11
8	Eddie Murray	2010	1917	93	8
9	Dave Winfield	1969	1833	136	13
10	Honus Wagner	1911	1732	179	16
11	Mel Ott	1896	1860	36	10
14	Lou Gehrig	1834	1995	-161	4
26	Jimmie Foxx	1717	1922	-205	7
36	Cap Anson	1631	2076	-445	3

e. Career OPS (at least 4,000 AB's) **

POS	PLAYER	ADJ-OPS	REAL-OPS	DIFF	RP†
1	Babe Ruth	1153	1164	-11	1
2	Ted Williams	1117	1116	1	2
3	Barry Bonds*	1053	1027	26	5
4	Lou Gehrig	1045	1080	-35	3
5	Rogers Hornsby	1036	1010	26	8
6	Mickey Mantle	1022	979	43	12
7	Joe Jackson	1009	940	69	T28
8	Ty Cobb	1001	945	56	24
9	Stan Musial	996	977	19	T13
10	Jimmie Foxx	993	1038	-45	4
11	Frank Thomas*	986	1006	-20	9
12	Hank Greenberg	985	1017	-32	6
13	Mark McGwire	983	986	-3	10
18	Manny Ramirez*	974	1014	-40	7

f. Career triples

POS	PLAYER	ADJ-3B	REAL-3B	DIFF	RP†
1	Paul Waner	190	191	-1	10
2	Stan Musial	183	177	6	T19
3	Roberto Clemente	177	166	11	T27
T4	Sam Rice	166	184	-18	14
T4	Goose Goslin	166	173	-7	22
6	Pie Traynor	160	164	-4	T30
7	Lou Gehrig	156	163	-7	T33
T8	Lou Brock	153	141	12	63
T8	Ty Cobb	153	295	-142	2
T10	Enos Slaughter	152	148	4	T54
T10	Kiki Cuyler	152	157	-5	T45
34	Tris Speaker	124	222	-98	6
T74	Sam Crawford	99	309	-210	1
T131	Honus Wagner	80	252	-172	3
T217	Fred Clarke	66	220	-154	7
T229	Jake Beckley	65	243	-178	4
T275	Roger Connor	60	233	-173	5
T338	Joe Kelley	54	194	-140	9
T348	Dan Brouthers	53	205	-152	8

* Active player
** OPS includes HBP but not SF
*** Player not in top 500
† Real position

Table 2. Adjusted seasonal leaders through 2002

a. Seasonal home runs

POS	PLAYER	YEAR	ADJ-HR	REAL-HR	RP†
1	Bonds, Barry*	2001	74	73	1
2	McGwire, Mark	1998	70	70	2
3	Sosa, Sammy*	1998	65	66	3
4	Sosa, Sammy*	2001	64	64	5
5	Maris, Roger	1961	63	61	7
6	McGwire, Mark	1999	62	65	4
7	Sosa, Sammy*	1999	59	63	6
8	Ruth, Babe	1927	58	60	8
9	McGwire, Mark	1997	57	58	10T
10	Mays, Willie	1965	56	52	22T
11T	Ruth, Babe	1921	55	59	9
19T	Foxx, Jimmie	1932	53	58	10T
29T	Greenberg, Hank	1938	52	58	10T
81T	Ruth, Babe	1919	46	29	***
95T	Freeman, Buck	1899	45	25	***
110T	Schulte, Frank	1911	44	21	***
165T	Cravath, Gavvy	1915	42	24	***
165T	Wilson, Chief	1912	42	11	***

b. Seasonal batting average

POS	PLAYER	YEAR	ADJ-BA	REAL-BA	RP†
1	Hornsby, Rogers	1924	.4288	.4235	5
2	Sisler, George	1922	.4128	.4198	6
3	Cobb, Ty	1922	.4011	.4011	25
4	Sisler, George	1920	.4006	.4073	14
5	Heilmann, Harry	1923	.4004	.4027	20
6	Hornsby, Rogers	1921	.4000	.3970	28
7	Williams, Ted	1957	.3995	.3881	48
8	Lajoie, Nap	1910	.3993	.3841	62
9	Williams, Ted	1941	.3991	.4057	16
10	Cobb, Ty	1918	.3981	.3824	69
13	Cobb, Ty	1911	.3972	.4196	7
32	Lajoie, Nap	1901	.3834	.4265	3
50	Delahanty, Ed	1899	.3764	.4096	9

c. Seasonal hits

POS	PLAYER	YEAR	ADJ-H	REAL-H	RP†
1T	Alou, Felipe	1968	253	210	193T
1T	Torre, Joe	1971	253	230	34T
3	Alou, Matty	1969	252	231	30T
4T	Rose, Pete	1968	251	210	193T
4T	Rose, Pete	1973	251	230	34T
6	Sisler, George	1920	250	257	1
7	Musial, Stan	1946	249	228	40T
8	Hornsby, Rogers	1920	244	218	97T
9T	Carew, Rod	1977	242	239	15T
9T	Lajoie, Nap	1910	242	227	42T
11T	Hornsby, Rogers	1922	241	250	5T
14T	Sisler, George	1922	239	246	8
18T	Suzuki, Ichiro*	2001	238	242	9
32T	Manush, Heinie	1928	234	241	10T
39T	Simmons, Al	1925	231	253	4
54T	Cobb, Ty	1911	226	248	7
54T	O'Doul, Lefty	1929	226	254	2T
66T	Terry, Bill	1930	224	254	2T
154T	Klein, Chuck	1930	215	250	5T
213T	Herman, Babe	1930	211	241	10T

d. Seasonal RBI

POS	PLAYER	YEAR	ADJ-RBI	REAL-RBI	RP†
1	Gehrig, Lou	1931	169	184	2
2	Gehrig, Lou	1927	168	175	4T
3T	Davis, Tommy	1962	161	153	34T
3T	Sosa, Sammy*	2001	161	160	21
5	Greenberg, Hank	1937	160	183	3
6	Torre, Joe	1971	159	137	115T
7T	Aaron, Hank	1963	158	130	174T
7T	Ruth, Babe	1927	158	164	17
9	Greenberg, Hank	1935	157	170	8T
10T	Killebrew, Harmon	1969	156	140	97T
10T	Medwick, Joe	1937	156	154	32T
10T	Sosa, Sammy*	1998	156	158	26
13T	Wilson, Hack	1930	155	191	1
17T	Ruth, Babe"	1921	153	171	7
23T	Gehrig, Lou	1930	151	174	6
33T	Foxx, Jimmie	1938	148	175	4T
46T	Foxx, Jimmie	1932	146	169	10
133T	Klein, Chuck	1930	134	170	8T

e. Seasonal OPS

POS	PLAYER	YEAR	ADJ-OPS	REAL-OPS	RP†
1	Bonds, Barry*	2002	1419	1383	1
2	Bonds, Barry*	2001	1394	1380	2
3	Ruth, Babe	1920	1355	1379	3
4	Ruth, Babe	1921	1305	1359	4
5	Williams, Ted	1957	1292	1259	7
6	Ruth, Babe	1923	1291	1309	5
7	Williams, Ted	1941	1269	1287	6
8	Ruth, Babe	1926	1243	1253	9
9	Ruth, Babe	1927	1236	1258	8
10	McGwire, Mark	1998	1225	1225	13T
14T	Ruth, Babe	1924	1219	1252	10

f. Seasonal triples

POS	PLAYER	YEAR	ADJ-3B	REAL-3B	RP†
1T	Cuyler, Kiki	1925	24	26	9T
1T	Myers, Hy	1920	24	22	32T
1T	Stirnweiss, Snuffy	1945	24	22	32T
4T	Combs, Earle	1927	23	23	22T
4T	Mitchell, Dale	1949	23	23	22T
6T	Hornsby, Rogers	1920	22	20	75T
6T	Musial, Stan	1943	22	20	75T
6T	Musial, Stan	1946	22	20	75T
6T	Waner, Paul	1926	22	22	32T
10T	Combs, Earle	1928	21	21	52T
10T	Daubert, Jake	1922	21	22	32T
10T	Johnson, Lance	1996	21	21	52T
10T	Mays, Willie	1957	21	20	75T
10T	Roush, Edd	1924	21	21	52T
10T	Terry, Bill	1931	21	20	75T
10T	Vaughan, Arky	1933	21	19	112T
10T	Walker, Curt	1926	21	20	75T
10T	Wilson, Willie	1985	21	21	52T
***	Wilson, Chief	1912	11	36	1

* Active player
** OPS includes HBP but not SF
*** Player not in top 500
† Real position

Table 3. Comparison of adjusted and real career statistics for selected players through 2002

TYPE	PLAYER	AB	R	H	2B	HR	RBI	BA	OBP*	SLG*	OPS*	SB	BB	RC**	RC25**
ADJ	Cap Anson	9496	1547	2637	554	310	1631	.2777	.3328	.4416	774	137	760	1389	5.06
REAL	Cap Anson	10277	1996	3418	581	97	2076	.3326	.3926	.4451	838	276	983	1788	6.52
ADJ	Willie Keeler	8213	1616	2554	307	259	746	.3110	.3559	.4529	809	249	463	1294	5.71
REAL	Willie Keeler	8591	1719	2932	241	33	810	.3413	.3878	.4146	802	495	524	1351	5.97
ADJ	Honus Wagner	10283	1936	3268	748	462	1911	.3178	.3810	.5409	922	359	930	2082	7.42
REAL	Honus Wagner	10430	1736	3415	640	101	1732	.3274	.3910	.4662	857	722	963	1868	6.66
ADJ	Sam Crawford	9502	1618	2893	579	488	1780	.3045	.3564	.5403	897	181	745	1823	6.89
REAL	Sam Crawford	9570	1391	2961	458	97	1525	.3094	.3616	.4522	814	366	760	1559	5.9
ADJ	Ty Cobb	11426	2476	4181	815	399	2141	.3659	.4325	.5688	1001	507	1246	2783	9.6
REAL	Ty Cobb	11434	2246	4189	724	117	1937	.3664	.4330	.5120	945	892	1249	2510	8.66
ADJ	Tris Speaker	10137	2026	3456	855	332	1638	.3409	.4238	.5480	972	247	1356	2326	8.7
REAL	Tris Speaker	10195	1882	3514	792	117	1529	.3447	.4279	.5003	928	432	1381	2157	8.07
ADJ	Babe Ruth	8302	2098	2776	505	733	2138	.3344	.4653	.6875	1153	115	1990	2643	11.96
REAL	Babe Ruth	8399	2174	2873	506	714	2213	.3421	.4739	.6897	1164	123	2062	2733	12.36
ADJ	Lou Gehrig	7861	1735	2581	508	467	1834	.3283	.4342	.6109	1045	99	1427	2072	9.81
REAL	Lou Gehrig	8001	1888	2721	534	493	1995	.3401	.4474	.6324	1080	102	1508	2250	1.66
ADJ	Jimmie Foxx	7960	1568	2472	428	496	1717	.3106	.4117	.5809	993	80	1355	1900	8.65
REAL	Jimmie Foxx	8134	1751	2646	458	534	1922	.3253	.4283	.6093	1038	87	1452	2119	9.65
ADJ	Joe DiMaggio	6774	1340	2167	379	351	1481	.3199	.3939	.5697	964	30	781	1506	8.17
REAL	Joe DiMaggio	6821	1390	2214	389	361	1537	.3246	.3983	.5788	977	30	790	1558	8.46
ADJ	Ted Williams	7704	1794	2652	520	522	1830	.3442	.4831	.6334	1117	24	2031	2348	11.62
REAL	Ted Williams	7706	1798	2654	525	521	1839	.3444	.4827	.6338	1116	24	2021	2347	11.62
ADJ	Stan Musial	11087	2047	3745	750	487	2043	.3378	.4257	.5702	996	81	1645	2676	9.11
REAL	Stan Musial	10972	1949	3630	725	475	1951	.3308	.4184	.5591	977	78	1599	2551	8.69
ADJ	Mickey Mantle	8271	1857	2584	369	571	1670	.3124	.4395	.5820	1022	159	1863	2112	9.29
REAL	Mickey Mantle	8102	1677	2415	344	536	1509	.2981	.4225	.5568	979	153	1733	1903	8.36
ADJ	Willie Mays	11092	2279	3494	555	699	2099	.3150	.4016	.5804	982	358	1562	2572	8.46
REAL	Willie Mays	10881	2062	3283	523	660	1903	.3017	.3867	.5575	944	338	1464	2333	7.67
ADJ	Hank Aaron	12637	2434	4044	671	811	2566	.3200	.3937	.5818	975	261	1504	2884	8.39
REAL	Hank Aaron	12364	2174	3771	624	755	2297	.3050	.3772	.5545	932	240	1402	2576	7.5
ADJ	H. Killebrew	8297	1437	2236	308	614	1773	.2695	.3955	.5344	930	20	1681	1741	7.18
REAL	H. Killebrew	8147	1283	2086	290	573	1584	.2560	.3786	.5085	887	19	1559	1556	6.42
ADJ	R. Clemente	9686	1602	3232	471	258	1478	.3337	.3792	.4988	878	88	674	1821	7.06
REAL	R. Clemente	9454	1416	3000	440	240	1305	.3173	.3616	.4751	837	83	621	1614	6.25
ADJ	Frank Robinson	10214	2044	3151	564	627	2028	.3085	.4093	.5622	971	215	1531	2289	8.1
REAL	Frank Robinson	10006	1829	2943	528	586	1812	.2941	.3924	.5370	929	204	1420	2052	7.26
ADJ	Willie Davis	9419	1417	2806	432	197	1211	.2979	.3338	.4384	772	435	455	1364	5.16
REAL	Willie Davis	9174	1217	2561	395	182	1053	.2792	.3142	.4118	726	398	418	1173	4.44
ADJ	Joe Torre	8073	1140	2541	371	273	1358	.3148	.3861	.4775	864	23	845	1464	6.61
REAL	Joe Torre	7874	996	2342	344	252	1185	.2974	.3669	.4521	819	23	779	1284	5.8
ADJ	Dick Allen	6516	1282	2032	352	387	1305	.3118	.4036	.5704	974	148	986	1495	8.34
REAL	Dick Allen	6332	1099	1848	320	351	1119	.2919	.3808	.5336	914	133	894	1282	7.15
ADJ	Pete Rose	14407	2459	4610	808	172	1487	.3200	.3958	.4319	828	214	1693	2436	6.22
REAL	Pete Rose	14053	2165	4256	746	160	1314	.3029	.3770	.4093	786	198	1566	2144	5.47
ADJ	Cal Ripken Jr.	11548	1644	3181	603	431	1690	.2755	.3436	.4473	791	36	1133	1757	5.25
REAL	Cal Ripken Jr.	11551	1647	3184	603	431	1695	.2756	.3436	.4474	791	36	1129	1758	5.25
ADJ	Tony Gwynn	9449	1500	3302	566	139	1230	.3495	.4036	.4725	876	339	834	1796	7.3
REAL	Tony Gwynn	9288	1383	3141	543	135	1138	.3382	.3915	.4585	850	319	790	1661	6.76
ADJ	Barry Bonds	8443	1948	2570	535	635	1752	.3044	.4417	.6114	1053	516	2002	2260	9.62
REAL	Barry Bonds	8335	1830	2462	514	613	1652	.2954	.4315	.5952	1027	493	1922	2120	9.03
ADJ	Mark McGwire	6186	1168	1625	253	578	1415	.2627	.3976	.5858	983	12	1312	1420	7.78
REAL	Mark McGwire	6187	1167	1626	252	583	1414	.2628	.3982	.5882	986	12	1317	1427	7.82
ADJ	Sammy Sosa	7031	1222	1960	297	497	1347	.2788	.3513	.5456	897	236	739	1332	6.57
REAL	Sammy Sosa	7026	1215	1955	297	499	1347	.2783	.3508	.5458	897	233	738	1330	6.56

* OBP, SLG, and OPS include BB, but not SF
** RC and RC25 use the basic RC formula

verted the RC based upon the players' team scoring 750 runs,[8] adjusting for park factors.[9] The adjusted runs created (RC_{ADJ}) was then substituted into the hits formula above, using the old ratios (with the conversions for pre-1920, where appropriate) for all other elements in the formula. This then gives us H_{ADJ}. The ratio between H_{ADJ} and H was then used to compute BB_{ADJ}, TB_{ADJ}, and most other counting statistics. The ratio between RC_{ADJ} and RC was then used to compute R_{ADJ}, and RBI_{ADJ}. As with the James's algorithms, games played, batting outs (AB – H) and strikeouts remained the same.

DISCUSSION

Not surprisingly, the players with the most change were the pre-1920 players, due to the Deadball Era algorithm (especially players in the early years of baseball, who because teams of that era scored so many runs, had their statistics decrease dramatically—except home runs, of course, which due to the Deadball Era algorithm still rose greatly). Of post-1920 players, the players most affected on the negative side were not surprisingly, players of the 1930s. On the positive side, also not surprisingly, players of the 1940s-1950s and 1960s-1970s were most affected. The players of today were not so greatly affected (except for park effects) because, except for a few exceptions in recent years, average runs per team in the leagues have been close to 750 per year. In addition, players who have played for longer have had any big league-wide run-producing years offset by lower league-wide run-producing years.

For post-1920 players, the most affected negatively overall seems to be Jimmie Foxx, who moves out of the 500 home run club (Real: .325/534/1,922/1,038 OPS (note OPS calculation includes hit by pitch, but not sacrifice flies) vs. Adjusted: .311/496/1,717/993 OPS). Foxx has the third largest decline in OPS (-45) among players with at least 1,000 career AB's—the first two being Todd Helton (-53) and Earl Averill (-52). The most affected positively overall seems to be Dick Allen (Real: .292/351/1,119 914 OPS vs. Adjusted: .311/387/1,305 974 OPS). The players closest to their original stats are probably Ted Williams (Real: .344/521/1,839 1116 OPS vs. Adjusted: .344/522/1,830 1117 OPS), Cal Ripken (Real: .276/431/1,695 791 OPS vs. Adjusted: .276/431/1,690 791OPS), and Sammy Sosa (Real: .278/499/1,347 897 OPS vs. Adjusted: .279/497/1,347 897 OPS).

In terms of famous records, Hank Aaron's HR record becomes 811. Three players join the 600 home run club (Frank Robinson (627), Harmon Killebrew (614), and Reggie Jackson (600)). Willie Mays just misses the 700 home run club with 699. Pete Rose gets 4,610 hits, 362 more than Ty Cobb (4,181). Hank Aaron comes much closer to Cobb than in real life with 4,044 hits. Overall, 25 players now have at least 3,000 hits. This includes Frank (3,151) and Brooks Robinson (3,091); the only players to move into the 3,000-hit plateau who are not there in real life. Two players move out of the 3,000 hit plateau: Wade Boggs (2,982), who had 3,010 hits in real life and Cap Anson (2,637), who had 3,418 hits in real life (a difference of almost 23%).

Ty Cobb still leads in career average (still at .366). Tony Gwynn moves all the way up to fourth (.350), and Rod Carew moves to sixth (.341). Table 1 shows the top 10 career leaders in various categories. Table 2 shows the leaders in various single-season categories (which are discussed below). The top five pre-1920 players (defined for career leaders as those players starting their careers before 1910 or ending their careers before 1920), with their position in the overall leaders, are included in the career and single season home runs list. In both of the tables I've also included the rest of the real top 10 and their position on the adjusted list, if they did not appear on the adjusted list already. Triples are included in both tables because they are, without question, the most affected statistic (in terms of leaders) due to the Deadball Era algorithm.

For single-season records, no asterisk was necessary for Roger Maris, who now hits 63 home runs in 1961, five more than Babe Ruth's 1927 total of 58. The closest to Ruth before Maris is now Ralph Kiner, who still hits 54 in 1949. Of the other players who came closest to Ruth in real life, Jimmie Foxx's total of 58 in 1932 becomes 53, Hank Greenberg's 58 in 1938 becomes 52, and Hack Wilson's 56 in 1930 becomes 49. Mark McGwire still hits 70 in 1998, but the current record is now 74 by Barry Bonds instead of 73. For RBI, Hack Wilson's former total of 191 in 1930 is now no better than a tie for 13th with George Foster in 1977 (155). Lou Gehrig has the top two spots in RBI (169 and 168 in 1931 and 1927, respectively). Sammy Sosa is now tied for third place with Tommy Davis (161 in 2001 and 1962, respectively). Based on the Deadball Era algorithm, five players hit 40 home runs or more prior to 1920 (Babe Ruth (46 in 1919), Buck Freeman (45 in 1899), Frank Schulte (44 in 1911), Chief Wilson (42 in 1912), and Gavvy Cravath (42 in 1915).

Only four players hit .400 in a season (rounded to the nearest thousandth) a total of six times. Rogers Hornsby leads with .428 in 1924, 15 points ahead of George Sisler's .413 in 1922. Hornsby and Sisler do it twice; Hornsby hits exactly .400 in 1921 and Sisler, in 1920, hits .401. The other players to hit .400 are Harry Heilmann (.401 in 1923), Ty Cobb (.401 in 1922 – while Cobb only hits .400 once, he hits over .390 no less than eight times), and Ted Williams (.3995) in 1957 (would Williams have considered that hitting .400?). Williams also hits over .399 in 1941 (.3991—Williams probably wouldn't have been happy about that, either). In recent years, Tony Gwynn's 1994 average becomes .397, George Brett's average in 1980 becomes .393, and Rod Carew's average in 1977 becomes .391. Also, Joe Torre hits .385 in 1971 and Barry Bond's 2002 average becomes .382.

See Table 3 for a comparison of real and adjusted statistics for selected players.

NOTES

1. In the 2003-04 McFarland Baseball Books catalog alone, more than one dozen books are available which compare players and/or teams from one era to another.
2. Thorn, John, et. al., *Total Baseball*, 6th Edition, Total Sports, 1999, p. 2,534.
3. James, Bill, "Runs Created," in Bill James, et. al., *Bill James Presents Stats All-Time Major League Handbook*, Stats, Inc., 1998, p. 7.
4. Note that, in this paper, the basic RC formula

$$((H + BB) \times (TB) \div (AB + BB))$$

is used for all years, regardless of the availability of data to complete the more advanced runs-created formulas.
5. For example, without going into the numbers, for many years of Babe Ruth's career if you try to adjust his home runs to league averages and then compute them for a typical home run year in baseball history, Ruth comes out with more home runs than at bats.
6. The algorithm is:
   1. Games played remain the same.
   2. Batting outs (AB - H) remain the same.
   3. The relationship between productivity as a hitter and the league average remains exactly the same.

   To complete (3) find the difference between team's runs scored vs. 750. Multiply this index by the player's real runs created to get the adjusted runs created. Then adjust this for park factor (which is modified based on the fact that half of the games are not played in that park). From there you enter the adjusted runs created in the hits formula (see methodology) to find the adjusted hits. Counting statistics rise and fall with hits. Productivity statistics (e.g., RBI, runs scored) rise and fall with runs created.
7. The Deadball Era algorithm includes the three elements of the first algorithm plus:
   4. 67% of triples become home runs.
   5. 3% of batting outs become home runs.
   6. 2% of batting outs become doubles.
   7. 50% of stolen bases disappear.
   8. Hits are pegged at whatever level creates the appropriate level of offense (the change in runs created).
   9. Everything else rises and falls with hits or total bases (as in the first algorithm).
   10. Note that in order for the hits to come out right, you also must assume that 5% of batting outs are taken away from singles (in order for the batting outs to remain the same, those 5% of batting outs which have been allotted to doubles and triples must come from somewhere). James does not mention this in the text, however, so it is possible that he might have figured out some other way to account for the change in batting outs.
8. James arbitrarily chose 750 runs as what seemed to him to be a "normal context" for runs scored. However, according to my calculations since 1920 the average number of runs scored per team in both leagues is very close to 700 (699.6). I used 750 anyway to remain consistent with James. Note that for players who switched teams during the year, the runs scored for the entire year is used even though a larger or smaller proportion of the runs may have been scored during the time the player was with the team. Also, the 750 runs are based upon 162 games; so in games-shortened seasons (such as for strike, war, or pre-expansion years), players will not have their statistics altered as if a 162-game schedule was played.
9. When I attempted to check some of my results against those of James's, the adjusted runs created were slightly off (59 for my analysis; 63 for James), thereby causing differences in the corresponding statistics. You can see these differences by checking the Adjusted Career Stats for Davis in Table 3 versus those in the James book on page 743. I believe this was due to the way James calculates park factors vs. the way the Lahman database does (I know that the Lahman database uses three-year park factors). When I plugged in the BPF for 1965 that James used for Davis (76 vs. 93 for the Lahman database) the adjusted runs created came out the same. Since 1965 was the only season that James mentions the BPF he used, it was the only season I could check.

PETER UELKES AND RON VISCO

Coincidences

On August 8, 1979, the visiting Milwaukee Brewers defeated the hometown Baltimore Orioles 8-4. The visitors scored a run in the top of the first inning, but the home team came back with three runs in the bottom half. Slowly, with a run in the fourth and another in the seventh, the visitors fought back to tie the game. Each team scored once in the eighth, but the visitors broke it open with four runs in the ninth to take the victory. The line score was:

```
VISITORS    100 100 114 - 8
HOME        300 000 010 - 4
```

A good game, close until the ninth, but unremarkable.

There are 6,370,650 different line scores that result in an 8-4 nine-inning win by the visiting team such as this one. An earlier article, "Let Me Count the Ways," (*BRJ*, No. 30, 2001) discussed the number of different line scores that could result in a game where each team scores nine runs or less. A small table summarized the number of "ways" for a team to score a given number of runs in nine innings (that is, for that number of runs to be distributed among the nine innings). Here is an expanded version:

Table 1. Possible ways for a team to score runs in 9 innings.

RUNS	WAYS
0	1
1	9
2	45
3	165
4	495
5	1,287
6	3,003
7	6,435
8	12,870
9	24,310
10	43,758
11	75,582
12	125,970

PETER UELKES *got a Ph.D. in particle physics and is currently working as a business analyst for the Vodafone group. He is a SABR member since 2001.* **RON VISCO** *works in the education department at the National Baseball Hall of Fame, and has been a SABR member since 1983.*

In this article, "nine innings" means a game which featured between 51 and 54 outs; this includes games which go to the bottom of the ninth where either the home team is ahead (and they do not bat) or where they score the winning run with one or two outs.

Using the 8-4 game above as an illustration, there are 12,870 ways for the visitors to score eight runs, and 495 ways for the home team to score four, so there are 12,870 × 495 = 6,370,650 ways to get an 8-4 visiting team victory (in nine innings): over six million ways! See the earlier article for more details.

This number not only exceeds the number of games ending in an 8-4 road win, but also far exceeds the number of games in major league history. Such results led the authors to hypothesize that games in which nine or more runs have been scored may have line scores unique in the history of major league baseball: that is, line scores in such relatively high-scoring games have never been duplicated.

In order to determine the *probability* that a line score is unique, more information is needed: the actual number of games played that resulted in the given score. For instance, how many nine-inning 8-4 games that resulted in visiting team victories have actually been played?

To address this issue, an analysis was conducted using a data set downloaded from Retrosheet; this online resource contains line scores from all major league games for the period 1978 to 2000 (inclusive). A number of analytic procedures, mainly scripts to be run in a Unix environment, were written. Their purpose was to determine, within the data set:

- The number of games with a given final score
- The *probabilities* for matching (duplicate) line scores for each given final score
- The *number* of matching line scores for each given final score

One focus, in particular, was to uncover matching line scores from games with high run totals (especially nine or more runs), where such matches might not be expected to occur. A match which is relatively unlikely to occur (but does) will be called a "coincidence."

To understand how to calculate the probability of a coincidence, consider an application of a mathematical concept often called the "birthday paradox." Suppose there are 25 people in a room (a party, classroom, etc.). What is the probability that at least two of them share the same birthday (month and day)? It's easier to first calculate the probability that no one shares a birthday. Choose someone and note the birthday. Then randomly choose another person. The probability that the second person has a different birthday is 364/365 (we'll ignore leap days for simplicity). The probability that a third person has a birthday different from the other two is 363/365. And so on. Thus, the probability that all 25 people have different birthdays is given by the product:

$$p = 364/365 \times 363/365 \times 362/365 \times \ldots \times 341/365$$

which is approximately p – .43; that's the probability of *no* birthday matches. We subtract this figure from 1 (because a probability of 1 is the highest possible, and means it's certain) to obtain the probability of at least two people in the room sharing a birthday: 1 – .43 = .57. In other words, there is a 57% chance that at least two people will share birthdays, a result which may seem surprising with only 25 people.

The same technique was applied to determining the probability of finding, or *not* finding, a coincidence in line scores for each given score. For this analysis, figures were tabulated separately for visitor wins and home wins. Only games with nine or more runs are considered here. Matches on line scores occur more often for low-scoring contests. For some scores it is certain that matches will occur. For instance, there are only nine different line scores for a 1-0 home victory, but 462 such outcomes were found in the available data set, so inevitably there are many matches.

Table 2 presents the findings for home team victories. The first two columns give the different possible outcomes, such as a 9-0 home win. The third column gives the number of possible line scores with that outcome. The fourth column shows how many such scores are known, based on the computer search. The fifth column shows the probability that *no* duplicate line

Table 2. Possibilities, probabilities, and coincidences (Cn) given a home team victory

	RUNS (V)	RUNS (H)	POSSIBLE LINE SCORES	KNOWN OCCURRENCES	PROB. OF NO Cn	KNOWN Cn
9 RUNS	0	9	11,440	87	.7205	0
	1	8	57,915	214	.6743	2
	2	7	154,440	420	.5654	3
	3	6	310,365	585	.5765	1
	4	5	637,065	1,002	.4549	1
10 RUNS	0	10	19,448	49	.9413	0
	1	9	102,960	177	.8595	0
	2	8	289,575	297	.8591	0
	3	7	593,505	471	.8298	0
	4	6	1,094,445	588	.8541	0
11 RUNS	0	11	31,824	40	.9758	0
	1	10	175,032	113	.9645	0
	2	9	514,800	219	.9547	1*
	3	8	1,061,775	328	.9507	0
	4	7	1,943,865	402	.9594	0
	5	6	3,864,861	719	.9354	0
12 RUNS	0	12	50,388	20	.9962	0
	1	11	286,416	63	.9932	0
	2	10	875,160	137	.9894	0
	3	9	1,887,600	227	.9865	0
	4	8	3,430,350	291	.9878	0
	5	7	6,073,353	361	.9894	0

* 100 000 100 2 by Cincinnati at LA, June 12, 1989
201 010 320 – 9 and Boston at Texas, August 21, 1999

Table 3. Possibilities, probabilities, and coincidences (Cn) given a visiting team victory

	RUNS (V)	RUNS (H)	POSSIBLE LINE SCORES	KNOWN OCCURRENCES	PROB. OF NO Cn	KNOWN Cn
9 RUNS	9	0	24,310	68	.9105	0
	8	1	115,830	211	.8258	0
	7	2	289,575	411	.7474	0
	6	3	495,495	553	.7348	2
	5	4	637,065	639	.7261	0
10 RUNS	10	0	43,758	59	.9616	0
	9	1	218,790	137	.9583	0
	8	2	579,150	265	.9414	0
	7	3	1,061,775	456	.9069	0
	6	4	1,486,485	548	.9041	0
11 RUNS	11	0	75,582	39	.9902	0
	10	1	393,822	92	.9894	0
	9	2	1,093,950	206	.9809	0
	8	3	2,123,550	313	.9773	0
	7	4	3,185,325	398	.9755	0
	6	5	3,864,861	439	.9754	0
12 RUNS	12	0	125,970	18	.9988	0
	11	1	680,238	79	.9955	0
	10	2	1,969,110	124	.9961	0
	9	3	4,011,150	203	.9949	0
	8	4	6,370,650	288	.9935	1
	7	5	8,281,845	349	.9927	0

score, or coincidence, will occur; this was calculated using the birthday paradox approach and the figures from the previous two columns. Finally, the last column gives the number of actual coincidences that were found in the search.

The hypothesis presented in "Let Me Count the Ways" proved to be incorrect! There have been a handful of coincidences among games with nine runs. Note that there have been so many 5-4 home team victories (over 1,000), despite over 637,000 possible line scores with that result, the probability is less than 50% that there would be no matches. Indeed, there is one. Keep in mind that the probability of matching a *particular* 5-4 line score would be very small, just as the probability of someone else in a room of 25 people sharing *your* birthday is quite small. The probability reported in Table 2 considers *any* 5-4 road win line score matches, not of matching one in particular.

What is more surprising is that we find line score matches among the 8-1, 7-2, and 6-3 home victories, which we do. None of these matches was likely, although none was unlikely to the extent found in higher-scoring games.

In passing, it's worth noting that within any of the "total runs" categories (9, 10, 11, and 12), the number of known games increases as the outcome gets closer. For example, among nine-run contests there are 87 9-0 outcomes, but the number of occurrences increases substantially as the score distribution changes to 8-1, 7-2, 6-3, and finally 5-4. Another way of stating that result, based on the findings presented here, is that the more *possible* ways to reach a given score or run total, the more times it has *actually* occurred.

As we move to 10, 11, and 12-run home team victories, Table 2 shows that the probability of a coincidence for a given case is quite small. Nonetheless, in an 11-run game where the probability of no matches is 95.5 we find a coincidence!

In Table 3 are the corresponding figures for visiting team victories. We find two coincidences for nine total runs, both happening in 6-3 games. However, the real surprise comes further down the table, where there are matching line scores in 12-run games.

On June 3, 1988, the visiting Houston Astros defeated the hometown San Francisco Giants 8-4. The visitors scored a run in the top of the first inning, but the home team came back with three runs in the bottom half. Slowly, with a run in the fourth and another in the seventh, the visitors fought back to tie the game. Each team scored once in the eighth, but the visitors broke it open with four runs in the ninth to take the victory. The line score:

```
VISITORS    100 100 114 - 8
HOME        300 000 010   4
```

It was a good game, close until the ninth, but not really remarkable. Except for one thing: this was the exact description, the exact line score, that had occurred in the Brewers-Orioles game nine years earlier. One chance in six million: now, *that's* a coincidence!

SHOELESS DAVE **Submitted by David MacDonald**

"Joe Jackson's brother Dave, whom Joe recommended to the Cleveland Baseball Club last season, is out of baseball. His left arm was broken last winter and is now paralyzed. Joe asserts that brother Dave would have been the star of the Jackson family but for the accident to his arm." —*Ottawa Journal*, April 2, 1912

Corrections

The following errors in previous journals have been noted by readers and/or the contributors themselves. We have not noted the occasional misspelled word or grammatical error, unless it is a misspelling of a name. If you come across errors in future or past journals, please contact the editorial director in writing and we will post it.

THE NATIONAL PASTIME #22

p. 20 Huntington is misspelled as Huntingdon in a reference to Huntington Grounds.

p. 41 Ray Kremer is misspelled as Ray Kramer.

p. 125 The sentence reads: "I asked [Greenberg] about the incident with Dick Bartell. It had been contended that Bartell held the relay too long, thus enabling the winning run to score in the seventh game." It was the tying run, not the winning run as stated. Adie Suehsdorf quotes Dick Bartell: "Slow as he is, I decided there is no play, no assurance that even a perfect throw will catch him. . . Remember, there were no outs and McCormick was only the tying run. Perhaps I should've made the throw, but it seemed safer not to."

BASEBALL RESEARCH JOURNAL #31

p. 1 Peer reviewer John Mathew's name is misspelled as John Mathews.

p. 61 Walter Johnson is listed as holding the record for pitching the most games (802) with one team. Elroy Face also pitched in 802 games for the Pittsburgh Pirates.

p. 64 Babe Ruth is listed as having the still standing American League record for runs scored with 2,174. Ty Cobb holds the record with 2,245.

p. 68 The sentence states that Frank Farrell, the first co-owner of the Yanks, played second base on the 1884 Providence Grays. It was Jack Farrell who played for the Grays.

p. 81 Second paragraph. Bill White was not the National League MVP in 1965.

p. 98 Equation (1) should be RBI/AB = .187 × OPP + .196 × AVG + .468 × ISO = .303. This change was made after adjusting for discrepancies that were discovered in the original data set after publication. Equation (2), some discussion and the table have also been adjusted. Details are available in SABRzine or <members.aol.com/cjmorong/myhomepage/WEB.htm> or call Cyril Morong at 210-734-4703 for the revised paper.

p. 106 The sentence misstates that Babe Ruth led the league in RBI eight times. He led six times.

THE NATIONAL PASTIME #23

p. 2 The text refers incorrectly to Billy Williams and Curt Flood pictured on the back cover. It is Lou Brock with Williams.

p. 3 In Art Ahrens' biography, it should read the Los Angeles Dodgers and not Brooklyn Dodgers.

p. 26 Third paragraph. The Cardinals won the 1942 World Series, not the Yankees.

p. 27 Second column, first paragraph. The last sentence should read Ralph Branca and not Cy Buker.

p. 55 In the article on the Robinsons in Montreal there is an omission. Montreal columnist Bertrand Raymond should be credited with writing the original version of this in French in 1987.

p. 57 First paragraph. The year should be 1862, not 1962.

p. 78 Last paragraph. Referring to the end of the World Series, the sentence reads, "Hoy passed away six weeks later on December 15, 1961." The time was about ten weeks later.

p. 81 The filler misstates that Gowdy (WWI and WWII) was the only major league vet to serve in those wars. There were others. Also, it is stated that the Marines were the only branch of the service to recall WWII vets for the Korean War. Walter Kephart, who served in WWII, says that this is not so.

p. 84 John Carden's 1948 record with Sioux City (3–0) was omitted. Also, the 1948 official Inter-State League averages show Carden with 139 hits, not the 130 in the article's record. In the sources for Carden, SABR member Davis Barker is misidentified as David Barker.

P. 113 Second paragraph. The sentence states that baseball has awarded home field advantage to the team with the better record "since the advent of divisional play in 1969." Starting in 1969, the ALCS began in the AL East with two games and then finished with one, two, or three games in the AL West in odd-numbered years and reversed in even-numbered years. The NLCS began in the NL West in odd-numbered years and in the NL East in even-numbered years. The regular season records were not a factor in selecting the home team.

SOCIETY FOR AMERICAN BASEBALL RESEARCH

Since August 1971, when sixteen "statistorians" gathered in Cooperstown to form the Society for American Baseball Research, SABR has been committed to helping people produce and publish baseball research.

Today, SABR has nearly 7,000 members worldwide. They come from all walks of life—the one thing they all have in common? A love for the game and its history.

Members receive the latest editions of SABR's research annuals, *The Baseball Research Journal* and *The National Pastime*. Also included is a subscription to *The SABR Bulletin*, special access to online newsgroups and research forums, and other special publications.

SABR membership is open to all those interested in baseball research. Annual dues are $50 US, $60 Canada and Mexico, and $65 overseas (US funds only). Student and senior discounts are also available. For details about the benefits of SABR membership, call (800) 969-SABR or visit **www.sabr.org** today!

SABR

SOCIETY FOR AMERICAN BASEBALL RESEARCH
812 HURON ROAD, CLEVELAND, OH 44115 (800)969-SABR

Pictured publications have been previously issued. For information about past SABR publications, call (800) 755-1105.

SABR BACK PUBLICATIONS ORDER FORM

THE BASEBALL RESEARCH JOURNAL

THE BASEBALL RESEARCH JOURNAL, the annual publication of the society, features some of the best member research. Articles range from statistical to biographical sketches, plus nearly every other topic in baseball.

- ☐ 1985 (88pp) **$6.00**
- ☐ 1987 (88pp) **$6.00**
- ☐ 1988 (88pp) **$7.00**
- ☐ 1990 (88pp) **$8.00**
- ☐ 1991 (88pp) **$8.00**
- ☐ 1992 (88pp) **$7.95**
- ☐ 1993 (112pp) **$9.95**
- ☐ 1994 (112pp) **$9.95**
- ☐ 1995 (144pp) **$9.95**
- ☐ 1996 (154pp) **$9.95**
- ☐ 1997 (144pp) **$9.95**
- ☐ 1998 (116pp) **$9.95**
- ☐ 1999 (144pp) **$12.00**
- ☐ 2000 (144pp) **$12.00**
- ☐ 2001 (136pp) **$12.00**

SABR BOOKS ON THE 19th CENTURY

☐ **NINETEENTH CENTURY STARS**
Biographies of America's First Heroes, Non-Hall of Famers (1988, 144pp)**$10.00**

☐ **BASEBALL'S FIRST STARS**
More biographies, includes Hall of Famers(1996, 183pp) **$14.95**

☐ **BASE BALL: HOW TO BECOME A PLAYER**
1888 reprint by John Montgomery Ward (1993, 149 pp) **$9.95**

THE NATIONAL PASTIME

THE NATIONAL PASTIME features articles by members more general in nature, although some volumes are arranged in a theme, as noted below.

- ☐ **#3, 19TH CENTURY PICTORIAL** (Spring 1984, (88 pp) **$7.00**
- ☐ **#4,** (Spring 1985, 88 pp) **$6.00**
- ☐ **#5,** (Winter 1985, 88 pp) **$6.00**
- ☐ **#7,** (Winter 1987, 88 pp) **$6.00**
- ☐ **#10,** (Fall 1990, 88 pp) **$8.00**
- ☐ **#11,** (Fall 1991, 88 pp) **$7.95**
- ☐ **#12, THE INTERNATIONAL PASTIME** (Summer 1992, 96 pp) **$7.95**
- ☐ **#13,** (Summer 1993, 96 pp) **$7.95**
- ☐ **#14,** (Summer 1994, 112 pp) **$9.95**
- ☐ **#15,** (Spring 1995, 156 pp) **$9.95**
- ☐ **#16,** (Spring 1996, 144 pp) **$9.95**
- ☐ **#17,** (Spring 1997, 144 pp) **$9.95**
- ☐ **#19,** (Summer 1999, 116 pp) **$12.00**
- ☐ **#20,** (Summer 2000, 132 pp) **$12.00**
- ☐ **#21** (Summer 2001, 124 pp) **$12.00**

SABR REVIEW OF BOOKS

- ☐ **VOLUME 1** (1986) **$6.00**
- ☐ **VOLUME 2** (1987) **$6.00**
- ☐ **VOLUME 3** (1988) **$7.00**
- ☐ **VOLUME 4** (1989) **$7.00**

☐ **ALL-STAR BASEBALL IN CLEVELAND**
(1997 Special Publication, 64pp) **$7.95**

☐ **BASEBALL FOR THE FUN OF IT**
A pictorial about the joy of baseball (1997, 92pp) **$14.95**

☐ **COOPERSTOWN CORNER**
Columns from **The Sporting News** by Lee Allen (1990, 181pp) **$10.00**

☐ **HOME RUNS IN THE OLD BALLPARKS**
Listings of top 5 HR hitters in parks no longer in use (1995) **$9.95**

BIOGRAPHIES BY SABR

☐ **LEFTY GROVE: AMERICAN ORIGINAL**
Bio of HOF pitcher Lefty Grove written by Jim Kaplan (2000, 315pp) **$12.95**

☐ **UNCLE ROBBIE**
Bio of HOF manager Wilbert Robinson by Jack Kavanagh and Norman Macht (1999, 200pp) **$12.95**

☐ **ADDIE JOSS: KING OF THE PITCHERS**
Bio of HOF pitcher Addie Joss by Scott Longert (1998, 141pp) **$14.95**

☐ **RUN, RABBIT, RUN**
Tales of Walter "Rabbit" Maranville (1991, 96pp) **$9.95**

☐ **BASEBALL HISTORICAL REVIEW**
Best of the 1972-74 BRJ (1981) **$6.00**

THE NEGRO LEAGUES BOOK
- ☐ Hardcover (1994, 382pp) **$49.95**
- ☐ Ltd. Edition, leather-bound, slipcase, autographed (1994, 382pp) **$149.95**

MINOR LEAGUE HISTORY JOURNAL
- ☐ **VOLUME 2** (54pp) **$6.00**
- ☐ **VOLUME 3** (72pp) **$7.00**

☐ **BATTING**
F.C. Lane reprint that shares the insights of baseball legends such as Cobb, Ruth, and Hornsby. New biographical foreword and an expanded index add much to the original work (2002, 230pp) **$14.95**

☐ **MEMORIES OF A BALLPLAYER: BILL WERBER AND BASEBALL IN THE 1930S**
Bill Werber is the last man alive who traveled with the '27 Yankees. His colorful anecdotes recall an era long gone (2001, 250pp) **$14.95**

NAME ______________________

ADDRESS ______________________

____

CITY, STATE, ZIP ______________________

DAYTIME PHONE ______________________

SEND YOUR ORDER TO: University of Nebraska Press, 233 North 8th Street, Lincoln, NE 68588-0255 or call 1-800-755-1105 weekdays from 8:00am to 5:00pm CST. You may also place orders online at **http://nebraskapress.unl.edu**
(12/2003)

BOOK TOTAL $ ________

SHIPPING $ ________
(Charges are $4.50 for the first book and 50 cents for each additional book)

NE RESIDENTS ADD SALES TAX $ ________

TOTAL $ ________

MASTERCARD & VISA ACCEPTED

CARD # ______________________

EXP. DATE ______________________