
By the Numbers

Volume 10, Number 1

The Newsletter of the SABR Statistical Analysis Committee

February, 2000

Welcome

Phil Birnbaum, Editor

This Issue

Two of our four articles this issue that are a bit different from the usual.

First, Rob Wood has put together an analysis on Darryl Strawberry's Hall of Fame chances. Typically, BTN has had articles on methods, or formulas, or relationships – that is, general truths about baseball. Rob's article is specifically about Darryl Strawberry. It breaks no new ground theoretically, but is an intelligent and informative look at Strawberry in the context of others like him.

At this point, we certainly have room for both kinds of articles – those on Sabermetrics, and those that apply Sabermetrics to intelligently answer specific questions. I hope you will join me in thanking Rob for his contribution.

Second, we have Tom Ruane's summary of the SABR-L debate on Mike Hoban's "HEQ" statistic. Those of you who subscribe to SABR-L are no doubt familiar with the past year's HEQ discussion. For you, the article will be a synopsis of what you've already read. For those unfamiliar with HEQ and its contentious history, Tom's piece is an excellent summary of what the debate was all about, in many fewer words.

I don't intend to have BTN become a forum for a continuation of the HEQ debate. Rather, we are running this summary as a review of the important discussion for those who have not seen it. BTN routinely runs reviews of work that has appeared elsewhere, as SABR's journal of statistical analysis, this fits right into our purview. But BTN will not be running anything more on HEQ.

Thanks, Tom, for your efforts.

Finally, as usual, I urge you to contact our contributors with questions or comments, or even just to mention that you enjoyed their work. Letters to the Editor are also welcome.

Contributions

A reminder that BTN needs your contributions. For the first time since 1998, I have no articles in reserve for next issue. If you were thinking of writing something up, please be assured that we'd all appreciate it.

In this issue

Welcome	Phil Birnbaum	1
Academic Research: "Stacking"	Charlie Pavitt	3
The Run Value of a Ball and Strike	Phil Birnbaum	4
The HEQ Chronicle	Tom Ruane	7
Squandering Hall of Fame Talent? The Case of Darryl Strawberry	Rob Wood	15

Deadline for the next issue is April 24. Thanks again.

Erratum

The POP formula on page 5 of the last issue of BTN was incorrect. This was my fault, and I apologize to Mike Sluss and those who were inconvenienced. The correct formula appears below.

$$P = \sum_{w=W}^A \frac{A!}{w!(A-w)!} L^w (1-L)^{(A-w)}$$

You can e-mail me at birnbaum@sympatico.ca. Or, you can write me at #608-18 Deerfield Dr., Nepean, Ontario, Canada, K2G 4L1. ♦

Informal Peer Review

The following committee members have volunteered to be contacted by other members for informal peer review of articles.

Please contact any of our volunteers on an as-needed basis - that is, if you want someone to look over your manuscript in advance, these people are willing. Of course, I'll be doing a bit of that too, but, as much as I'd like to, I don't have time to contact every contributor with detailed comments on their work. (I will get back to you on more serious issues, like if I don't understand part of your method or results.)

If you'd like to be added to the list, send your name, e-mail address, and areas of expertise (don't worry if you don't have any - I certainly don't), and you'll see your name in print next issue.

Expertise in "Statistics" below means "real" statistics, as opposed to baseball statistics - confidence intervals, testing, sampling, and so on.

Member	E-mail	Expertise
Jim Box	im.box@duke.edu	Statistics
Keith Carlson	kcarlson@stlnet.com	Economics/Econometrics/Statistics
Rob Fabrizio	rfabrizio@bigfoot.com	Statistics
Larry Grasso	l.grasso@juno.com	Statistics
Tom Hanrahan	HanrahanTJ@navair.navy.mil	Statistics
Keith Karcher	kckarcher@compuserve.com	General
John Matthew	john.matthew@home.com	Apostrophes
Duke Rankin	RankinD@montevallo.edu	Statistics
John Stryker	johns@mcfely.interaccess.com	General
Steve Wang	Steve.C.Wang@williams.edu	Statistics

Receive BTN by E-mail

You can help save SABR some money, and me some time, by receiving your copy of *By the Numbers* by e-mail. BTN is sent in Microsoft Word 97 format; if you don't have Word 97, a free viewer is available at the Microsoft web site (<http://support.microsoft.com/support/kb/articles/Q165/9/08.ASP>).

To get on the electronic subscription list, send me (Phil Birnbaum) an e-mail at birnbaum@sympatico.ca. If you're not sure if you can read Word 97 format, just let me know and I'll send you this issue so you can try

If you don't have e-mail, don't worry—you will always be entitled to receive BTN by mail, as usual. The electronic copy is sent out two business days after the hard copy, to help ensure everyone receives it at about the same time.

E-Mailing the Editor

I have been told that my iname.com and philbirnbaum.com e-mails often don't work. Those go through internet forwarding services before they reach me, and it seems these services are unreliable.

My "real" e-mail address – for now – is birnbaum@sympatico.ca. If that ever fails to work – who knows, I may change service providers again someday – try either phil_birnbaum@iname.com, or BTN@philbirnbaum.com.

Sorry for any inconvenience.

Academic Research: "Stacking"

Charlie Pavitt

The author reviews recent academic research on "stacking," the tendency for positional segregation based on race.

This is the one of what I foresee as occasional reviews of sabermetric articles published in academic journals. It is part of a project of mine to collect and catalog sabermetric research, and I would appreciate learning of and receiving copies of any studies of which I am unaware. Please visit the Statistical Baseball Research Bibliography at www.udel.edu/johnc/faculty/pavitt.html, use it for your research, and let me know what I'm missing.

Benjamin Margolis and Jane Allyn Piliavin, "Stacking" in *Major League Baseball: A Multivariate Analysis*, *Sociology of Sport Journal*, Volume 16, 1999, pages 16-34.

"Stacking" is the technical term that sociologists of sports have given to position segregation by race or ethnicity. Research has given us evidence of stacking in several sports: college and pro football, Australian and English rugby, English soccer, cricket, women's basketball and volleyball, pro hockey, and, of course, baseball. Several explanations have been proposed for the phenomenon, and all probably contain at least a grain of truth. First and foremost, stereotypes of, for example, African-Americans as relatively incapable of "leadership" or "thinking" leads to their underrepresentation in those positions in which such skills are believed to be vital. Second, children admiring role models of the same ethnicity would gravitate to the positions of those role models even if these stereotypes were to disappear; for example, a young black boy seeing Ken Griffey Jr. in center field and envisioning himself as a future center fielder. Third, intentional specialization will occur if young players notice a relevant demand; Dominicans, for instance, believe that scouts search their country for middle infielders and as a consequence prepare themselves for those positions.

The research literature on stacking in baseball has become rather large, although most of it has done little more than replicate past findings in attempts to evaluate the proposed explanations. The first important study (Loy and McElvogue, *International Review of Sport Sociology*, Volume 5, 1970, pages 5-23) found a two-tiered structure (overrepresentation of whites in the "central" positions in the infield, including pitcher and catcher, and blacks in the "peripheral" positions in the outfield). Pattanyak and Leonard (*Sociology and Social Research*, Volume 76, 1991, pages 3-9) found a preponderance of Hispanics in the infield, supporting a three-tiered structure. The present article describes a senior thesis by the first author at the University of Wisconsin under the direction of the second author, a well-respected social psychologist.

Margolis based his analysis on data from the 1992 season from Topps baseball cards and The 1993 Great American Stat Book (I, as a former Project Scoresheeteer, am pleased to see our book so well employed). His study makes two specific contributions to the baseball stacking research literature. First, he distinguished six levels of centrality: catcher, middle infield, the infield corners, center field, and the outfield corners, and designated hitters; thus pitchers, who we know are overwhelmingly white, are ignored in the study. Second, he used multiple regression to examine the joint impact of ethnicity and group differences in variables such as height, weight, age, speed, and the like. For example, if Hispanics are as a group smaller than whites or non-Hispanic blacks, then their preponderance in the middle infield could be accounted for by size rather than the usual explanations for stacking.

The results favor a four-tiered structure, with catchers (and, based on earlier research, pitchers) overwhelmingly white, the middle infield overrepresented by Hispanics, the infield corners somewhat overrepresented by whites, and the outfield and designated hitter overrepresented by non-Hispanic blacks. The regression analysis adds some important insights to these basic findings. First, Hispanics are only shown to be more relegated to peripheral positions after age, height, and weight are controlled for. In other words, Hispanics are discriminated against in position assignment, but the effect of that discrimination is not obvious if one ignores the fact that their average smaller size makes them particularly well suited to the middle infield. Second, the consignment of non-Hispanic blacks to the outfield is shown to be partly, although not completely, the result of better speed as measured by percentage of successful steal attempts. I was reminded of a little study done by Bill James (*Baseball Abstract* 1987, pp. 68-71) suggesting that non-Hispanic blacks as a group maintain their speed, along with other skills, longer than whites and Hispanics.

Margolis and Piliavin's work is the most sophisticated study of stacking in baseball to date, and any further work on the subject must begin where they leave off.

Charlie Pavitt, 812 Carter Road, Rockville, MD, 20852, chazzq@udel.edu ♦

The Run Value of a Ball and Strike

Phil Birnbaum

Recent studies suggest that the catcher can have a substantial impact on his team's ERA. How much does framing pitches have to do with this impact? To answer this question, the author calculates the value, in runs, of a ball and strike, and determines that even the occasional ball turned into a strike can account for much of the catcher's impact.

In the studies on Catcher ERA that I've seen (Craig Wright's first chapter of *The Diamond Appraised*, and Tom Hanrahan's article in the August, 1999 issue of BTN), the authors conclude that the catcher's pitch handling can affect the pitcher's ERA by as much as half a run in either direction. That is, a pitcher with a "natural" ERA of 4.00 can see that ERA rise to as much as 4.50 with a poor catcher, or drop to 3.50 with a very good catcher.

It seems to me that there are three ways a catcher can influence a pitcher's effectiveness: (a) by calling the right pitches for the pitcher's "stuff"; (b) by calling the right pitches to take advantage of the batter's weaknesses; or (c) by framing the pitch so that the umpire is more likely to call it a strike.

Thinking about this, I started to wonder about how much effect (c) could actually have. How many runs are saved if a ball is called a strike? Could the catcher's skill in this regard actually save half a run a game?

Results by Count

To find out, I ran a quick study of play-by-play data in the 1988 American League. Using Project Scoresheet data, I identified every plate appearance in which a particular count occurred. Then, I computed the average linear weight run value based on the outcome of those plate appearances.

For instance, if a 2-0 count occurred 3 times (it actually occurred 9,114 times, but let's keep the example simple), and the eventual results were a single (linear weight .465), double (linear weight .775), and strikeout (linear weight -.273), the value of the 2-0 count would be .322 runs, the average of the three weights.

In reality, the value of the 2-0 count was .083 runs. Here's the entire set of results:

	0 strikes	1 strike	2 strikes	3 strikes
0 balls	.0000	-.0365	-.0874	-.2736
1 ball	.0288	-.0119	-.0680	-.2734
2 balls	.0829	.0290	-.0306	-.2732
3 balls	.1858	.1252	.0578	-.2733
4 balls	.3137	.3137	.3135	

The numbers in the table are denominated in incremental runs – that is, the difference in expected runs vs. starting that plate appearance over at 0-0. Some technical details are in this footnote.¹

A couple of quick notes about the table:

- The "3 strikes" column is -.2733 because that's the value of a strikeout. The "4 balls" row is .3137 because that's the value of a non-intentional walk. The values vary slightly because of errors in the data (for instance, there were 3 instances where the Project Scoresheet file included four strikes).

¹ Technical details: (a) I used linear weight values I calculated myself from the 1988 American League. Using the standard Pete Palmer's weights would make the results vary only slightly. (b) A plate appearance was included in the stat line for any count it passed through. So an at-bat that went Ball/Strike/Strike/Home Run would have the homer included in the 0-0, 1-0, 1-1, and 2-1 stat lines. (c) since linear weights is defined to make the average event zero, the 0-0 count must work out to zero, since every at-bat passes through that count. (d) any situation in which an intentional ball was thrown was ignored.

- The “even” counts 1-1 and 2-2 favor the pitcher. This makes sense, because both counts are closer to a strikeout (which requires only three strikes) than a walk (which requires four balls).
- Unsurprisingly, the best count for a pitcher is 0-2. The surprise is that it takes three of those 0-2 counts to save as many runs as one strikeout. On the other hand, a 3-0 count is more than half as good as a full-fledged walk.

Here’s the data in another form: a batting line for each of the counts. The RC27 column is Runs Created per 27 outs (25.5 batting outs), and the last column is the record of a team who scored that many runs and gave up the league average 4.34.²

	ab	h	2b	3b	hr	bb	k	avg	slg	RC27	Record
0 and 0	77007	19965	3556	425	1901	7190	12322	.259	.391	4.34	81- 81
0 and 1	33524	7780	1295	157	633	1627	8001	.232	.337	2.99	52-110
0 and 2	12302	2189	348	39	161	382	4942	.178	.252	1.58	19-143
0 and 3	2211	0	0	0	0	0	2211	.000	.000	0.00	0-162
1 and 0	31011	8279	1546	194	896	5076	4210	.267	.416	5.36	98- 64
1 and 1	28699	6929	1225	139	687	2581	6177	.241	.366	3.74	69- 93
1 and 2	18874	3570	580	71	298	1050	7131	.189	.275	2.00	28-134
1 and 3	4374	0	0	0	0	0	4371	.000	.000	0.00	0-162
2 and 0	9114	2518	484	75	298	3707	1085	.276	.444	7.59	122- 40
2 and 1	14260	3611	678	75	394	2935	2603	.253	.394	5.12	94- 68
2 and 2	14171	2886	516	59	265	1856	4852	.204	.304	2.88	50-112
2 and 3	3577	2	0	0	0	0	3575	.001	.001	0.00	0-162
3 and 0	1546	405	63	22	55	2630	189	.262	.438	11.00	140- 22
3 and 1	4304	1171	206	44	153	3105	610	.272	.447	9.04	132- 30
3 and 2	6210	1389	256	37	154	2781	1840	.224	.351	5.35	98- 64
3 and 3	1828	0	0	0	0	0	1828	.000	.000	0.00	0-162
4 and 0	0	0	0	0	0	1573	0	.000	.000	0.00	
4 and 1	0	0	0	0	0	2135	0	.000	.000	0.00	
4 and 2	1	0	0	0	0	2742	0	.000	.000	0.00	
4 and 3	0	0	0	0	0	0	0	.000	.000	0.00	

With every batter spotted a 3-0 count, the average team would score 11 runs a game on their way to a 140-22 record. With every pitcher getting ahead of the count 0-2, the team would score only 1.58 runs per game, and lose 143 games per season.

More interesting, of course, are the less extreme cases. If every pitcher threw one extra strike per batter, on the first pitch, its team would win 110 games. And throwing one extra first-pitch ball, falling behind 1-0 on every batter, would reduce an average team to a 64-98 team.

Strike Value

So how much is a strike worth?

The first answer is that it all depends on the count. If a 1-0 count is worth .0288 runs, and a 1-1 count is worth -0.119 runs, the strike has saved the pitcher .0407 runs – about one twenty-fifth of a run.

A strike on 3-2, however, turns a .0578 run situation into a strikeout at -.2733, a difference of a third of a run.

The 3-2 strike is the highest-valued strike in the table, worth about 9 (!) first-pitch strikes. The old saw has that the best pitch in baseball is strike one, but, clearly, the best pitch in baseball is actually strike three.

So not all strikes are created equal, nor are all balls. But we can easily get an *average* value for a strike. Again for the entire league, I calculated the run reduction for all the strikes, and averaged them. (This average is weighted by count frequency, because an 0-0 strike happens much more often than an 0-2 strike.)

The result:

A strike is worth -0.0829 runs.

² Using Pythagorean Projection with exponent 2.

I did the same calculation for balls:

A ball is worth +0.0560 runs.

Surprisingly, or perhaps not, a strike is worth about 1/3 of the value of a strikeout, and a ball is worth close to ¼ the value of a walk (actually, a bit less).

The Catcher

If a catcher is successful in framing a ball so that the umpire calls it a strike, the defensive team saves the difference between the two values:

Turning a ball into a strike saves .1389 runs, or about 1/7 of a run.

So, does it seem reasonable that a difference of half a run per game can be obtained just by framing pitches better? Well, 0.5 runs per game is about four extra balls turned into strikes – four more than an average catcher. Does that seem plausible? I don't really watch enough games, or pay enough attention to pitch framing, to answer that question. But my intuition would suggest that one or two pitches seems reasonable.

One pitch a game better than average is about 20 runs a season. That's two extra games in the win column, or about a third to a half the added value of a superstar free agent. It can be as big a consideration as the catcher's batting line and throwing arm put together.

Phil Birnbaum, 18 Deerfield Dr. #608, Nepean, Ontario, Canada, K2G 4L1, birnbaum@sympatico.ca. ♦

Submissions

Submissions to *By the Numbers* are, of course, encouraged. Articles should be concise (though not necessarily short), and pertain to statistical analysis of baseball. Letters to the Editor, original research, opinions, summaries of existing research, criticism, and reviews of other work (but no death threats, please) are all welcome.

Articles should be submitted in electronic form, either by e-mail or on PC-readable floppy disk. I can read most word processor formats. If you send charts, please send them in word processor form rather than in spreadsheet. Unless you specify otherwise, I may send your work to others for comment (i.e., informal peer review).

I usually edit for spelling and grammar. (But if you want to make my life a bit easier: please, use two spaces after the period in a sentence. Everything else is pretty easy to fix.)

Deadlines: January 24, April 24, July 24, and October 24, for issues of February, May, August, and November, respectively.

I will acknowledge all articles within three days of receipt, and will try, within a reasonable time, to let you know if your submission is accepted.

The HEQ Chronicle

Tom Ruane

In September, 1998, SABR member Mike Hoban introduced a new player rating statistic, HEQ, on SABR-L, the SABR internet mailing list. The result was more than a year of heated debate on the merits of the method. Here, one of the participants summarizes that discussion in the words of the combatants themselves.

On September 30, 1998, in a post titled "HEQ Rating System/1998 season", Mike Hoban introduced HEQ ("a set of simple fan-friendly formulas that establishes how good a season a player had") to SABR-L, the organization's moderated newsgroup. Thus began an often tumultuous 14-month long discussion of his method, a lively (some would argue too lively) debate that showed the workings of SABR-L at both its best and worst. Given the fact that HEQ is now the subject of a book published by McFarland & Company, and that the issue was a primary topic on the newsgroup off and on for more than a year, I thought that a description of what transpired on SABR-L would be of interest to the readers of By The Numbers. As one of the principal combatants in the story, I make no claim to being unbiased. Still, almost all of what follows is part of the public record (being taken directly from the SABR-L posts themselves) and I have taken pains to accurately represent the discussion, both in the quoted material as well as in my summaries of what was written.

Mike's initial post outlined how his method worked. He wrote that an offensive HEQ score of 600 was outstanding, as was a defensive score of 400, and that a player's career score was taken by averaging his ten best seasons together. The offensive formula was described:

$$\text{Total Bases} + \text{Runs} + \text{RBIs} + \text{Stolen Bases} + (\text{Walks} / 2)$$

He wrote:

Since this results in a sum, the number of games played is an important factor in the player's total score. What this means, of course, is that if we had two players of equal ability, then the one who plays 160 games will have a higher score (be more effective) than the one who plays 130 games. This is why the HEQ is superior to any percentage-based statistic such as the batting average or fielding percentage. That is, it automatically takes playing time into account.

While he didn't fully describe how the defensive score was obtained, he did write that it was:

obtained by combining putouts, assists, double plays and errors in a manner that is different for each position. A score of 400 represents a great fielding season regardless of the position played.

Much of the rest of the post consisted of lists showing his method in action as well as his opinion of HEQ's value. It was "free from bias and hype" and "superior to any percentage-based statistic such as the batting average or fielding percentage." HEQ represented "a whole new approach to comparing players - away from the dependence on inadequate percentage-based statistics to a more mathematically sound approach."

He pointed out that HEQ showed Sammy Sosa's 1998 season to be the top offensive season in 60 years, Alex Rodriguez' performance in 1998 to be the best all-around season by a shortstop in history and wrote that "Only a comprehensive comparative system like the HEQ enables us to make such comparisons with the level of confidence that numbers convey."

He didn't have to wait long for a reaction. The next day, I sent a post to SABR-L. I complained that his method assumed that 40% of a player's value, regardless of the position he played, came from his defense, which certainly didn't seem to be the case for first basemen and left-fielders. I quoted his offensive formula and commented:

What? A walk is half as valuable as a stolen base? Stolen bases are counted without regard to caught stealing? Total bases are counted without regard to outs made?

I next showed his method in action by listing a series of paired players, where the bottom player in each group had the higher offensive HEQ score. A few examples:

Name	Team	Year	AB	R	H	2B	3B	HR	RBI	BB	SB	AVG	SLG	OBP
Hank Aaron	ATL	N 1973	392	84	118	12	1	40	96	68	1	.301	.643	.402
Omar Moreno	PIT	N 1980	676	87	168	20	13	2	36	57	96	.249	.325	.306
Mark McGwire	OAK	A 1995	317	75	87	13	0	39	90	88	1	.274	.685	.441
Don Blasingame	STL	N 1957	650	108	176	25	7	8	58	71	21	.271	.368	.343
Ted Williams	BOS	A 1955	320	77	114	21	3	28	83	91	2	.356	.703	.496
Alan Wiggins	SD	N 1984	596	106	154	19	7	3	34	75	70	.258	.329	.342

A few more excerpts from my response:

Another fundamental problem with this method is its heavy reliance upon runs scored and RBIs. This gives a big boost to players fortunate enough to be surrounded by good hitters. It also doesn't take into account park effects.

His decision to evaluate a player's career by ignoring all but their ten best seasons makes a player with fifteen great seasons the equal of one with only ten. His method of counting defense stats without regard to games (or innings) played gives a shortstop with 600 chances in 150 games (who would normally be considered to have poor range) the same defensive score as one with 600 chances in 100 games (or one with great range).

All was quiet for five days. Then on October 6th, Mike Hoban wrote a response. My criticism didn't surprise him. "When someone spends three years building what he knows to be a glass house, he cannot be too surprised when some visitors are tempted to throw stones." He would begin to answer the question I had,

But before doing that, I want to make what I believe to be the most important comment about the HEQ rating system. IT WORKS! That is, it may not be perfect (what is?) but it gives the best picture to date of a player's complete season contribution - hitting, fielding and playing time.

He then proceeded with over 200 lines dedicated to showing how well HEQ worked. All of the 16 top players according to their HEQ career ranking were in the Hall of Fame. Ruth, Mays, Gehrig and Foxx were at the top. Brooks Robinson was the best career defensive third baseman, Rabbit Maranville the top shortstop, and so on. As for the points I'd raised in my note, he wrote:

1. Someone thought that I had assumed that 40% of a player's contribution comes from his defense. This is a misunderstanding. One of the strengths of the system is that no pre-conceived ratios were used - rather a player's offensive to defensive ratio is based on his performance. For example, Lou Gehrig was a much better hitter than fielder - so his career HEQ of $750 + 314 = 1064$ is 70% offensive and 30% defensive. But, Brooks Robinson's career numbers are $462 + 451 = 913$, 51% offense and 49% defense.
2. One of the major goals of the HEQ was to be fan-friendly and to keep things as simple as possible. You can argue weighting systems all day and not get anywhere. The total bases concept gives one point for a single, two for a double, etc. A walk gets half a point compared to one point for a single. The three other components: runs, RBIs and stolen bases each get one point. Data like caught stealing cannot be used because they were not kept uniformly throughout the 20th century - one of the conditions needed to compare all players.
3. One last note at this time. If you compare Mark McGwire's 1995 season with that of Don Blasingame's 1957 season (as one respondent did), you find that Blasingame had an offensive HEQ score of 462 while McGwire had 427 (neither of which is very effective). But Blasingame had the more effective (productive) season. Why would this surprise anyone when you consider that Blasingame batted 650 times compared to only 317 for McGwire (over twice as many at-bats)?

We are dealing with total production - not "per-at-bat." Remember that we are not talking about who had the "better season." I might be inclined to say that Mac had the better season since he achieved an almost equal HEQ score with only half as many at-bats. That's what BA, OBP, SLG, etc. are all about - percentages. As a mathematician, I cannot use undefined terms and "better season" is an undefined term. That is why I have defined "offensive effectiveness" (offensive HEQ) precisely. I cannot say with any precision who had the "better season." But I can say that Blasingame had the more effective season.

This would be a source of confusion for many on the list over the next fourteen months: what exactly did HEQ measure? Sometimes, Mike would use "effective" and words like "better", "outstanding" or "great" interchangeably; on other occasions (like the quote above) he would seem to be saying that all HEQ measures is HEQ.

Chip Hart was the first to respond. He felt that Mike had not actually addressed any of the points I had raised. "Why on earth are SBs worth 2x walks? How in the world do we get 30-40% of a player's worth from defense? Etc." He was also unimpressed by Mike's claim that HEQ works. "Your logic is flawed here. Just because this tool brings up names we recognize doesn't mean it's effective."

Mike Hoban posted again on October 7th, discussing the problem he saw with rate statistics in general and batting average in particular. Batting average, according to Mike:

... is a helpful tool to give us some idea of how a player is hitting as the season progresses. However, once the season is over, the BA has served its purpose and should be put to rest in favor of a statistic that measures the actual accomplishments of the player for the season. To a mathematician, this fact seems very apparent but, evidently, it does not seem that obvious to others. When the BA is used after the season is over, it becomes a potentially misleading guide.

He followed that with three examples. He showed that Mike Schmidt had a higher offensive HEQ than George Brett in 1980, despite a much lower batting average, that Yogi Berra had a higher career HEQ score than Bill Dickey, and that Shoeless Joe Jackson "does not have Hall of Fame numbers - and, in no way (other than myth), does he belong in the company of the greatest players", despite having one of the highest batting averages in history.

He continued:

Any attempt to create a post-season comparative measure for baseball players that is to be fair must take at least three elements into account: hitting, fielding and playing time. Such a measure must exclude all in-season percentage-based references such as the BA. It must add together in a thoughtful and valid manner the actual performance numbers that a player produced during the season. It must be completely objective - free from bias or hype. But is such a measure possible? And, if so, why do we not have one in common use at the present time?

One problem as he saw it was that the creation of such a statistic:

would take an enormous amount of time researching the numbers in the record books and coming up with a set of formulas that yield results that are both accurate and intuitively pleasing. But that is exactly what I have done with the HEQ rating system.

And he finished with the question:

Is baseball ready for an understandable, accurate post-season comparative rating system? Who knows - only time will tell.

I posted a response later the same day. Two excerpts:

He seems very impressed that his system had Babe Ruth and Willie Mays at the top of the career list. I'm more interested in the fact that it placed Charlie Gehringer far ahead of Ty Cobb and Ted Williams, and Earl Averill ahead of Mickey Mantle. I'm sure any of us could come up with a system in a couple of minutes that'll put the Bambino on top of the heap. How about slugging percentage? OPS? The hard part is getting Earl Averill ahead of Mickey Mantle.

But what about defense? Who had the best defensive seasons in 1998? Well, he didn't give us a complete listing, but among the top hitters, the choice was clear (as well as one I'm sure many of us would have come up with ourselves if given enough time). I'm talking, of course, about Rafael Palmeiro and Jeff Bagwell. With 463 and 462 points, Palmeiro and Bagwell tower over such lesser defensive players as Nomar Garciaparra (372) and Barry Bonds (295). Much like the way he handled the failures of HEQ in evaluating offense, Hoban reports these findings as if they are revelations rather than simple mistakes.

I mentioned his assertion that Blasingame had been more effective than Mark McGwire in the seasons I had listed and asked:

Does he really believe this? In 1957, Blasingame got on-base 62 more times than McGwire and had 22 more total bases. He also made 207 more outs. Doesn't that seem like an awful lot of outs to give up for those extra singles? I think most people would consider 39 home runs and 88 walks in 317 at-bats extremely effective. I guess one of those people is not Mike Hoban.

John Stuart wrote that he didn't think that HEQ was an improvement over an even simpler formula, runs scored plus RBIs divided by two. James Tuttle thought that "it doesn't mean much if you post a list of famous players and assert they're the best because they're at the top of your list. You need to validate your method first."

The next day, Mark Armour tied the HEQ discussion into an earlier debate on SABR-L concerning the value of an average player. That debate had started when I pointed out that Pete Palmer's Total Player Rating produces a number which approximates the number of victories a player was responsible for above what an average player would have produced at his position. I (and others) argued that for some types of analyses (evaluations of trades, for example) this is not very helpful, since an average player ends up with the same value (zero) as a player who never played in the majors at all. I suggested that for these types of discussions we need to compare players, not against average major leaguers, but against replacement players who would be available in the high minor leagues. Mark can take it from there:

One side (Total Baseball and others) measures all players as a difference from the average player at his position. What you measure is how much a player brings a team up from a .500 team.

The other side (my side) measures all players from a theoretical replacement player who is less than the average player. A team of replacement players might win a third of their games.

A new, third, position is that espoused by Hoban in his book. He believes (apparently) we should measure from zero. Omar Moreno (playing every day) is better than (an occasionally injured) Mark McGwire if you assume that McGwire would be occasionally replaced by a player who struck out every single time he batted, and let every ball drop next to him or go through his legs.

On October 8th, Mike sent another post to SABR-L. He admitted that HEQ was "a work in progress" and that he was looking for feedback.

Unfortunately, it appears that I have begun this attempt in an inappropriate way. I have evidently posted an inadequate amount of information in a manner similar to waving a red flag at a bull. Since my manuscript for the book runs to 220 pages and a good part of that establishes the rationale for the HEQ, I realize that posting a few lists of results and some very limited explanation has had the effect of shedding more heat than light.

He then proposed to mail sections of his manuscript to anyone who was interested. In addition, his post contained the complete defensive formulas used in HEQ:

$$\begin{aligned} OF &= PO + 4A + 4DP - 2E \\ 2B &= (PO + A + DP - 2E) \times .460 \\ 3B &= (PO + A + DP - 2E) \times .888 \\ SS &= (PO + A + DP - 2E) \times .548 \\ 1B &= (2A + DP - 2E) \times 1.29 \\ C &= (PO + 3A + 2DP - 2E) \times .445 \end{aligned}$$

He admitted that the 1B and C formulas weren't working too well yet and stated that the "multiplication factors derive from a study of Gold Glove winners and those fielders who had a reputation for excellence and the establishment of an 'outstanding season score' for each position."

Finally, he attempted to address a few of our concerns. Why use only a player's ten best seasons in deriving career scores? Two reasons: that's what the Hall of Fame uses for eligibility, and the fact that we need to level the playing field to include players who had shorter careers than others. As for Averill being a greater player than Mantle, he responded:

Sure, why not? All this really says is that if you look at the ten best seasons each had, Mantle was a slightly better hitter and Averill had better fielding numbers. Both of these men were great players. But keep in mind that much of Mantle's place in the fan's perception is influenced by the hype of the NY media. That's one reason why an objective comparative tool like the HEQ is helpful.

Jim Furtado joined the discussion, arguing that HEQ was hardly an objective method. "You decided who the best defensive players were. Then, you developed a method which produces results which agree with your own subjective ideas. How is that objective?"

And I wrote about his assertion that "effective" is only defined in terms of HEQ and does not necessarily mean that someone was "better" than someone else:

While I must applaud the modesty in this approach, I think it severely limits HEQ's usefulness. Most of us expect these kinds of measurement systems to give us some indication of a player's relative worth. Since HEQ doesn't pretend to do this, I'm not sure why any of us should care what scores it comes up with.

Bob Allen preferred to discuss HEQ's defensive formulas.

Now that we have the formulae for appraising defensive contributions according to Dr. Hoban's HEQ system, it is possible to see how glaring the weaknesses are.

If you apply the formulae for each position to the league averages for 1997, you get the following results (per 162 games):

	AL	NL
1B	430	456
2B	398	407
3B	369	362
SS	436	435
OF	397	370
C	580	635

Accordingly, an average first baseman who plays the whole schedule makes a defensive contribution equal to or better than an average performer at 2B, 3B or SS. Is this logical?

He concludes that "it is the logic that bothers me. A first baseman wins and keeps his job by hitting, not with his glove, and the opposite is true for a shortstop. To suggest that their defensive contributions are in any way equal is absurd."

Mike Hoban responded to a few of the many points being raised. On defense:

But the essential question is how much should defense count when compared to offense if we are trying to identify the "most complete" players (offense and defense) to have ever played the game - which is what the HEQ is trying to do. Of course, I do not know the answer to this question - and neither does anyone else. Does it seem reasonable to say that Lou Gehrig's contribution to his team was in the ratio of 70% offense to 30% defense (as the HEQ indicates)? Sure. Could some other ratio seem reasonable? Quite likely.

On validating his approach:

Validate or prove any rating system? There is no empirical way to validate or prove any system - whether it is the TPR or the HEQ.

On adjusting for the era or park that each player played in:

Playing in different eras and in different parks (and other factors) certainly affect a player's career numbers - even though these factors cannot be quantified and "adjusted for."

He then argued that a value of HEQ was that it enabled him to make statements like: "In 1998, Alex Rodriguez had the most effective all-around season for a shortstop in this century: $700 + 423 = 1123$." And followed that with a list of the most effective shortstops of the century.

That post brought more people into the discussion. Gary Collard argued that measurement systems are validated all the time by testing their correlations with runs scored or winning percentage. He also stated that "a large body of work out there" contradicted Mike's statement that park and era affects can not be adjusted for. Chip Hart made similar points in a post of his own, while F.X. Flinn had this analysis of the defensive component of HEQ:

Reading between the lines, I suspect the bias in HEQ is that its core starting point was to make the legendary outfielders of the game earn around 1000 points. Once that system was in place, adjustments were added to bring other players in line with that target. So there is no on-the-field, game-play foundation for the "value" of the derived points.

John Jarvis talked about how some other systems were validated, referring to a paper he had written on the subject and concluding:

At the level of understanding now achieved in the player rating game a new method that does not explicitly establish its accuracy by comparison to the statistical record probably should not be taken seriously.

I wrote:

We've talked before about how the basic weights Hoban assigns to the various offensive events don't make much sense. But what if he had ignored runs scored and RBIs? What if the weights for stolen bases, hits, walks, caught stealing and outs had been more in line with current research? And what if he had taken into account both the era and the park the hitter played in? Well, in that case, he'd have reinvented something close to Pete Palmer's Batting Runs. In other words, there's nothing really new in the offensive portion of HEQ; the only thing unusual about his method is how simple and inaccurate it is.

As for the defensive portion, I pointed that the "most effective" defensive players in the NL from 1990 to 1998 consisted of five first basemen, three catchers and an outfielder. "Have you ever heard of a good-field/no-hit first baseman?"

The most "effective" defensive players of all-time? In the AL, the king was Bill Buckner of the 1985 Red Sox. He was 35 that year and mighty slow, precisely the type of guy I'd figure to be the top defensive player the AL has ever produced.

I continued:

My feeling is that there's got to be a good reason to come up with yet another measurement system. One reason would be if that system gave us some insight into a player's contribution not available through other means. Usually, the way I look for these is by examining the cases where the system disagrees with others and seeing if the new method gives us cause to reevaluate the conventionally held belief. Well I've looked at a lot of instances where HEQ produces wildly different results than other approaches, and in EVERY case the reason for the difference was some obvious failure in HEQ.

Right around this time, Jeff Nibert was wondering whether posting to SABR-L was the "most effective way to do peer review for one's research?" Mike Hoban responded:

I think it is fair to say that if one's research suggests an alternative way to look at the performances of the great players and, in so doing, questions some of the theories put forth over the past twenty-five years and requires a book of better than 200 pages to explain fully - then the answer is that SABR-L is probably not the appropriate place to present such research. Essentially, it is not fair to ask people to react to pieces of information when they have not had the opportunity to examine the "big picture."

It has become obvious to me that meaningful discussion of some of the important concepts that the HEQ system is questioning is not possible until interested people have had the opportunity to examine the book in its entirety. So, I think it is best to continue our discussion after the book is available.

And so ended the first phase of the HEQ discussion. The book was originally due to come out that December but was delayed. It still wasn't available when Mike Hoban's article on Joe Jackson appeared in The Baseball Research Journal around the end of 1998. Adam Katz and Rob Wood, newcomers to SABR-L, both wrote to the list about it. Rob Wood was very critical of the article, finding it "fundamentally flawed" (among other things), while Adam Katz didn't want to comment on Mike's method "because, frankly, I don't know much about it."

This prompted Mike to reintroduce HEQ to the list a short time later:

Adam Katz, in commenting on my piece in the SABR Research Journal regarding the Hall of Fame numbers of Shoeless Joe Jackson, seemed to suggest that he would like to know more about the HEQ rating system.

I thought that some members might be interested in seeing a rating of 112 players based on a three-year study by a mathematics professor who has been a student of the game for more than fifty years, although a member of SABR for only two. The HEQ rating system is one fan's attempt to create a relatively simple and understandable total-season statistic to compare the season production of position players. Is it perfect? Of course not, what is? But it is a serious effort by an experienced researcher to return the enjoyment of baseball numbers to the average fan for whom concepts such as Runs Created and TPR are driven by formulas that are just too complicated. Is the HEQ research going to have an impact on how some fans view the players? I cannot be sure, but I think it will.

Two of the defensive formulas had changed:

$$1B = (.25PO + 3A + DP - 2E) \times .51$$
$$C = (PO + 3A + 2DP - 2E) \times .445 \text{ with PO not to exceed } 800$$

But everything else was unchanged by last fall's discussion. As a result, there were a flurry of posts by F.X. Flinn, John Pastier, Chip Hart, Scott Fischthal and myself basically rehashing the things said the previous October. Richard Zitrin probably spoke for many of the list subscribers when he wrote:

There have been numerous long posts about HEQ, as there were last fall. I'd love to see us put this issue to bed before we all fall asleep. Enough is enough, and all that can be said has been said.

Mike Hoban continued to send posts concerning HEQ to the list at regular intervals throughout the year, but little new was said about it until a post of his on September 8th reasserted his claim that it was a bad idea to adjust for a player's era and park.

Clearly, the era and the park in which a player performed had an effect on his performance. That is not the issue. The hocus-pocus begins when someone suggests that through some sort of mathematical manipulation we can, in a fair and unbiased manner, CHANGE A PLAYER'S ACTUAL PERFORMANCE NUMBERS to reflect the era and the park in order to COMPARE that player with another. That notion is pure fantasy.

In the discussion that followed, Mike Hoban expanded upon this notion, describing the difference between "packaging" and "projecting" numbers.

To package numbers means to take the player's actual performance numbers and to put them together in an attempt to make them more meaningful. Batting average, fielding percentage and on-base percentage are simple examples of this.... The HEQ system is an example of this sort of packaging which includes many offensive and defensive numbers. Some people may object to the packaging of actual baseball numbers for one reason or another but it is clear that the outcome is based on the player's actual numbers.

To project numbers in baseball research means to change existing numbers and/or create new numbers in an effort to make the player's actual numbers more meaningful. Exercises such as adjusting for era, adjusting for park effect or projecting runs created are examples of this endeavor. Many researchers find these sorts of exercises helpful for various reasons.

The essential distinction, and my point, is that packaged numbers are based completely on actual numbers and the result is a 'real' representation of what a player did. Whereas, projected numbers are based on created numbers and are a 'fantasy' representation of what the player might have done.

This elicited the following response from Clay Davenport:

And yet these "fantasies" of Michael's are also based on real numbers – it's just that instead of looking only at the player, and his real numbers, the analyst is extending his gaze to include the equally real numbers of what other people were doing at the same time. Why are they less "real" than those of a single, isolated player?

Why, Mike, is a limited reality preferable to a more inclusive one?

A fundamental tool, in all fields, is to normalize values from disparate backgrounds into a standardized setting: "all figures translated to 1990 dollars", the entire metric system, every calendar system.

If you get a reasonable assessment of what a player was worth, relative to his peers; and you convert that into a line that has equal value, but relative to a different baseline; the resulting numbers may well be a "projection fantasy", but the *value* of those numbers will be unchanged from the original, and are far easier to compare to other seasons than the numbers alone.

And it is the value of the numbers, not the numbers themselves, that are important.

In November, Mike posted a list of his of "the top players of the 90s according to the HEQ system (based on their five best offensive and defensive seasons)." Once again, there was a heated reaction to this and his previous posts. Old hands at the discussion like Chip Hart, John Pastier and F.X. Flinn joined with newer contributors like James Fraser, Stephen Lyman, Jay Roberts and even Richard Zitrin (who had pleaded to stop the HEQ discussion earlier in the year) to note problems they saw in the method. Most of what was being said was no

longer new and several subscribers to SABR-L wrote to Ted Turocy, the list moderator, asking for him to put a stop to the discussion. On November 18th, 1999, Ted did just that, imposing a moratorium on HEQ on SABR-L until the book came out in early December. As I write this, the book has been out for a couple of weeks and the moratorium remains in effect.

At the top of this article, I wrote that the HEQ experience showed SABR-L at its best and worst. I'd argue that the discussion represented above, especially the time and energy that a wide range of people put into their review of Mike's work, shows how SABR-L can be an extremely effective way of getting a peer review of one's research. And how did this show SABR-L at its worst? Well, I think that the discussion could have been more civil. Part of this may have been due to Mike's manner; many felt that much of what he wrote read more like ad-copy than research. There was also a certain lack of humility and dismissive attitude on both sides of the debate. I don't believe that this was a discussion that aged particularly well either. As time went on (and Mike introduced the subject again and again), several people including myself, became testy and irritable. We sniped and were petty. Worst of all, we got boring and repetitious. We made points that no longer needed making. In retrospect, I think we all would have been better off had we trusted the judgements of SABR-L subscribers more and let much of what was submitted pass without comment.

In the final analysis, what distinguished the HEQ discussion from so many others on SABR-L was the one-sided nature of the debate. No one except for Mike Hoban defended the method, while literally dozens of subscribers attacked it. Why? Mike would no doubt argue it was because of the revolutionary nature of his work. In a *By The Numbers* piece last year he compared himself to Martin Luther, and I think he felt that many of the members on the list had been too brainwashed by the work done by Pete Palmer, Bill James and others to give his startling new approach a fair reading. He seemed confident that this would change once the book came out and serious baseball researchers had a chance to read a full-length treatment of his ideas. Well, the book has now been published and, to quote Mike, only time will tell.

Tom Ruane, 92 College Ave., Poughkeepsie, NY, 12603, truane@bestweb.net. ♦

Mike Hoban Responds

Much of what is printed in this article reflects the understanding of the individual at the time he posted the comments (some of which may not have been accurate). Two examples of this follow.

FX Flinn said "Reading between the lines, I suspect ... [p. 10]" This is, of course, pure speculation and has no basis in reality.

Tom Ruane said "The most effective defensive players of all time... [p. 11]" This simply does not reflect what the HEQ says. Bill Buckner has never been anywhere near the better defensive first basemen of all time. The HEQ has always shown that Keith Hernandez was the best.

It is important to note that much has changed in the HEQ since some of these comments were made. The HEQ is alive and well. I am willing to send up-to-date information on the HEQ to any member who requests it. Just e-mail me at mhoban@monmouth.edu.

Mike Hoban

Squandering Hall of Fame Talent? The Case of Darryl Strawberry

Rob Wood

Darryl Strawberry was once considered a serious Hall of Fame candidate. Is that still true? Or has his career degenerated to the point where he no longer has a chance? Here, the author takes a stab at these questions, and also investigates other players whose later seasons failed to match the promise of their early career.

Introduction

During last year's World Series, a media pundit observed that these could well be the final games in Darryl Strawberry's rollercoaster career. More to the point, he argued, Strawberry was at one time a great player surely destined for the Hall of Fame. Now, he concluded, due to drugs, injuries, and cancer, Strawberry has become a bench player who nobody would even consider for the Hall. He wondered how unique Strawberry's precipitous fall was in the annals of baseball history.

In this article I want to explore three aspects of the pundit's comments. First, what are Strawberry's Hall of Fame chances now? Second, was Strawberry really at one time a sure-fire Hall of Famer-to-be? Third, how unique is Darryl's fall from grace?

Strawberry Now

I think it can be safely argued that Strawberry is not a serious Hall of Fame candidate based solely upon the statistics he has accumulated to date. After the 1999 season, Darryl has 1401 hits, 335 home runs, 221 stolen bases, 898 runs, 1000 RBI, a .359 on-base percentage (OBP), and a .505 slugging percentage (SLG).

According to the Hall of Fame monitoring system developed by Bill James, Strawberry has 49 points. James designed his system so that a fully deserving Hall of Famer has 130 or more points; a player deserving of no Hall of Fame consideration has fewer than 70 points; and players with 70-130 points are in the "gray area" (100 points was designed to be the 50/50 demarcation). Thus, we see that James's system predicts that Strawberry is well short of any Hall of Fame consideration.

Suffice it to say, in my opinion a great deal of "sympathy" would have to be given to Strawberry for his injuries (and other afflictions) before a Hall of Fame case could be made. I just don't see that happening. So I agree with the pundit that Strawberry will not make the HOF.

Strawberry Then

Strawberry's last good season before his troubles was 1991, his first year with the Dodgers after eight seasons with the Mets. At the end of the 1991 season, Darryl was 29 years old (he was born in March 1962), and he had career totals of 1159 hits, 280 home runs, 201 stolen bases, 748 runs scored, and 832 RBI, with a .359 OPB and a .516 SLG.

It could be argued that these stats, impressive as they may be, do not do justice to how great a player Strawberry actually was. First, Darryl was an all-around player; no single stat can capture his contributions. Second, offense in the 1980's was somewhat moderate, nothing like the offensive explosion witnessed in the 1990's (especially the late 1990's). Third, Strawberry played his home games in Shea Stadium and Dodger Stadium. Although Shea is nowhere near as favorable to pitchers as Dodger Stadium is, Mets hitters were nonetheless at a slight disadvantage.

One way to attempt to take into account these three points is to review Strawberry's "Total Player Rating" (TPR) from Total Baseball. At the end of the 1991 season, Strawberry's TPR was 28.0, meaning that Darryl's contributions led to 28 additional team wins over his career to that point, over and above what a league-average right fielder would have contributed.

Historical Counterparts

How does this rate? Does a TPR of 28.0 after age 29 indicate a sure-thing Hall of Famer? Table 1 presents the top 10 TPR among 20th century hitters after their age-29 season. (Following custom, a player's playing age was deemed to be his age as of July 1 of that year, roughly the midpoint of the season.)

Of course, these players are among baseball's all-time greats. The fact that Darryl Strawberry's TPR total through age 29 does not put him among this elite company is no criticism.

But Strawberry's TPR total of 28.0 through age 29 is still noteworthy. In the entire 20th century, only 32 hitters had accumulated a higher TPR at that point in their careers, and 29 of them were ultimately elected to the Hall of Fame.

Let's investigate how players' TPR through age 29 correlate with their eventual Hall of Fame status. Table 2 presents that data.

Table 1: Top 10 TPR through age 29 (among 20th century hitters)

	TPR through age 29	Final Career TPR
Babe Ruth	66.5	124.8
Rogers Hornsby	60.7	81.2
Mickey Mantle	58.1	76.1
Ty Cobb	56.4	91.0
Tris Speaker	51.3	86.5
Barry Bonds*	49.5	79 +
Willie Mays	46.1	92.2
Hank Aaron	45.7	89.8
Rickey Henderson*	45.7	80 +
Ted Williams	45.2	85.7

* active

The table should be easy to read. The first row indicates that 12 hitters in the 20th century had TPR exceeding 40 after their age-29 season, and all 12 were ultimately elected to the Hall of Fame.

Generally, TPR through age 29 is a good predictor of Hall of Fame membership. Accumulating a TPR of 40 or more through age 29 has thus far guaranteed future Hall of Fame election. Of course, only 12 of the all-time greats have managed this feat. In addition to Ruth, Hornsby, Mantle, Cobb, Speaker, Mays, Aaron, and Williams from Table 1, this list includes Eddie Collins, Mel Ott, Jimmy Foxx, and Eddie Mathews.

You will note from Table 2 that one player accumulated a TPR in excess of 35 through age 29 but is not a Hall of Fame member. In addition to this player, there are 2 other non-HOFers who had 30-35 TPR through age 29, and still another who had 25-30 TPR. I will leave it to the reader to think who these players may be. (Hint: all four played since World War II; their identities will be revealed below.)

Table 2: TPR through age 29 vs. Hall of Fame (among 20th century hitters)

TPR through age 29	Hall of Fame Members	Total # Eligible for HOF	Percent
40 +	12	12	100%
35 - 40	6	7	86%
30 - 35	7	9	78%
25 - 30	12	13	92%
20 - 25	15	32	47%

So there is a general tendency toward making the HOF after compiling a hefty TPR total through age 29. But what is Darryl Strawberry's cohort? His TPR through age 29 of 28.0 would lead me to believe that he is not truly comparable to Ruth, Hornsby, and Mantle, but to the players with TPRs in the 20's.

Let's start from the bottom. What about the 20-25 TPR'ers? This seems to be the fulcrum from which about half the players go on in their thirties to fulfil their Hall of Fame potential, and the other half kind of peter out and don't quite live up to Hall of Fame expectations. Table 3 lists the 15 players who had 20-25 TPR and eventually made the Hall of Fame, along with their relevant TPR totals.

For comparison, Table 4 lists the 17 players who had 20-25 TPR and did *not* make the Hall of Fame, along with their relevant TPR totals.

By looking at Tables 3 and 4, you can see that there really isn't all that much difference between the quality of the players. Sure the players in Table 3 generally performed better in their 30's and are now deemed better than the players in Table 4, but it isn't all that obvious. The salient point is that, through age 29, a TPR of 20-25 seems to represent the Hall of Fame "gray area".

Now let's discuss the players who had even higher TPR totals through age 29. I've kept you in suspense long enough. Who are those players who started out like gangbusters, accumulating sure-fire Hall of Fame TPR totals, only to be deemed not worthy of Cooperstown when they retired? What happened to them? Did they fall apart in their thirties? Did injuries (or other afflictions) curtail their careers? Does TPR greatly exaggerate their values relative to other systems or experts? I would answer yes to all of these questions.

As some of you may have guessed, Ron Santo is the player who accumulated the most TPR through age 29 (36.2) and is not in the Hall of Fame. His final career TPR of 39.5 reflects his relative lack of productivity in his final seasons, and the fact that he retired "young" (at age 34). I won't rehash the Hall of Fame debate surrounding Santo. Personally, I am not displeased that Santo is not in the HOF.

Next we have Bill Mazeroski (30.4, 36.3) and Cesar Cedeno (30.4, 28.3). Many argue that Maz's whopping TPR totals are greatly exaggerated by an overvaluation of turning the double play. Personally, I have never considered Mazeroski to be a strong Hall of Fame candidate. Cedeno, as you remember, was seriously hurt in 1978 at age 27, effectively ending the productive portion of his career, even though he played irregularly for eight more seasons. Thus Cesar was not able to rack up impressive career totals, and was never considered to be a serious Hall of Fame candidate late in his career. In his one time on the Hall of Fame ballot, he received only two votes.

Finally, Dick Allen (27.1, 35.3) is the only player with 25-30 TPR through age 29 who was not voted into the Hall of Fame. Allen's strange and twisted career is well-known. He has not received serious Hall of Fame support during his years on the ballot. This is probably appropriate.

Table 3: Players with TPR 20-25 through age 29 who are in the HOF (among 20th century hitters)

	TPR through age 29	Final Career TPR
Richie Ashburn	23.7	30.2
George Brett	23.9	39.0
Mickey Cochrane	23.0	31.4
Sam Crawford	24.5	30.3
Joe Cronin	23.9	39.4
Hank Greenberg	23.6	30.0
Travis Jackson	24.6	24.0
Chuck Klein	24.1	21.4
Rabbit Maranville	23.4	12.5
Johnny Mize	21.5	36.2
Joe Morgan	24.4	56.3
Al Simmons	23.1	25.6
Ozzie Smith*	21.1	43.0
Duke Snider	23.1	22.3
Paul Waner	21.1	36.1
Average	23.3	31.8

* sure Hall of Famer to be

Table 4: Players with TPR 20-25 through age 29 who are not in the HOF (among 20th century hitters)

	TPR through age 29	Final Career TPR
Dick Bartell	21.1	27.9
Bobby Bonds	23.1	29.5
Gary Carter	20.1	29.6
Jack Clark	20.7	31.1
Harlond Clift	20.4	22.7
Rocky Colavito	20.8	25.2
Jim Fregosi	23.4	20.6
Joe Gordon	23.2	28.2
Bobby Grich	21.8	44.0
Heinie Groh	24.3	27.2
Keith Hernandez	21.0	34.4
Charlie Keller	23.0	23.5
Sherry Magee	22.9	26.5
Tony Oliva	23.4	24.9
Dave Parker	21.9	23.0
Reggie Smith	20.5	35.0
Jimmy Wynn	22.8	33.2
Average	22.0	28.6

Concluding Remarks

It can be safely argued that Strawberry is not going to make the Hall of Fame. And, before the onset of his troubles, Darryl was a virtual lock for the Hall of Fame. However, his fall from grace is not unique. I would argue that there are a few historical counterparts who were similar to Strawberry. Among those players who had great early careers, yet never lived up to their Hall of Fame talent, I would cite Dick Allen, Bobby Bonds, and Dave Parker. Each has a basketful of reasons why it is a shame that they are not in the Hall of Fame. In conclusion, the media pundit was right to include Strawberry on this short list.

Rob Wood, 2101 California St. #224, Mountain View, CA, 94040-1686, rob.wood@us.pwcglobal.com. ♦

Book Reviews Wanted

Every year, a number of books and magazines are published with a Sabermetric slant. Many of our members have never heard of them. Our committee members would like very much to hear when this kind of stuff comes out.

If you own a copy of any baseball book of interest, we'd welcome a summary or a full-length review. The only restriction, please: the book should have, or claim to have, some Sabermetric content.

See John Matthew's review in the previous BTN for the kind of thing we're looking for.

Send reviews to the usual place (see "Submissions" elsewhere in this issue). Drop me a line if you want to make sure no other member is reviewing the same publication, although multiple reviews of the same book are welcome, particularly for major works. Let me know which book you're doing, so I don't assign the same book twice.

And if you're an author, and you'd like to offer a review copy, let me know – I'll find you a willing reviewer. Ron Shandler's "Baseball Forecaster" appeared in BTN based on a review copy, and your book could be next.

Convention Presentations Wanted

Doug Lehman

SABR members are invited to submit presentations to be given at SABR 30 in June. Research presentations should be 20 minutes in length, with a 5-minute question-and-answer period to follow. The presentations typically tend to be statistical or historical in nature.

Send a 250-word abstract of your presentation, along with biographical and address information, to:

Doug Lehman
11271 N. Kendall Dr. Apt. J109
Miami, Florida
33176

or e-mail dlehman@mdcc.edu.

The deadline for submissions is April 30, 2000.