

BASEBALL ANALYST

JOURNAL OF SABERMETRICS
June, 1987 Vol. 30

We have seven articles for y'all this time. I like that, like having a bunch of little pieces rather than three or four longer articles. The seven articles are:

Improving the Runs Created Formula

by David H. Robinson
Page 2

Miracle Teams and Long Shots: What are the Odds?

by Sandy Sillman
Page 7

Some Comments on Charles Pavitt's "Bias In Fielding Evaluation 1: Pitcher Handedness"

by Dallas Adams
Page 11

All-Time Greats Consensus

by Daniel Greenia
Page 12

The Implications of Leadership Research on Baseball:
2. Current Theory

by Charles Pavitt
Page 13

On the Inaccuracy of the Pythagorean Equation at Extreme Scoring Ratios

by Dallas Adams
Page 17

Strikeout to Walk Ratios and Winning Records

by Russ Eagle
page 19

Another thing I like about this issue is that several articles included accomplish something that was one of the initial goals of this project, which was to provide a place for an on-going discussion of sabermetric issues. Dallas Adams' first article responds to a comment six issues ago by Charles Pavitt. Russ Eagle's research responds to a question I posed in the last issue. Charles Pavitt's article extends his own work begun in issue 27. David Robinson's leadoff piece responds to prior sabermetric research published in other places. That's what we're about. Thanks.

Bill James

IMPROVING THE RUNS CREATED FORMULA

by David H. Robinson

Bill James' Runs Created formula has fascinated me ever since I first picked up a copy of the Abstract. Why does it work? What is it saying about other baseball statistics? How do we use it? Well, I'm not going to attempt an explanation of all these issues here. Like other statisticians, when given an estimator of any kind, one of the foremost questions in my mind is "How can I improve it?"; that is, how can I tweek the Runs Created formula to make it come even closer to the actual number of runs that a team scores? That's what I'll attempt to show in this article.

The simplest formula, and the only one I've ever committed to memory, is presented by James as:

$$RC = (H + W) * (TB) / (AB + W)$$

This formula consists of three components:

H + W = number of times reached base from hits or walks

TB = total bases

AB + W = number of plate appearances (almost)

With a slight amount of inaccuracy, we can also write the formula as:

$$RC = AB * SLA * OBA,$$

where OBA = on-base average = $(H + W)/(AB + W)$ approximately,
and SLA = slugging average = TB/AB .

This form for RC is very appealing to me, as it combines SLA and OBA in a meaningful, but simple way. But this simple form for RC is not perfect, as James has noted in presenting two more technical versions for calculating RC. The two technical versions are somewhat more accurate than the simpler RC, but lose some appeal in the increased number of statistics necessary to calculate them. In addition, Paul Johnson, in the 1985 Abstract, has presented his own formula for RC, also with three versions ranging from simple to technical, though none is as simple as James' original RC formula. Johnson's versions appear to be slight improvements over James' RC, but they too require more complete statistical data. In case you don't happen to have the Abstract handy as you read this, the formulas for the three James estimators and three Johnson estimators are given below:

$$BJ1 = \frac{(H + W) * (TB)}{AB + W}$$

$$BJ2 = \frac{(H + W - CS) * (TB + .55*SB)}{AB + W}$$

$$BJ3 = \frac{(H + W + HP - CS - GIDP) * (TB + .26*(W - IW + HP) + .52*(SH + SF + SB))}{AB + W + HP + SH + SF}$$

$$PJ1 = .16*(2*(TB + W) + H + SB - .615*(AB - H))$$

$$PJ2 = .16*(2*(TB + W) + H + SB - .610*(AB + SB/4 - H))$$

$$PJ3 = .16*(2*(TB + W + HP) + H + SB - .605*(AB + CS + GIDP - H))$$

I began my research on RC by taking the team batting data from the Elias Baseball Analyst for the years 1984, 1985, and 1986. The data contains team SLA and OBA broken down according to 5 categories: leading off, runners on, runners in scoring position, runners on with 2 out, runners in scoring position with 2 out. From these, other categories can easily be computed; in fact, 13 different SLAs and OBAs can be obtained from the Elias data. They are:

- (1) total
- (2) bases empty
- (3) leading off
- (4) bases empty, not leading off
- (5) runners on
- (6) runners on with less than 2 out
- (7) runners on with 2 out
- (8) runner on first only
- (9) runners in scoring position
- (10) runner on first only with less than 2 out
- (11) runners in scoring position with less than 2 out
- (12) runner on first only with 2 out
- (13) runners in scoring position with 2 out.

Not being sure which categories might be most influential on RC, I took the formula $RC = AB * SLA * OBA$, and computed it 169 times, using all combinations of the 13 SLAs and 13 OBAs given above. I did this for each of the 78 teams from the years 1984, 1985, and 1986; I then computed the average error and average squared error for each of the 169 RCs, in predicting the 78 teams' runs for the three seasons. Three combinations of SLA and OBA outperformed the rest; they were

RC1 = $AB * SLA1 * OBA1$ (the original RC formula)
 RC2 = $AB * SLA5 * OBA1$ (SLA5 = slugging average with runners on)
 RC3 = $AB * SLA5 * OBA2$ (OBA2 = on-base average with bases empty)

These three RC formulas performed roughly the same for predicting the 78 teams' runs. The results are given below in terms of mean bias (MB = average amount over-estimating or under-estimating), mean absolute error (MAE = average amount of error in estimation, without regard to sign), and mean square error (MSE = average amount of squared error). For comparison purposes, I have also included the results for BJ1, BJ2, BJ3, as well as PJ1, PJ2, PJ3.

	MB	MAE	MSE
RC1	-3.1	21.8	724.3
RC2	10.3	21.1	648.9
RC3	-9.1	20.7	682.4
BJ1	-4.5	21.6	724.3
BJ2	-4.4	20.7	657.3
BJ3	-0.0	19.2	560.0
PJ1	1.9	19.4	560.7
PJ2	2.3	19.5	562.0
PJ3	-0.1	19.1	550.4

The 3 PJ estimators, as well as BJ3, are doing the best. The first 5 estimators (RC1, RC2, RC3, BJ1, and BJ2) are roughly equivalent, but inferior to the others. This is somewhat expected, as the more technical estimators use more statistical information from each team. The problem I considered then was how to improve on any of the RC's (using SLA and OBA) to make them comparable to the more technical versions of BJ and PJ.

As I examined the results from RC1, RC2, and RC3, I noticed that RC3 tended to underestimate runs scored (MB = -9.1), while RC2 tended to overestimate runs scored (MB = 10.3). Why is this? I'm guessing that RC3 is biased downward slightly, because it uses the OBA with bases empty, which occurs more often with good pitching than it does with poor pitching. That is, $OBA2 < OBA1$, generally. On the other hand, RC2 is biased upward slightly, because it uses the SLA with runners on base, which occurs more often with poor pitching than with good pitching. That is, $SLA5 > SLA1$, generally. (Actually, RC3 uses both OBA2 and SLA5, but apparently the downward bias of OBA2 is stronger than the upward bias of SLA5.) The actual number of runs scored tends to be somewhere in the middle, though this is not true in every case. This suggests that a better estimate of runs scored can be derived by averaging the estimates RC1, RC2, and RC3. I'll call this average RCA. The results were as follows (again from predicting the 78 teams' runs):

	MB	MAE	MSE
RCA	-0.6	17.9	514.0

This estimator is better than any of the others! I wasn't expecting such a big improvement, though in retrospect maybe I should have. It is often the case in statistics, that if you have several estimators of the same quantity, and the average bias is not too severe, then by averaging the estimators together you get a reduction in the error of estimation.

Is this a practical thing to do? Does it make sense? I wrestled with this for some time, and I must admit that I'm not sure. My guess is that if someone wanted to average all 9 of the estimators given in the table above, and use that for an estimate of runs scored, you would do even better! Certainly there is a limit to the sensibility of averaging different estimators together, particularly when they come from different estimation procedures.

I started looking for other ways to adapt the runs created formula. It appears that SLA and OBA have different effects on runs created when a batter is batting with runners on base as opposed to batting with the bases empty. Maybe what is needed is an appropriate mix of SLA2 and SLA5, along with an appropriate mix of OBA2 and OBA5. In reality, the overall slugging average and on-base average (SLA1 and OBA1) for any particular team are weighted averages of SLA2 and SLA5, and OBA2 and OBA5, respectively.

Combining the American League and National League statistics over the last three years gives us an equation like this:

$$SLA1 = .57 * SLA2 + .43 * SLA5,$$

where .57 is the proportion of at-bats that occurred with the bases empty, and .43 is the proportion of at-bats that occurred with runners on base. A similar equation holds for on-base average as well:

$$OBA1 = .55 * OBA2 + .45 * OBA5.$$

But are these proportions the right ones to use in computing runs created? If a batter makes a more important contribution by getting a hit with runners on base, it seems natural that the weight applied to SLA5 should be heavier than the proportion of at-bats with runners on base. Similarly, if OBA2 seems to be important, (since it did show up in the RC3 formula earlier), perhaps a heavier weight should be given to OBA2.

The upshot of this thought pattern was that I should try to come up with new weights for the slugging and on-base averages. I examined equations of the form:

$$RC = AB * SL * OB$$

where $SL = C1 * SLA2 + (1 - C1) * SLA5$,
and $OB = C2 * OBA2 + (1 - C2) * OBA5$,
and the weights C1 and C2 were chosen to minimize the mean error of estimation across all 78 teams. To minimize MAE, the best values of the constants turned out to be

$$C1 = .11 \quad \text{and} \quad C2 = .73, \quad \text{giving MAE} = 17.6.$$

To minimize MSE, the best values of the constants turned out to be

$$C1 = .18 \quad \text{and} \quad C2 = .66, \quad \text{giving MSE} = 502.3.$$

Since the values of C1 and C2 came out slightly different for minimizing the two types of error, I decided to compromise and use values of

$$C1 = .15 \quad \text{and} \quad C2 = .70, \quad \text{giving MB} = 1.1, \\ \text{MAE} = 17.7, \\ \text{and MSE} = 504.8.$$

You can see that we don't gain alot in additional reduction of MAE or MSE over what was already gained by averaging RC1, RC2, and RC3 together. However, I feel this last method for computing RC is helpful, as it sheds some light on the relative importance of slugging average with runners on-base, and on-base average with the bases empty. The same type of reasoning should help us in evaluating players individually. Thus, a good measure of a player's "slugging contribution" would be given by

$$SL = .15 * (\text{slugging average with bases empty}) + \\ .85 * (\text{slugging average with runners on base}).$$

A good measure of a player's "on-base contribution" would be given by

$$OB = .70 * (\text{on-base average with bases empty}) + .30 * (\text{on-base average with runners on base}).$$

Even though about 55% of the average batter's plate appearances occur with the bases empty, it appears that those at-bats are more important (70%) in determining the average batter's on-base contribution toward runs created. It also appears that these same bases-empty at-bats are of much less importance (15%) in determining the slugging contribution toward runs created.

One effect of using this new runs created formula is that it helps to separate two distinct functions of a hitter trying to create runs for his team. First of all, if the bases are empty, the hitter's primary goal is to reach base somehow, so that a teammate can knock him in. If there are runners on base, however, then a batter's goal to get a hit which advances the runners. The relative importance of each of these functions is shown in the calculation of SL and OB. If a hitter has a large OB value, then he should be hitting leadoff for his team. On the other hand, a hitter with a high SL value should probably be batting cleanup.

Traditionally, the two separate functions of getting on base and knocking in runs have been measured by counting runs scored and RBIs. But the idea here is not to have two separate statistics to measure these contributions, but rather to use the statistics jointly to measure RC. That is, the new SL and OB should really have little use or meaning by themselves; they are primarily useful together in the calculation of RC.

It might well be argued that the percentages (C1 and C2) used for an individual player should depend on how much that particular player batted with the bases empty. But the method for adjusting the percentages to fit the player needs more study. I hope to pursue this in the near future.

The data for computing these new slugging and on-base averages are readily available from either the Elias book or the new Great American Baseball Stat Book from Project Scoresheet. Below are the new slugging averages (SL), the new on-base averages (OB), and the new runs created figures (RC) for several prominent players from 1986. I've also shown the overall slugging average (SLA1), the overall on-base average (OBA1), and the runs created figured from them (RC1).

PLAYER	SL	OB	RC	SLA1	OBA1	RC1
J. Barfield	.530	.365	114	.559	.368	121
G. Bell	.514	.347	114	.532	.349	119
W. Boggs	.515	.443	132	.486	.453	128
J. Canseco	.567	.304	103	.457	.318	87
J. Carter	.500	.339	112	.514	.335	114
D. Mattingly	.536	.400	145	.573	.394	153
K. Puckett	.565	.366	141	.537	.366	134
J. Rice	.505	.376	117	.490	.384	113
M. Schmidt	.575	.378	120	.547	.390	118

MIRACLE TEAMS AND LONG SHOTS: WHAT ARE THE ODDS?

Sandy Sillman

Every spring when the baseball predictions are made, certain teams come out listed as 100-1 shots or 200-1 shots. And each year some of us wonder how bad the odds really are against a long-shot or miracle team winning the pennant. After all, in 1986 three of the four worst teams in baseball -- Cleveland, Texas and San Francisco -- arose to shake their respective leagues, and one is left with the feeling that, with a little bit of luck, either of these three could have won it all. Bill James (1984 and 1987 Baseball Abstracts) has speculated that no team should ever be a 100-1 shot, no matter how bad they are, since 'miracle teams seem to come along more like one time in 40 rather than once in 200.' Yet even Bill James wrote that he wondered how often it happens that a team rises to win a championship after finishing 25 games out or after finishing 44 games out the previous year (BA, p.108,139 and 219). And if Bill doesn't know, who does?

Well, you could look it up, as ol'Case would say -- and that's what I've done. I looked back through the 86 years of baseball history in order to find out what the odds really are against a long shot. I looked up how often a last-place team rises to finish first the following year. I looked through history in order to see what sort of records pennant-winning teams have the year before they won. I found out how often a team wins after finishing 25, 35, or 45 games out the previous year. And I've used this info to calculate what the odds should be -- the average odds, of course -- against each team based on its previous years' record.

Let me go through all this in the following order:

- (1) How often does a last place or near-last place team rise to win a championship?
- (2) What is the largest deficit in games behind that a championship team has overcome from one year to the next?
- (3) What are the average odds against a team based on its previous year's record?

Here goes....

(1) HOW OFTEN DOES A LAST PLACE OR NEAR-LAST PLACE TEAM RISE TO WIN A CHAMPIONSHIP?

This one is easy. In 86 years of baseball, there has been NO team that has won a championship after finishing in last place the year before -- not one. However there have

been four times when a team almost did it. The 1945 Washington Senators, a cellar-dweller in 1944, finished just one game behind the pennant-winning Detroit Tigers. And the '69 Mets, '67 Red Sox and '59 Los Angeles Dodgers won pennants after having finished within 2 games of last place the year before.

A next-to-last place team has gone on to win a pennant on 8 occasions. In addition to the three mentioned above, these rapid-rising pennant winners were: the '84 Cubs, '71 Reds, '46 Red Sox, '33 New York Giants, and '26 Yanks.

Of course, there are last-place teams and there are last-place teams. A team may finish last or next-to-last and still have a halfway-decent record, as Los Angeles did in '59 or Baltimore last year. Or a team may finish last with a terrible record, such as the 1985 Cleveland Indians. Rather than look at the standings, it is better to look at how far out a team finishes, in games behind, in order to determine what the odds are against it winning.

(2) WHAT IS THE LARGEST DEFICIT IN GAMES BEHIND THAT A CHAMPIONSHIP TEAM HAS OVERCOME FROM ONE YEAR TO THE NEXT?

The record for the largest deficit ever made up by a championship team in a single year belongs to the 1954 New York Giants. In 1953 they finished 35 games behind their arch-rivals, the Dodgers, before rising to win the pennant in '54. The '54 Giants were not a 'miracle' team, and unlike their kinsmen of 1951 were not long remembered for their feat. In 1954 they were a regular pennant contender coming off an off-year, having finished 5th in 1953.

The second-greatest deficit ever made up by a pennant-winning team -- 32 games -- was by the 1914 Boston Braves, the 'miracle' Braves.

These represent the only occasions when a team has won a championship after finishing more than 30 games behind the previous year. A total of 5 teams have won after finishing between 26 and 30 games out, and nine others have won after finishing between 21 and 25 games out. Here is a complete listing, (with games behind in parenthesis):

1954 Giants (35), 1914 Braves (32), 1945 Cubs (30), 1961 Reds (28), 1926 Yankees (27), 1920 Brooklyn Dodgers (27), 1944 St. Louis Browns (26), 1967 Red Sox (25), 1940 Tigers (25), 1969 Mets (24), 1934 Tigers (24), 1918 Cubs (24), 1924 Senators (23), 1912 Red Sox (23), 1907 Tigers (22), 1959 Los Angeles Dodgers (21).

Some of these teams were miracle teams. Others, like the '54 Giants and '26 Yankees were 'comeback' teams, regular pennant contenders recovering from an off year. Still others like the '40 Tigers and '45 Cubs were teams that picked up the pieces after the powerhouse from the previous year (the Yankees and Cardinals) collapsed. Incidentally, of these 18 teams only 2 repeated as champions the following year -- the '26

Yankees and '34 Tigers.

I have looked up the records of all 206 pennant and division-winning teams in the year before winning, going back to 1901. The record, in terms of the previous years' finish, forms a nice consistent pattern. The further back a team finishes, the less likely it is to go on and win. A team finishing 20 games back is less likely to win, historically, than a team finishing 15 games back. A team finishing 10 games back is less likely to win than a team finishing 5 games back. And a team that WON by 5 games is less likely to win again than is a team that won by 10 games.

Using this fact, we are now able to give the average odds against each team, based on its previous-year finish.

(3) WHAT ARE THE AVERAGE ODDS AGAINST A TEAM GIVEN ITS PREVIOUS YEAR'S RECORD?

In baseball history since 1901, a total of 140 teams have finished the season 40 games behind or more. Of these teams, not one has gone on to win a championship or even come close the following year. In view of this, it seems that a team that finishes 40 or more games out would historically rate odds of 200:1 or worse, unless there are strong indications that the team will improve.

There have been 220 teams that have finished between 31 and 40 games behind. Of these, only 2 have gone on to win a championship. A team that finishes between 31 and 40 games behind historically rates 100:1 odds.

Approximately 160 teams have finished between 26 and 30 games out, and of these teams 5 have won championships the following year. The historical odds against them: 33:1.

Here is a complete listing of the historical odds against a team based on its previous year's finish:

(1st-PLACE FINISHERS -- GAMES AHEAD)

15+	10-14	5-9	0-4
---	----	---	---
1:4	1:1	2:1	3:1

(GAMES BEHIND)

1-5	6-10	11-15	16-20
---	---	----	----
2.5:1	5:1	6:1	9:1

21-25	26-30	31-40	41+
----	----	----	---
15:1	32:1	100:1	200:1

(You may notice that there is one exception to the rule that a better finish makes for better odds. Teams that lose championships by slim margins are more likely to win the following year than are teams that WIN by slim margins. This fact should be no surprise to any baseball fan in the '80's. However 1st-place teams still win more often than 2nd-place teams.)

In view of all this, it seems historically justifiable to give odds against certain teams of 100:1 or even 200:1. Especially because this table shows only the historical odds; it does not account for what you know as an observer. A guy on the scene usually have some idea whether to expect improvement in a team, and this enables him to give better -- or worse -- odds than the team's record shows.

However it is also true that odds of 100:1 or more are regularly hung on teams that don't deserve it. This is especially true of a perennial loser -- it may be listed at 100:1 regardless of the quality of the current team. The Seattle Mariners in 1987 are being rated as 200:1 shots or worse -- 1000:1 in USA Today. Yet they finished only 24 games out in 1986, a finish which should rate them 15:1 odds. In fact the team probably rates even better odds than this since it is a young team who most people pick to improve.

By contrast, the Chicago Cubs are being listed as 10:1 or 20:1 shots in 1987. Yet the Cubs finished 32 games back in 1986. In addition the Cubs are an old team with few farm prospects, so that improvement may be even less likely than the record indicates. I would have no qualms against listing the Cubs as 100:1 shots this year.

The difference between Chicago and Seattle, of course, is that the Cubs are closing in on the franchise record for most years without a winning season. However the team record in 1984 has little effect on its chances for 1987. Based on the 1986 records and 1986 teams, the 'long-shot' Mariners are a potential dark horse -- and the Cubs, considered as a possible dark-horse, are really long shots.

SOME COMMENTS ON CHARLES PAVITT'S "BIAS IN FIELDING EVALUATION I: PITCHER HANDEDNESS"

By Dallas Adams

The June 1986 Baseball Analyst presented a paper by Charles Pavitt, in which he investigated the effect of pitching handedness on the assists per game ratio of infielders. Pavitt theorized that as, for example, a team's percentage of innings pitched by righthanders increased, the number of lefthanded batters in the opposing lineup would also increase. Further, because most batters pull the ball, more balls would be hit to the right side of the infield, fewer to the left side; and hence assists by the shortstop and third baseman should decrease while those by the first and second basemen should increase. Pavitt studied data for the 1980 and 1981 major league seasons, treating each infield position as a separate case. He writes:

"Thus, based on the correlations between assists and righthanded innings, we find evidence supporting the existence of the predicted bias in three of four cases. As expected, the greater the proportion of innings pitched by righthanders, the more balls are hit to the second basemen, and the less balls are hit to the shortstop and third basemen. The exception is first basemen, for whom assists have an insignificant negative correlation with righthanded innings. Pete Palmer has suggested to me that this surprising finding may be due to first basemen's positioning. Against lefthanded batters, they may be more likely to be playing near the base, allowing them a greater opportunity to make plays themselves (and getting putouts). Against righthanded batters, they are more likely to be far from the base, forcing them to flip the ball to the pitcher (and getting assists)."

Palmer's conjecture certainly sounds reasonable. If it were true, however, one would expect that lefthanded pitchers (Pavitt postulates that these will be facing a higher percentage of righthanded batters than are faced by righthanded pitchers) would compile higher putout totals per inning pitched than do righthanded pitchers.

Unfortunately this appears not to be the case. I took pitcher fielding data and innings pitched for 1980 and 1981, the two seasons which Pavitt used in his analysis, and tabulated them on the basis of pitcher handedness. As can be seen from the table below, lefthanded pitchers make fewer putouts (and more assists) per inning pitched than do righthanders!

	-----LEFTHANDED PITCHERS-----						-----RIGHTHANDED PITCHERS-----					
	IP	PO	A	E	DP	FA	IP	PO	A	E	DP	FA
1980 NL	4589	213	802	41	45	.961	12841	834	1867	147	131	.948
1980 AL	7089	321	1074	49	71	.966	13243	879	1888	142	158	.951
1981 NL	2944	142	548	25	24	.965	8692	577	1281	89	79	.945
1981 AL	4383	216	742	37	50	.9628	9077	637	1372	78	109	.9626
TOTALS	19105	892	3166	152	190	.964	43853	2927	6408	456	477	.953
TOTALS/IP		.0467	.166	.008	.010			.0667	.146	.010	.011	

Although my primary interest was in putouts for pitchers, I tabulated the other fielding statistical categories as well. It is interesting to note that clean chances (putouts plus assists) per inning pitched are essentially equal for both lefthanders (.21241) and righthanders (.21287). Moreover, pitcher double play frequency is also very similar for both groups. On the other hand, error frequency is much lower (and hence fielding average is much higher) for lefthanders. Indeed, in each of the four "league seasons" studied, lefthanded pitchers had the higher fielding average. This is a bit mystifying since I would have expected it to be easier to make an error on a throw than on a catch; yet it is the lefthanders, doing more throwing and less catching (as far as batted balls are concerned), who have the fewer errors. Thus, perhaps it is the more difficult pickoff throw which is the main source of additional errors for righthanders.

This survey was inspired by McCleery & Woods' article in issue 28. My intent was to identify baseball's greatest year by answering the question, In what season were the most great players in their primes?

First I made a consensus list of the fifty greatest players. Actually, I was able to compile the list up to eighty players but did not feel compelled to include all of these in the analysis. I used Bill James' lists in the Historical Abstract, Peter Palmer's lists in The Hidden Game and McCleery & Woods' list from the Analyst. I believe this method gives me a very accurate top fifty, although I wasn't too exact about the rankings.

Next, I carefully determined each player's five best years, again with reference to James' and Palmer's books. I decided that for non-pitchers these years must fall within a ten year span. I assumed that a player could not retain his peak skills beyond that period, especially hard-to-measure defensive skills. I made one exception, for Mize, due to the war. (See accompanying list.)

The result is that no one season stands above the rest. That suggests to me that the list is probably fair, with no bias towards extreme hitting or pitching years.

Five years had seven players: 34, 36, 69, 70, 77. Two years had six: 32 and 38. Twelve years had five, none before 1932. Six years before 1983 had none: 19, 26, 43, 44, 45, 78.

The number of player-seasons by decade is as follows:

90's: 3 10's: 28 30's: 49 50's: 33 70's: 42

00's: 17 20's: 18 40's: 19 60's: 35 80's: 6

Looking at five year stretches, the best is 1934-38 with thirty player-seasons. The players comprising that "Golden Age" are Greenberg and Gehrig (4 years); Foxx, Ott, Appling and Hubbell (3 years); Gehrig, DiMaggio, Mize and Waner (2 years); Grove and Cronin (1 year). Other leading eras are 1973-77 with 27 seasons and 1966-70 with 25. Before 1932 the top stretch is 1909-13 with 19.

Except for 1941-45, the worst five year periods began three years after the wars: 1922-26 (5) and 1948-52 (7). This suggests that many great careers were lost due to missed development years during the wars.

Among teams, the Giants have six of the fifty players. However, one of these is the borderline inclusion of Perry, who actually hit his peak with Cleveland. The Yankees are next with five. The Indians would have five if you credit them with Young and Perry. Surprisingly, the Dodgers have none. (Koufax is #62.)

Finally, the leagues are evenly represented. If Young is counted as NL, each league has 25 players.

- | | |
|--------------------------------------|---|
| 1. Ruth, Yankees 20-21-23-24-27 | 26. Yastrzemski, R. Sox 63-67-68-69-70 |
| 2. Cobb, Tigers 09-11-12-15-17 | 27. Kaline, Tigers 55-56-59-61-63 |
| 3. Aaron, Braves 57-59-61-62-63 | 28. Young, Red Sox 92-93-95-01-02 |
| 4. Mays, Giants 54-55-57-58-62 | 29. Bench, Reds 70-72-74-75-77 |
| 5. Williams, R. Sox 41-42-46-47-49 | 30. Gehrig, Tigers 29-34-35-36-37 |
| 6. Wagner, Pirates 05-06-07-08-09 | 31. Berra, Yankees 50-52-53-54-56 |
| 7. Gehrig, Yankees 27-30-31-34-36 | 32. Carlton, Phillies 72-76-77-80-82 |
| 8. Hornsby, Cards 20-21-22-24-25 | 33. Rose, Reds 68-69-70-73-76 |
| 9. Johnson, Senators 10-12-13-15-18 | 34. Greenberg, Tigers 34-35-37-38-40 |
| 10. DiMaggio, Yankees 37-38-39-40-41 | 35. Spahn, Braves 47-51-53-56-57 |
| 11. Musial, Cardinals 46-48-49-51-53 | 36. J. Jackson, Indians 11-12-13-16-20 |
| 12. Grove, Athletics 28-30-31-32-35 | 37. B. Robinson, Orioles 62-64-65-66-67 |
| 13. Speaker, R. Sox 12-13-14-16-17 | 38. Feller, Indians 39-40-41-46-47 |
| 14. Mantle, Yankees 55-56-57-58-61 | 39. Mize, Cardinals 37-38-39-40-47 |
| 15. E. Collins, A's 09-10-13-14-15 | 40. R. Jackson, A's 69-73-74-75-77 |
| 16. Mathewson, Giants 03-05-08-09-11 | 41. Gibson, Cardinals 66-68-69-70-72 |
| 17. Schmidt, Phillies 74-77-79-80-81 | 42. P. Waner, Pirates 27-28-32-34-36 |
| 18. Lajoie, Indians 01-03-04-06-10 | 43. Appling, W. Sox 33-35-36-37-40 |
| 19. Foxx, Athletics 32-33-34-36-38 | 44. Hubbell, Giants 32-33-34-35-36 |
| 20. Seaver, Mets 69-71-73-75-77 | 45. Brett, Royals 76-77-79-80-85 |
| 21. F. Robinson, Reds 59-61-62-64-66 | 46. McCovey, Giants 65-66-68-69-70 |
| 22. Morgan, Reds 72-73-74-75-76 | 47. Palmer, Orioles 70-73-75-76-77 |
| 23. Ott, Giants 29-32-34-36-38 | 48. Banks, Cubs 55-57-58-59-60 |
| 24. Alexander, Phils 11-15-16-17-20 | 49. Cronin, Senators 30-31-32-33-38 |
| 25. Mathews, Braves 53-54-55-59-60 | 50. G. Perry, Giants 66-69-70-72-74 |

The next 30: Clemente, Carew, Marichal, Santo, Cochrane, Ford, Dickey, Crawford, Roberts, Frisch, Hartnett, Koufax, Killebrew, Goslin, Fingers, Vaughan, F. Baker, Wilhelm, Stargell, D. Allen, P. Niekro, Doerr, Boudreau, M. Brown, Winfield, A. Simmons, Dean, B. Williams, Wallace, Gossage.

The Implications of Leadership Research on Baseball:
2. Current Theory
Charles Pavitt

In *Analyst* 27 I began a two-part series of articles concerning the match between manager and team, specifically regarding Bill James' claim in the 1985 *Abstract* (p. 214); "Managers have different skills . . . It is never a question of their goodness or badness, but of their suitability to the needs of the organization at this moment." In the earlier article, I discussed evidence in the social science journals suggesting that changing managers has a negative impact on a team's performance to the extent that the change is "disruptive" (occurs during the season and includes a replacement from outside of the organization). In this concluding article, I wish to describe some work in leadership theory that stands in strong support of Bill's general claim while giving us with some insight into the manager/team match.

Scientific research on group leadership began in the 1920's. At the time, scientists were apparently under the influence of what has come to be known as the "Great Man Theory of Leadership," the idea that certain people have special characteristics that set them apart from the common folk, so that they will be successful leaders in any situation. The scientific task was to discover what these characteristics are. First, scientists tried physical characteristics. While they found a disproportionate tendency, for example, for taller people to become leaders, neither height nor any other physical characteristic had any effect on a leader's performance. By the 1940's, scientists had turned to personality characteristics, correlating measures of some trait hypothesized to be relevant to leadership (such as "intelligence" or "initiative") with measures of leadership performance. Correlations were rarely as high as .25, which is too small a relationship to have any practical value. Further, starting in the late 1940's, research appeared showing that leaders successful in one type of task could not be expected to be successful in a different type of task. These studies struck at the heart of the "Great Man" theory, revealing its basic inadequacy.

Let me clarify the implications of all of this. While many, if not most, people seem to believe that some people are born leaders and everyone else is not, the available evidence suggests otherwise. Leadership is situational. One should not expect a leader who has been successful in one type of situation to be successful in another type. This conclusion, however, begs at least two questions. First, what are the possible types of situations in which leadership is relevant? Second, what types of leaders would be successful in each of them?

Several answers to these questions have been proposed. The most successful proposal was made about twenty years ago by social psychologist Fred Fiedler. He suggested that situations can be classified according to their favorableness, the degree to which the leader can act and expect these actions to succeed. Situational favorableness is a function of three dimensions. The most important dimension is leader-member relations, the degree to which the leader has the support and loyalty of the subordi-

nates. As with the other dimensions, leader-member relations is categorized at either end of a continuum; in this case, it can be either good (favorable) or bad (unfavorable). The second dimension, task structure, is the degree to which the requirements for completing the task are clearly specified. Task structure can be either clear (favorable) or unclear (unfavorable). Task structure is considered one-half as important as leader-member relations. The third dimension is position power, the degree to which the position of leadership (independently of the person filling the position) allows the leader to reward or punish subordinates. Position power is either strong (favorable) or weak (unfavorable). Position power is considered one-half as important as task structure, and thus one-fourth as important as leader-member relations.

A typology of situations can be formed by assigning to each situation a numerical weight analogous to its relative importance (4 for leader-member relations, 2 for task structure, 1 for position power) and assigning each situation the summed weight of its favorable features. This results in:

Situation	Leader-Member Relations	Task Structure	Position Power	Overall Favorableness
I	Good (4)	Clear (2)	Strong (1)	Very High (7)
II	Good (4)	Clear (2)	Weak	High (6)
III	Good (4)	Unclear	Strong (1)	High (5)
IV	Good (4)	Unclear	Weak	Medium (4)
V	Bad	Clear (2)	Strong (1)	Medium (3)
VI	Bad	Clear (2)	Weak	Low (2)
VII	Bad	Unclear	Strong (1)	Low (1)
VIII	Bad	Unclear	Weak	Very Low (0)

Fiedler's theory has thus provided a possible answer for the first question above: what are the possible types of leadership-relevant situations? His research has provided a possible answer to the second question: what types of leaders are successful in each of them? Fiedler proposed that leader performance is a result of leadership style, which is in turn a result of the leader's characteristic manner of viewing subordinates. Most simply, in circumstances of low favorableness, the most successful leaders tend to be those who only evaluate their subordinates according to the subordinate's performance in the group. For these leaders, a bad worker is a bad person. These leaders believe that an orientation toward the task, even at the expense of the feelings of the group members, is conducive to good group performance. In circumstances of medium favorableness, the most successful leaders are those who can evaluate their subordinates as people independently of their performance. For these leaders, a bad worker may very well be a good person. These leaders believe that good interpersonal relationships among group members, even at the expense of group efficiency, is conducive to good group performance. In circumstances of high favorableness, surprisingly, the "bad worker is a bad person" kind of leader is most successful. Apparently, now surrounded by good workers, the leader sees them all as good people and maintains strong interpersonal relations with them. Research has shown correlations

between trait measures of this leader characteristic to correlate in the order of .4 to .6 with leader performance within each situation, high enough to engender some trust in their relationship.

Now, let us relate all of this specifically to baseball. First, it is important to note that the task structure of baseball is always unclear. This is obvious when viewed relative to what would be considered a clear situation, say an assembly line, where we know that if we do steps 1 to 10, a Ford will pop out. The plan for winning baseball games cannot be analogously specified, not when there is an opposing team intent to thwart it, and not when one has to ad lib strategy throughout the game depending on the score. A Fiedlerian typology for baseball would require only:

Situation	Leader-Member Relations	Position Power	Overall Favorableness
III	Good	Strong	High
IV	Good	Weak	Medium
VII	Bad	Strong	Low
VIII	Bad	Weak	Very Low

Now an attempt to provide some scenarios for these situations. Imagine a team with a rich and impatient owner prone to public criticism of his players and quick to send young players back to the minors and old players out to pasture. Such behaviors may bring about in players a mistrust for the entire organizational leadership chain (bad leader-member relations). The manager is likewise subject to attack from and removal by the owner (weak position power). This situation is very unfavorable, and as a result requires a "bad worker is bad person" manager. Let us call this manager "Billy M." [1] He is successful, so the owner steps back and leaves the players alone. As a result, leader-member relations improves. But wait - good leader-member relations and poor position power means a situation with medium favorableness. So Billy M. starts having problems, and its time for him to go, and be replaced by a "bad worker may be good person" manager. Let us call this manager "Yogi B." So far so good - until the team slumps. Then the owner's behavior reverts, leader-member relations nosedives, and its time to bring back Billy M.

Imagine a second scenario. A team has a lot of young players respectful of the managerial position (good leader-member relations). The manager is also young and unproven, working under a "probationary" one-year contract (weak position power). This means medium favorableness and a "bad worker - good person" manager ("Chuck C."). The team shows promise, so the manager sticks around a few years, maybe gets a multi-year contract (strong position power). Now we have high favorableness and a need for a "bad worker - bad person" manager ("Dick W.").

Let me try one more. An organization has a long history of success, along with a reputation for fair-handedness and loyalty to its players (resulting in high leader-member relations). It may only give its manager a one-year contract, but it has been giving the same manager one-year contracts for twenty years (strong position power). The situation is highly favorable, sug-

gesting a "bad worker - bad player" manager. Let us call this manager "Walter A."

My attempt here to account for managerial success through Fiedler's theory does not always work. It fails to explain why a "bad player - good person" manager ("Tommy L.") has also succeeded in the last situation. As for those situations I have attempted to explain, I may have misrepresented matters to some degree; I am sure I have oversimplified them. More basically, there is no question that Fiedler's theory is an oversimplification; but it is a good early attempt to account for the fact that leadership is situation-dependent, and his research is more support for this fact. We should always remember that baseball is no exception to this fact, no matter what the best theory turns out to be. One interesting implication of all this is that because of its public visibility, we can learn more about leadership in baseball and other major sports than about leadership in other areas. As a result, we might be able to use this knowledge to help us propose better general theories of the leader-situation relationship than Fiedler's. It would be a great boon to sabermetrics if our work could be shown to have demonstrable relevance to other areas.

For further reading on leadership in general, consult Bernard M. Bass' Handbook of Leadership. More specifically relevant to this article are two articles by Fiedler in Advances in Experimental Social Psychology, edited by Leonard Berkowitz; the first in Volume 1 (1964), the second in Volume 11 (1978).

Footnote 1 - Any similarities between these fictional characters and actual people is totally intentional.

ON THE INACCURACY OF THE PYTHAGOREAN EQUATION AT EXTREME SCORING RATIOS

By Dallas Adams

The "Pythagorean Equation" of Bill James is familiar to readers of The Baseball Analyst. It states that the ratio of a team's wins and losses is approximately equal to the ratio between the squares of the team's runs scored and runs allowed.¹ Symbolically:

$$WL\% = \frac{R^2}{R^2 + OR^2}$$

which can be rewritten as

$$\frac{W}{L} = \left(\frac{R}{OR} \right)^2$$

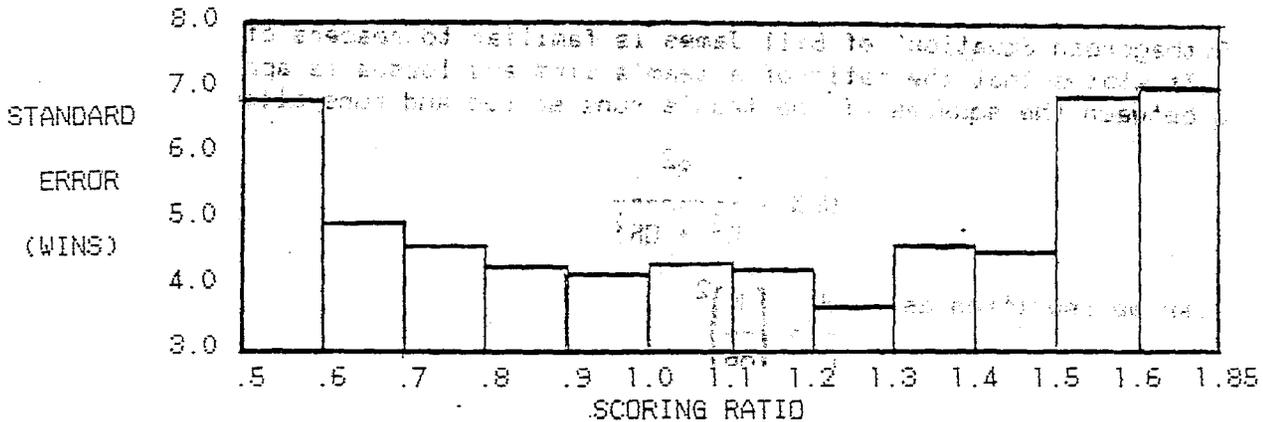
It is recognized that the equation generally works well when applied to teams. That is, when dealing with team scoring ratios (SR), where SR=R/OR, the equation is usefully accurate. If it is applied to all 20th century teams, the Standard Error is approximately 4.2 wins. Thus, about two times in three the number of wins predicted by the equation will be within 4.2 of the team's actual wins total; and about 19 times in 20 the predicted wins will be within twice 4.2 (i.e. 8.4 wins) of the actual figure.

Note, though, that team scoring ratios tend to be reasonably close to 1.00. All but 11 major league teams this century have had SR's between 0.6 and 1.6; and the remaining 11 haven't been far outside those limits. For example, the modern (i.e. since 1900) major league record scoring ratio is 1.848 by the 1906 Chicago Cubs.

There is, however, at least one application where the Pythagorean Equation is used on scoring ratios far beyond the major league team range. This application is the computation of what Bill James calls Offensive Winning Percentage,² the estimate of a player's offensive value as described in terms of his theoretical probability of winning versus an average performance within his league. The problem here is that often the batters whose Offensive Winning Percentages are of particular interest (e.g. the outstanding batters in baseball) often create runs at two or three times (or even more) the average league rate. Hence, in applying the Pythagorean Equation to the Offensive Winning Percentage of such batters, one is using SR's of, say, 3.0 or 3.5.

Because the equation has been determined to be reasonably accurate for teams (i.e. for SR's between 0.6 and 1.6) it has hitherto been assumed to retain its accuracy at far higher (and far lower) scoring ratios. Unfortunately this is not the case. I tested the accuracy of the equation by sorting all major league teams from 1901 to 1980 into groups of similar SR, each group (except the last) covering a range of .10 in scoring ratio. Then, after applying the Pythagorean Equation to all teams, I computed the Standard Error for each group. The results are shown below in both tabular and graphical form.

SCORING RATIO RANGE	STANDARD ERROR	NUMBER OF TEAMS
0.50-0.59	6.9092	6
0.60-0.69	4.9416	47
0.70-0.79	4.6098	146
0.80-0.89	4.2350	212
0.90-0.99	4.1455	259
1.00-1.09	4.2634	271
1.10-1.19	4.2090	229
1.20-1.29	3.7221	132
1.30-1.39	4.5620	71
1.40-1.49	4.5358	27
1.50-1.59	6.8596	9
1.60-1.85	7.0054	5



It is evident that the equation is most accurate for SR's between 1.20 and 1.30; and the further one gets from that range, the greater the Standard Error. That error is less than 4 wins in the 1.20-1.30 SR group, yet it is approximately 7 wins when dealing with SR's in the 0.50-0.60 and 1.60-1.85 ranges.

The scoring ratio interval of 0.50-1.85 covers the entire range of 20th century major league play. By utilizing 19th century major league and 20th century minor league data it is possible to extend the range coverage slightly. For example, I was able to find seven teams³ with SR's exceeding the 1.848 record of the 1906 Cubs. Treating these seven teams as a group (their composite SR equals 2.0) and applying the Pythagorean Equation as before (and prorating the results to a 154 game season), one finds that the Standard Error is 8.36 wins.⁴

For the other extreme, I located another seven teams⁵ with SR's below 0.50. Repeating the calculations for this group (their collective SR is 0.46) gives a Standard Error of 7.04 wins in a 154 game season.

At both ends, then, these hyper-extreme cases extend the error trend noted for 20th century major league teams: the more extreme one gets in scoring ratio, the greater the standard error in the Pythagorean Equation's prediction. This casts much doubt on the accuracy of Offensive Winning Percentage computations for great batsmen when those calculations involve the use of the Pythagorean Equation.

-
1. Bill James, 1986 Baseball Abstract, page 338
 2. Bill James, The Bill James Historical Abstract, pages 293-297
 3. 1876 Chicago NL (2.43 SR), 1884 St. Louis UA (2.07), 1885 New York NL (1.87), 1947 Havana FLA INT (1.852), 1973 Texas GCL (1.855), 1975 Texas GCL (2.09), 1979 Paintsville APP (1.849)
 4. which corresponds, in 154 games, to a .054 Standard Error in WL%
 5. 1876 Cincinnati NL (.41 SR), 1883 Philadelphia NL (.49), 1890 Pittsburgh NL (.48), 1899 Cleveland NL (.42), 1946 Portland N ENG (.47), 1946 Fall River N ENG (.48), 1946 Walden N ATL (.43). Note, there were also three teams in the 1884 Union Association with SR's below 0.50; but these teams played only briefly and disbanded early in the season, hence I did not include them.

STRIKEOUT-TO-WALK RATIOS AND WINNING RECORDS

by Russ Eagle

One of the questions Bill James asked in his "Suggestions for Areas of Research" in the March edition of the Baseball Analyst was "What is the connection between strikeout-to-walk ratio and a winning record?" I've constructed the chart on the next page, similar to the one Mr. James suggested, in an effort to answer this question.

I began by identifying all pitchers during the period 1970-1984 who pitched at least 162 innings. For strike seasons this requirement was lowered to one inning pitched per team game. There were 1274 pitchers who qualified (I think I got them all, I may have missed one or two). For each pitcher I recorded their won-lost record, calculated their K/W ratio, and recorded their team's won-lost record. The source used was The Sports Encyclopedia: Baseball.

Rather than try to explain each column of the chart, I'll just describe one row, the one marked by an asterisk. This line tells you that there were 46 pitchers with a strikeout-to-walk ratio between 2.75/1 and 2.99/1, inclusive. Thirty of these pitchers had winning records. That's 65%. Eleven had losing records, while five had .500 won-lost percentages. As a group these pitchers were 674-502, a winning percentage of .573. Thirty-two pitchers, or 70% of the 46, pitched for winning teams. Fourteen pitched for losing clubs. None pitched for teams with a .500 record.

I decided to record the team data as I went along because, for one thing, it was convenient to do so, but also because I thought it might clarify things a bit further. The correlation between K/W ratio and won/lost record suggested by the chart may be exaggerated a bit by the quality of team the pitcher plays for. For example, the pitchers with the lower K/W ratios are going to be on losing teams more often than not. A losing team is much more likely to give 162 innings of work to such a pitcher. The team data in the final four columns of the chart may be of use to anyone who wishes to study the issue further.

As the chart shows, the percentage of pitchers with winning records does increase as K/W ratio improves. A more liberal grouping of the data yields the following:

<u>K/W RATIO</u>	<u># PITCHERS</u>	<u>% WINNING RECORDS</u>	<u>WINNING %</u>
4.00/1 or better	8	100%	.648
3.00/1 - 3.99/1	57	88%	.610
2.00/1 - 2.99/1	359	69%	.559
0.50/1 - 1.99/1	850	50%	.510

Some may be surprised initially that 50% of the latter group had winning records, but remember that we're talking about pitchers with 162-plus innings. A guy's not likely to get that many innings unless he has something going for him. In fact, the 1274 pitchers in this study have a combined winning percentage of .530, and they are 1895 games over .500. That means the pitchers with less than 162 innings are 1895 games below .500.

There's a lot more that can be done with the data in the chart. I plan to use it first of all as part of a study I'm working on on 20-game winners. For now I present it as is, in hopes that it can be of some use to others.

STRIKEOUTS TO WALKS: EFFECTS ON WINNING

K/W RATIO	# PITCHERS	WINNING RECORDS	% WINNERS	LOSING RECORDS	.500	WINS	LOSSES	PCT.	WINNING TEAMS	LOSING TEAMS	.500 TEAMS	% WITH WINNING TEAMS
5.00/1 or better	2	2	100	0	0	49	25	.766	2	0	0	100
4.75/1 - 4.99/1	1	1	100	0	0	24	10	.706	1	0	0	100
4.50/1 - 4.74/1	2	2	100	0	0	42	26	.618	2	0	0	100
4.25/1 - 4.49/1	0	-	-	-	-	-	-	-	-	-	-	-
4.00/1 - 4.24/1	3	3	100	0	0	38	22	.633	2	1	0	67
3.75/1 - 3.99/1	9	9	100	0	0	137	98	.582	3	5	1	33
3.50/1 - 3.74/1	10	9	90	1	0	159	104	.605	6	4	0	60
3.25/1 - 3.49/1	18	15	83	2	1	315	185	.630	15	2	1	83
3.00/1 - 3.24/1	20	17	85	2	1	331	215	.606	15	5	0	75
*2.75/1 - 2.99/1	46	30	65	11	5	674	502	.573	32	14	0	70
2.50/1 - 2.74/1	76	54	71	16	6	1106	856	.564	51	24	1	67
2.25/1 - 2.49/1	102	71	70	27	4	1489	1158	.563	62	34	6	61
2.00/1 - 2.24/1	135	92	68	34	9	1902	1558	.550	75	53	7	56
1.75/1 - 1.99/1	184	104	57	65	15	2338	2105	.526	107	75	2	58
1.50/1 - 1.74/1	203	108	53	82	13	2531	2392	.514	96	95	12	47
1.25/1 - 1.49/1	247	131	53	99	17	2978	2818	.514	130	113	4	53
1.00/1 - 1.24/1	162	70	43	77	15	1894	1914	.497	70	86	6	43
0.75/1 - 0.99/1	49	12	24	32	5	491	591	.454	16	31	2	33
0.50/1 - 0.74/1	5	1	20	4	0	42	66	.389	1	4	0	20

One final comment. The line marked with the asterisk is the only one that seems out of line with the others. The group winning percentage of .573 seems to be about right (in relation to the others on the chart), but the percentage of pitchers with winning records does not. Only 65% of these pitchers had winning records, even though they had a good K/W ratio and 70% of them pitched for winning teams. I double-checked everything, and the data is correct. I went back and looked at the team records for the pitchers involved, compiling the winning percentage of the 46 teams combined. Again it fit neatly in with the values for the other groups. I really don't have any idea why only 65% of these pitchers were winners. I did notice that 11 of the 16 non-winning pitchers were within at least 2 games of .500. If just four of these had managed to win another game or two, the percentage of winners would have jumped to 74%, so maybe I'm making too much of it.

20