By the Numbers

The Newsletter of the Statistical Analysis Committee of the Society for American Baseball Research

Volume 5, Number 2 June, 1993

Greetings

The baseball season is well under way, providing us with a great deal of enjoyment, surprises, and fodder for statistical analysis. I hope the Statistical Analysis Committee can make the same claim!

As you know, SABR's national convention is upcoming in San Diego on June 24-27. The Committee meeting is scheduled for Friday afternoon (June 25) from 12 noon to 2 p.m. I hope to see you there. If you have any topics that we should discuss, please let me know.

In the past, the newsletter following the convention included material of a statistical nature which was presented at the convention (typically by committee members). This allows members who are unable to attend the convention to see the research. If you are planning on a presentation at the convention, please give me a copy of the material if you would like to have it appear in the next issue of the newsletter.

The current issue of the newsletter is chock full of interesting and thought-provoking articles. In fact, there is a welcome backlog of material that we didn't have room to include in this issue. Keep sending in material. If the backlog continues to grow, we can publish the newsletter more frequently.

Following this greeting, we have a section on news and views from committee members. We then have eight articles.

Pete DeCoursey kicks things off with a piece on the soundness of the Atlanta Braves pitching staff. You may be surprised at the results of his analysis (written before the season started).

Willie Runquist discusses the conventional methodology used to estimate park effects.

Daniel Levitt investigates the effect of strikeouts, walks, and ground-to-air ratios on ERAs.

Don Coffin analyses past expansion years to help us set expectations for this year of expansion.

Bill Gilbert responds to **Barry Codell's** introduction of the Diamond Weight statistic by way of another measure of offensive performance, Bases per Plate Appearance.

I write on the hidden impact of the DH, arguing that it has detrimentally altered the makeup of the two leagues.

Willie Runquist checks in again with a piece on the variability in baseball players' performances.

Dick Adams follows up on an earlier piece by **Dave Smith** dealing with the quality start statistic and a "Moss Klein" pitcher.

Please send material, comments, etc., to my address: **Rob Wood**, 2101 California St. #224, Mountain View, CA 94040. My home number is (415) 961-6574, and my daytime number is (415) 854-7101.

****

Notes from committee members

In this section of the newsletter, I will pass along news/ideas/information I receive from committee members.

Murray Browne writes that his 4th Annual Game Score Report (covering the 1992 season) is available. Send $2.50 to Murray at 645 Midway, Holland, MI 49423, or call him at (616) 335-8708.

Dan Heisman has enhanced **Bill James'** Brock-2 spreadsheet model to predict hitter's performance.

Contact Dan at 1263 Paso Fino Dr., Warrington, PA, 18976; (215) 343-6033 evenings.

Jon McEntyre has written an interesting report on hitters' Run Production (runs plus RBI) through the 1992 season. Contact Jon at 14 Harbor Drive, Corte Madera, CA 94925; (415) 927-7531.

Bruce Stone has compiled team by team data on earned vs. unearned runs and games won/lost due to unearned runs. The impact is greater than you might first believe. If you would like a copy of the data, or if you have worked in this area, contact Bruce at 5054 Chowen Ave. So., Minneapolis, MN 55410; (612) 926-6087.

****

ARM WARNINGS

by Pete DeCoursey

The Atlanta Braves may think their trifecta of Greg Maddux, Tommy Glavine and John Smoltz is poised to bring home 50 victories per year for the next five years, but instead, they have assembled a troika of hurlers whose history and use pattern threatens their careers.

The Braves are not alone. While many teams have attempted to protect the massive investments they make in star hurlers, there are still many stars and promising prospects whose tomorrows were short-sightedly traded for a few seasons worth of yesterdays.

Several years ago, Craig Wright published ground-breaking work in *The Diamond Appraised* on why pitchers who had early success often break down and are virtually done by the age of 30.

Wright isolated two reasons for pitcher breakdowns: overuse in total innings pitched before age 24 and overuse by way of pitches per game (or batters faced per game) within a given season.

His conclusions have long seemed worthy of an update because if they are proven to be true over the years, they will be genuinely useful to anyone who wants to know why apparently flourishing pitchers suddenly decline.

Wright suggests that one way to distinguish between pitchers whose good times end with their 20s and those who last well into their 30s is to look at their early use totals.

Those who pitched more than 1200 professional innings or so by age 24 were more likely to have finished their effective careers by age 28 than those who were used gently, particularly from ages 18-22, Wright found.

In fact, some pitchers like Jim Palmer benefited from early injuries which kept their arms from being overloaded by managers eagerly seeking a 300-inning ace.

After considering Wright's theories, I noticed that pitchers like Dwight Gooden, Bret Saberhagen and Fernando Valenzuela pitched more than 1200 innings before the end of the season when they were 24 years old, and all seemed to pay a high price for their early excellence.

The price? Their primes have been riddled with injuries and ineffectiveness which mock their early promise. Like Marley's Ghost, you can hear the chains on their pitching arms clank as they approach.

Valenzuela is the ghost of Overworked Pitchers Past. Five years ago, Valenzuela was 26 and had just recorded a 21-11 record with 242 strikeouts and a 3.14 ERA for the Dodgers. But by his 24th birthday, he had pitched 1482 professional innings and faced 30-32 batters per game, over Wright's limit of 28 or so for developing arms.

What happened then? He never had a winning record after his 27th birthday and was finished by the age of 31.

The Mets have two candidates for Ghost of Overworked Pitchers Present: Gooden and Saberhagen. The former Royal pitched 1253 pro innings before he turned 24, and has been injured and inconsistent ever since.

Gooden worked 1441 pro innings before his 24th birthday rolled around in 1988. He then lost half of 1989, set a personal record for worst ERA at 3.83 in 1990 followed by the 2nd and 3rd highest ERAs of his career in 1992 and 1991, respectively. Since 1989, his strikeouts per inning are down while his hits and walks per inning are up.

These three made me wonder how pronounced the effect was, which called for a study. I looked at all major league pitchers from 1991 and 1992 and made a list of everyone who had amassed 1025 or more professional innings before the end of the season of their 24th year.

The answer appears to be that early overuse seems to affect fastball-slider pitchers more than curveball

pitchers. As you will see in the chart (on the next page) of early over-used pitchers, flame-throwers like Gooden, Valenzuela and Saberhagen are joined by other "I relied on my heat" types like Mark Gubicza, Dan Petry, Richard Dotson and Mike Witt. Hard slider hurler Andy Hawkins is also on the list, but how junkballer Dave LaPoint made the tally, I don't know.

They are balanced by a group of finesse and curveball pitchers who did succeed in their 30s: Bert Blyleven, Jerry Reuss, Dennis Martinez, Bill Gullickson, Bill Wegman, Bob Knepper and Mike LaCoss.

The link, as noted earlier, appears to be twofold: for Blyleven, Gullickson, Reuss, Knepper and LaCoss, their careers depended on incredible curve balls.

Martinez, by the way, did not escape the consequences of early overuse. After winning 83 games and losing 57 with ERAs under 4 to the age of 27, Martinez fell apart from ages 28 to 31, compiling a 29-42 record with ERAs well over 5.

But at 32, having arrived in Montreal, it all came back as smoothly as the motion that started producing the speed and location that made the Expo hurler one of the best pitchers around, again. He has won 81 games against 57 losses since 1987.

This shows it is possible to return from the abyss of arm abuse careless management may drop a pitcher into, but it seems to be a rare occurrence.

Dennis Martinez, by the way, has come back so far that after striking out three hitters per nine innings for most of his career, the 38-year-old has averaged striking out six hitters per nine innings over the past four seasons.

But perhaps the best way to see that there is something to the dangers of early overuse is that five and ten years ago, there were dozens and dozens of pitchers in the majors who had been overused in their pro careers by Wright's standards.

Now there are just a few. The Dodgers are a good example of the change. By the time Ramon Martinez reached the age of 21, he had thrown 735 professional innings. By the end of his 24th birthday season last year, Martinez had tossed 1202 pro innings, and used up what looked like a career's worth of hard-to-hit fastballs before he turned 25.

Do the Dodgers blame it on overuse? Well, Ramon's brother Pedro turned 21 last year, but has only pitched 388 pro innings so far, slightly more than half of the innings Ramon hurled to that point in his career.

Other organizations, however, may still have a lot to learn. Craig McMurtry of the Braves lost his career to early overuse, but the team never seemed to recognize it.

McMurtry threw 202 innings for AA Savannah at the age of 21, and then threw 210 innings for AAA Richmond. Two major league seasons and 408 major league innings later, he had won 26 games and lost 28 and his career was over at the age of 24. He won two games and lost 15 after that, and managed only a 28-26 record in the minors.

But the fact that Tommy Glavine and John Smoltz both have pitched well over 1000 innings in their formative years may show the Braves have not learned that slow and easy builds the ace.

Smoltz, a fastball-slider devotee, could be headed for severe problems by the age of 28, a la Gooden, Saberhagen and Valenzuela, but Glavine, a finesse pitcher with lovely mechanics, may have a greater immunity to the effects of overuse.

But the Braves either don't see or don't agree that Smoltz is in danger, or else they wouldn't have imported a similar pitcher with the same skills and even more early abuse.

New Brave and current Cy Young award winner Greg Maddux may look good now, but the dance is almost over and the bill is about to come due.

He is a fastball pitcher working in a hitter's park, and it says here that after another year or maybe two, his years as one of the best pitchers in the big leagues will be over, at least for a while.

Maddux is my choice for the Dennis Martinez award, since I think he has the brains, finesse and make-up to adjust and emerge from his coming overuse slump after a few ugly seasons. But his pre-eminence in the NL will be over by 1995, at the age of 28, and you read it here first.

The Braves' aces are not the only current star pitchers who appear headed for peril. Ramon Martinez, Tommy Greene and Melido Perez are all fastball-slider pitchers whose early use totals make me wonder about their futures.

On the other hand, breaking ball types like Bill Wegman and Sid Fernandez probably will employ the Blyleven exception to the overuse curse, and not show many effects.

I THOUGHT THEY PASSED YOUTH LABOR LAWS

(pitchers who threw more than 1025 innings up through the season in which they turned 24)

Pitcher	Age	Pro IP before 24	ML W-L before 24	ML W-L after 28
Over-Pitched Early and Gone				
Fernando Valenzuela	32	1758	128-103	13-15
Mark Gubicza	30	1088	88-72	16-18
Mike Witt	32	1096	100-89	14-25
Richard Dotson	33	1173	94-88	17-25
Dan Petry	34	1425	97-73	28-31
Mark Eichorn	32	1056	29-23	9-12
Andy Hawkins	33	1207	56-52	28-39
Roy Smith	31	1019	10-11	20-20
Curveballers Do It Longer				
Bert Blyleven	42	1734	148-132	139-118
Jerry Reuss	44	1255	105-92	115-99
Bob Knepper	38	1318	61-70	85-85
Bill Gullickson	34	1289	101-86	44-36
Mike LaCoss	37	1187	50-53	48-50
Dave LaPoint	34	1066	53-53	27-33
Anomalies				
Dennis Martinez	38	1317	89-73	104-83
Dennis Eckersley	39	1355	120-99	61-46
Sid Fernandez	30	1177	78-59	15-14
Yet To Be Seen (pitchers with a history of early overuse who have not yet pitched after age 28)				
Greg Maddux	26	1332	95-75	--------
Dwight Gooden	28	1307	142-66	--------
Tommy Glavine	27	1182	73-60	--------
Bret Saberhagen	29	1254	113-83	--------
Ramon Martinez	25	1202	52-37	--------
Melido Perez	27	1146	58-62	--------
Tommy Greene	26	1048	20-15	--------
Jose Mesa	27	1042	17-28	--------
John Habyan	29	1042	18-18	--------
Mark Grant	29	1270	22-31	--------

One list we are missing is that of pitchers who were overworked by college coaches rather than major league organizations.

I don't know how to assess the kind of overuse which threatens the careers of pitchers like Tyler Green of the Phillies, because the complete records are not currently available, and it is hard to look at the available records and know what they mean, without day-to-day use charts.

Without those records, however, we can still monitor the use of young pitchers, and since I figured some of you

probably are interested in young prospects, here are 14 minor league prospects who are on the road to arm abuse.

Keep in mind that overused young pitchers normally break onto the major league scene with success, only to lose their abilities a few years later, as they turn 27-30.

The minor league hurlers in the chart below, then, are somewhere between Craig McMurtry and Bret Saberhagen, but where they are, what they throw or which group they fit into is still unknown. What is known about them? Only this: their odometers are spinning rapidly.

Misspent Youths (pitchers on the path to professional overuse before age 24)

Pitcher	Age	Team	IP before 24*	Arm Abuse Years Left**
Nate Minchey	23	Red Sox	857	2
Royal Thomas	23	Padres	841	2
Mike Hampton	20	Mariners	375	5
Johnny Ruffin	21	Chi Sox	539	4
Brons Patrick	22	A's	654	3
Tavo Alvarez	21	Expos	391	4
Salomon Torres	21	Giants	373	4
Jim Converse	21	Mariners	363	4
Greg Hansell	22	Dodgers	522	3
Kurt Miller	20	Pirates	344	5
John Roper	21	Reds	381	4
Jose Martinez	22	Marlins	500	3
Joel Adamson	21	Marlins	307	4
Carlos Pulido	21	Twins	280	4

* IP before 24 is the number of innings pitched before the season in which they were 24 for most of the year
** Arm abuse years left is the number of years before their 24th year season

Another indicator Wright targeted was pitches per start. He and fellow STATS Inc. colleague and author Don Zminda have said throwing more than 130 pitches in a game abuses a pitcher's arm.

So what do you think will happen to pitchers like the Phillies Ben Rivera, who averaged throwing 130 pitches in each of his 14 starts? or to Curt Schilling of the Phils, who averaged 125 pitches in each of his 26 starts?

For some pitchers, however, that magic level where the arm turns to mush is much lower than 130. Phillies pitchers Jose DeJesus and Tommy Greene both set personal highs in starts and pitches per game in 1991, with 105 pitches per start, and both bankrupted their arms doing so.

Their use pattern was a factor in the injuries which forced Greene to miss almost all of 1992 and cast into doubt whether DeJesus will even have a career.

David Cone is another example of a pitcher who suffered from overly ambitious pitch totals. Throwing more than 160 pitches in one start last year with the Mets triggered Cone's 1-4 4.65 ERA slump during August.

After he arrived in Toronto, home of the strong bullpen and the confident starter, however, Cone's arm was asked to do less, and immediately did better.

With the exception of pitchers like Tommy Greene, who received early overuse and consistently high levels of seasonal use, the evidence, as our chart will attest, seems to be that too many pitches per start is more likely to trigger short slumps than permanent trouble.

The chart on the next page presents the 1992 major league leaders in pitches per start, whom you should keep an eye on, if your hopes are resting on their arms.

1992 WORKHORSES (Top 25 major league leaders in pitches per start, minimum 14 starts)

Pitcher	Team	GS	W-L	ERA	P/start
Ben Rivera	Phils	14	7-4	3.07	130
Curt Schilling	Phils	22	14-11	2.35	125
Randy Johnson	Mariners	31	12-14	3.77	121
David Cone	Mets/Jays	34	17-10	2.81	120
Roger Clemens	Red Sox	32	18-11	2.41	119
Mike Gardiner	Red Sox	18	4-10	4.75	119
Jack McDowell	Chi Sox	34	20-10	3.18	117
Kyle Abbot	Phils	19	1-14	5.13	117
Jack Armstrong	Indians	23	6-15	4.64	115
Cal Eldred	Brewers	14	11-2	1.79	114
Melido Perez	Yanks	33	13-16	2.87	112
Bill Wegman	Brewers	35	13-14	3.20	110
Chuck Finley	Angels	31	7-12	3.96	109
Charles Nagy	Indians	33	17-10	2.96	109
Mark Langston	Angels	32	13-14	3.66	108
Frank Viola	Red Sox	35	13-12	3.44	108
Doug Drabek	Pirates	34	15-11	2.73	108
Kevin Brown	Rangers	35	21-11	3.32	107
John Smoltz	Braves	35	15-12	2.85	107
Sid Fernandez	Mets	32	14-11	2.73	107
Mike Mussina	Orioles	32	18-5	2.54	106
Julio Valera	Angels	28	8-11	3.73	106
Greg Maddux	Cubs	35	20-11	2.18	106
Bob Walk	Pirates	19	10-6	3.20	106
Juan Guzman	Blue Jays	28	16-5	2.64	106

It would be wrong and foolish to assume that because a pitcher makes one of our lists, he is doomed to injury. But if a fastball-reliant young ace made both of our lists, you would want to be cautious.

Put it another way. If you were a major league owner, exactly how many millions of dollars would you want to spend on pitchers who have suffered repeated in-game arm abuse, after a history of early arm overwork?

The Braves gave us their answer when they showered dollars and long-term contracts onto Maddux, Glavine and Smoltz.

The Braves might be right to ignore the developing literature on pitcher use and its effect on careers, but if they are wrong, it is a $40 million mistake, the kind that dooms a franchise for decades.

And if they make it, they will wish for a long time that they had read the Wright stuff.

Pete DeCoursey is a baseball columnist for the Reading (PA) Eagle/Times. His address is P. O. Box 220, Paoil, PA 19301

****

PARK EFFECTS RECONSIDERED

by Willie Runquist

The conventional methodology for measuring park effects confounds genuine park effects with differences in overall level of home team performance. Spurious park

effects can be calculated in the absence of any real effect other than an overall advantage to the home team.

Assume that there are two teams, the Sluggers and the Wimps who play each other in identical indoor stadiums. Because of travel, etc., each team scores 5% more runs at home, and 5% fewer on the road. Note that this is not a "park effect", but a "home cooking" effect that exists in both parks.

The Sluggers overall score 600 runs, 315 at Slugger Stadium, and 285 of them at Wimps Field. On the other hand, the Wimps score 400 runs, 210 at home and 190 when they play in Slugger Stadium.

The conventional method for calculating park effects looks at the total number of runs scored by both teams in both parks. In Slugger Stadium the two teams score 315+190=505 runs, and in Wimps Field the two teams score 285+210=495 runs. Via this method, Slugger Stadium appears to be a hitter's park, while Wimps Field appears to be a pitcher's park. We have "created" an apparent park effect when in fact there is none. There is only a general home team advantage and overall differences between the two teams' ability to score runs.

This method is subject to a bias caused by adding proportional effects based on different baselines. This does not mean that there are no park effects, but it does mean that the simplistic use of difference in scoring in different parks might be biased by the overall level of performance of the home team.

Willie Runquist, The Union Bay Oyster Chucker, P.O. Box 289, Union Bay, B.C. V0R-3B0 CANADA

****

STRIKEOUTS, WALKS, & GROUND-TO-AIR RATIOS

by Daniel R. Levitt

We all know that in general pitchers that have high strikeout ratios are more effective than those with low ratios; pitchers that walk few batters are in general more productive than those who walk many. My interest lay in quantifying the above ideas and in examining the interaction between walks, strikeouts, and ground to air out ratio. For example, can a pitcher overcome low strikeouts if he's a groundball pitcher?

The table on the next page lists the average ERA of starting pitchers classified by their ground to air ratio (GAR), strikeout ratio per nine inning game (SO/G), and walk ratio per nine inning game (BB/G). Three different categories (A, B, and C) were defined for each of the these three pitching statistics.

The table allows us to investigate the impact of these variables on a pitcher's ERA, as well as their interaction effects. The table indicates the conventional wisdom is correct: high strikeout pitchers have an ERA 0.68 lower than their low strikeout counterparts (3.56 vs. 4.24), while low walk pitchers have an ERA 1.16 lower than high walk pitchers (3.29 vs. 4.45). As one would guess, walks cause a greater difference than strikeouts.

When looking at the breakdowns by ground to air out ratio, some interesting numbers come to light. Although not very pronounced, high GAR pitchers have a lower ERA than flyball pitchers (3.71 vs. 3.93). A groundball pitcher who is also a low walk pitcher can be successful even with low strikeouts (ERA of 3.17). However, a flyball pitcher with low walks and low strikeouts will probably not be successful (ERA of 4.72).

A groundball pitcher with low strikeouts and high walks exhibits the highest ERA of any of the 27 possible combinations (ERA of 5.40). High walk pitchers show a respectable ERA if they are also flyball, high strikeout pitchers (3.71). The additional effect on ERA by GAR (i.e., beyond the information imparted by strikeout and walk data) is greatest for low strikeout pitchers.

I hope this short article has demonstrated that there are insights than can be gleaned from an analysis of a pitcher's strikeout, walk, and ground to air out ratios. I'd be happy to hear from others who have looked into the same issues.

Daniel R. Levitt, 2516 West 22nd Street, Minneapolis, MN 55405; (612) 377-7275

STARTING PITCHER ANALYSIS
BASED ON WALKS, STRIKEOUTS & GROUND TO AIR OUT RATIO

ERA

GAR: A

@DAVG(	SO/G A	SO/G B	SO/G C	SO/G T
A	4.72	3.26	3.26	3.48
BB/G B	4.20	3.90	3.67	3.82
C	4.68	4.76	3.71	4.24
T	4.49	4.03	3.66	3.93

GAR: B

@DAVG(	SO/G A	SO/G B	SO/G C	SO/G T
A	3.45	3.52	2.91	3.36
BB/G B	4.50	3.75	3.60	3.81
C	4.92	4.40	4.21	4.39
T	4.25	3.89	3.60	3.84

GAR: C

@DAVG(	SO/G A	SO/G B	SO/G C	SO/G T
A	3.17	3.20	2.71	3.06
BB/G B	4.02	3.56	3.53	3.65
C	5.40	5.40	4.36	5.30
T	4.05	3.77	3.26	3.71

GAR: ALL

@DAVG(	SO/G A	SO/G B	SO/G C	SO/G T
A	3.52	3.39	2.88	3.29
BB/G B	4.29	3.73	3.61	3.77
C	5.01	4.63	3.93	4.45
T	4.24	3.90	3.56	3.84

NUMBER

GAR: A

@DCOL	SO/G A	SO/G B	SO/G C	SO/G T
A	2	9	2	13
BB/G B	4	9	14	27
C	4	11	14	29
T	10	29	30	69

GAR: B

@DCOL	SO/G A	SO/G B	SO/G C	SO/G T
A	5	23	9	37
BB/G B	8	23	20	51
C	3	23	10	36
T	16	69	39	124

GAR: C

@DCOL	SO/G A	SO/G B	SO/G C	SO/G T
A	6	10	6	22
BB/G B	5	12	8	25
C	4	5	1	10
T	15	27	15	57

GAR: ALL

@DCOL	SO/G A	SO/G B	SO/G C	SO/G T
A	13	42	17	72
BB/G B	17	44	42	103
C	11	39	25	75
T	41	125	84	250

	BB/G	SO/G	GAR
A	0 – 2.6	0 – 4.3	0 – 1.00
B	2.61 – 3.5	4.31 – 5.9	1.01 – 1.5
C	> 3.5	> 5.9	> 1.5
T	TOTAL	TOTAL	TOTAL

NOTES:

(1) DATA IS FOR ALL STARTING PITCHERS WHO HAD 9 OR MORE STARTS IN 1992 PLUS ANY OF THOSE SAME PITCHERS WHO ALSO HAD 9 OR MORE STARTS IN 1991.

(2) ERAs IN THE ABOVE TABLES ARE SIMPLE AVERAGES (i.e. THEY ARE NOT WEIGHTED BY INNINGS PITCHED).

WHAT TO EXPECT FROM EXPANSION

by Donald A. Coffin

As the National League prepares for expansion, both baseball fans and baseball insiders begin (again) to ask the question, "What can we expect to happen when expansion occurs?" Fortunately, we now have some evidence we can use to answer that question. We have observed three periods of expansion in major league baseball--1960/61, 1969, and 1977--and we know what happened then. Whether these things will happen again we do not know, but the evidence remains our best clue.

We can easily describe what happened--runs per game, home runs per game, batting average, and extra-base power (slugging average minus batting average) all increased slightly immediately following expansion (see Table 1 on the next page), regardless of whether expansion was in one league or in both. In short, expansion has generated greater offense. However, these increases in offense have not persisted. By three years following expansion (1964, 1972, 1980), runs per game always falls, as does extra-base power. Home runs per game have increased only once (in the AL, from 1961 to 1964, 1.89 to 1.91); batting average also increased only once (also in the AL, from 1977 to 1980, .266 to .269).

The league leaders in batting average, on-base average, slugging average, and home runs generally also improved their numbers (see Table 2). Except in the National League from 1976 to 1977, the batting average of the five league leaders improved with expansion (the average for the three expansions, for both leagues, is .0125). The leading slugging averages rose every time (an average increase of .067). The leading on-base averages also rose every time (an average increase of .026). These increases were uniformly larger than the increases for the leagues as a whole, suggesting that the spread between the best and worst performances in the majors also got larger with expansion.

But, again, after three years, these gains in batting performance did not persist. The league-leading batting averages fell in the three years following expansion by an average of .0095--giving up about 75% of the gains associated with expansion. League-leading slugging averages fell by an average of .037 (55% of the expansion gains). League-leading on-base averages fell by an average of .021 (82% of the average gain). The losses among the league-leaders were, again, larger than the losses for the leagues as a whole. This suggests that the spread between the best and worst performances also got smaller as major league baseball adjusted to expansion.

The five league leaders in home runs in each league also benefited from expansion. The average gain for the five leaders in each league was 8.5 home runs--up from an average of 33.1 home runs to an average of 41.6. The fall-off again occurred, with the five leaders in each league losing 6.5 home runs, falling back to 35.1 after the adjustments to expansion.

The only exception to declines in offense occurred in the American League between 1977 and 1980--batting averages, slugging averages, and home runs for the league leaders continued to rise (and on-base average did not fall). Indeed, the increases among the leaders in the AL suggest that the spread between the best and worst offensive players continued between 1977 and 1980.

What is interesting here is that the effects, in general, occurred in both leagues, even if expansion occurred only in one (the AL in 1961--one year before the NL expansion in 1962--and in 1977).

Don Coffin, 3400 Broadway, Indiana University NW, Gary, IN 46408; (219) 980-6630

Table 1: Runs per Game, Home Runs Per Game, Batting Average, Extra Base Power, and Expansion.

	National League				American League			
Year	RPG	HRPG	BA	EBP	RPG	HRPG	BA	EBP
1960	8.48	1.68	.255	.133	8.77	1.76	.255	.132
1961	9.05	1.93	.262	.143	9.05	1.89	.259	.141
1964	8.03	1.49	.254	.120	8.12	1.91	.247	.135
1968	6.86	1.10	.243	.098	6.81	1.36	.246	.123
1969	8.11	1.51	.250	.119	8.18	1.69	.250	.129
1972	7.81	1.46	.248	.116	6.93	1.26	.239	.104
1976	7.96	1.15	.255	.106	8.02	1.16	.256	.105
1977	8.80	1.68	.262	.132	9.06	1.78	.266	.139
1980	8.07	1.28	.259	.115	9.01	1.63	.269	.124

Source: Calculated from data in John Thorn and Pete Palmer, ***Total Baseball***, Revised Edition, Warner Books (New York, NY: 1991).

Table 2: League Leaders in Selected Batting Categories (Average of Each League's Five leaders)

	National League				American League			
Year	BA	OBA	SA	HR	BA	OBA	SA	HR
1960	.317	.401	.566	37	.313	.405	.547	37
1961	.336	.410	.596	39	.325	.440	.644	50
1964	.327	.395	.566	35	.311	.401	.571	38
1968	.319	.381	.510	32	.290	.390	.517	34
1969	.337	.436	.587	39	.312	.413	.583	45
1972	.322	.400	.551	36	.310	.399	.535	29
1976	.330	.413	.529	33	.327	.397	.480	27
1977	.323	.416	.583	41	.341	.427	.556	36
1980	.316	.392	.532	35	.349	.427	.574	37

Source: Calculated from data in John Thorn and Pete Palmer, ***Total Baseball***, Revised Edition, Warner Books (New York, NY: 1991).

A TRUE MEASURE OF OFFENSIVE PERFORMANCE

by William C. Gilbert

Barry Codell's Diamond Weight, described in the February, 1993 issue of *By the Numbers*, is clearly a better measure of offensive performance than On Base Average + Slugging Average. However, there is another measure that is more precise in determining the offensive value of a player's performance. This measure is Bases per Plate Appearance (BPA).

The formula for calculating Bases per Plate Appearance is (TB + BB + HP + SB - CS - GIDP) / (AB + BB + HP + SF). This measure is similar in most respects to the Diamond Weight calculation, [(Reached Base + Total Bases) / Total Plate Appearances], in that it includes the basic elements of On Base Average and Slugging Average related to Plate Appearances. The denominators are identical except that BPA does not include sacrifice hits in plate appearances because the batter is sacrificing the opportunity to gain a base in that plate appearance.

However, there are two significant differences in the numerator. The most obvious is that BPA gives credit for baserunning by including stolen bases reduced by caught stealing and grounding into double plays. The less obvious, but more significant difference, is that Diamond Weight gives double value to hits since hits are included in both the Reached Base (RB) calculation and the Total Bases (TB) calculation. This tends to inflate the Diamond Weights for players like Kirby Puckett and Ken Griffey Jr. who get a lot of hits but relatively few walks and penalizes players like Mark McGwire and Darren Daulton whose hit totals are a lesser part of their total offensive production.

Following is a list of the top 10 major league players in 1992 in Diamond Weight and Bases per Plate Appearance (minimum of 502 plate appearances):

Diamond Weight (DW)	
1. B. Bonds	.938
2. G. Sheffield	.908
3. E. Martinez	.889
4. F. Thomas	.871
5. M. McGwire	.863
6. F. McGriff	.861
7. K. Griffey, Jr.	.851
8. A. Van Slyke	.834
9. R. Sandberg	.825
10. K. Puckett	.823

Bases per Plate Appearance (BPA)	
1. B. Bonds	.734
2. M. McGwire	.625
3. D. Daulton	.605
4. F. McGriff	.601
5. F. Thomas	.588
6. G. Sheffield	.578
7. E. Martinez	.575
8. D. Tartabull	.574
9. B. Anderson	.571
10. R. Lankford	.554

Several observations can be made from a comparison of the two lists. Barry Bonds was the clear leader by both measures. However, his advantage over the rest of the field is significantly greater in BPA. Edgar Martinez is the top AL player in DW, whereas Mark McGwire is the top AL player in BPA. The last four players on the DW top ten list failed to make the BPA top ten list. Ken Griffey (.553 BPA), Andy Van Slyke (.543), Ryne Sandberg (.552) and Kirby Puckett (.512) were replaced by Darren Daulton (.819 DW), Brady Anderson (.742), Danny Tartabull (.800) and Ray Lankford (.790). Both lists include 5 AL players and 5 NL players.

The biggest shifts among players who appear on both lists involve Gary Sheffield and Mark McGwire. Sheffield, with 184 hits and only 48 walks, is helped much more by double counting hits in the DW calculation than is McGwire with 125 hits and 90 walks.

Three other players who failed to accumulate 502 plate appearances in 1992 achieved a BPA rate which would have put them in the top ten. However, none would be in the top ten in DW: Rickey Henderson (.788 DW, .628 BPA), Rob Deer (.817 DW, .587 BPA), and Deon Sanders (.806 DW, .562 BPA). These three players scored relatively high in BPA because of unique aspects of their offensive games. However, they did not fare as well in DW because these features were diluted when hits were counted twice in calculating DW.

A limitation to both of these measures is that they require information which does not appear regularly in Baseball Weekly or USA Today's weekly averages (HP, GIDP, SH and SF). Following is a simplified version of BPA which can be calculated using data readily available, and the top 10 players in 1992 according to SBPA = (TB + BB + SB - CS) / (AB + BB):

		SBPA	BPA
1.	B. Bonds	.755	.734
2.	M. McGwire	.650	.625
3.	F. McGriff	.627	.601
4.	F. Thomas	.622	.588
5.	D. Daulton	.613	.605
6.	G. Sheffield	.612	.578
7.	E. Martinez	.603	.575
8.	D. Tartabull	.590	.574
9.	K. Griffey, Jr.	.576	.553
10.	B. Anderson	.576	.571

The difference between SBPA and BPA is typically 20-25 points, depending primarily on the number of GIDPs. In Anderson's case, the difference was just 5 points because he grounded into only 2 double plays in 1992. In the cases of Thomas and Sheffield, the difference was 34 points because they each grounded into 19 double plays.

Note that Barry Bonds is on target to top the list again in 1993 by an even wider margin. In early May, Bonds' SBPA is .961 - more than .300 higher than anyone else achieved in 1992.

Bill Gilbert, 5113 Glen Haven Dr., Baytown, TX 77521

****

THE HIDDEN IMPACT OF THE DH RULE

by Rob Wood

The two most significant changes in major league baseball in the last twenty years both occurred about 20 years ago: the advent of free agency and the adoption by the A.L. of the designated hitter. In this article, I wish to explore the possible interrelationship and consequences of these two historical partners.

I have been interested in this topic for many years, but only recently begun to explore the impact it has on today's game. Player movement exploded this off-season with abatement nowhere in sight. One player move in particular caught my eye and prompted this mini-study.

Andre Dawson and his aching knees have moved from the N.L.'s Chicago Cubs to the A.L.'s Boston Red Sox. Part of the reasoning was surely the opportunity Boston has to use Dawson as its DH.

There is a long and storied history of aging stars migrating from the N.L. to the comforts of the A.L. and its DH role. Notables defining this category include Hank Aaron, Frank Robinson, Orlando Cepeda, Rico Carty, Billy Williams, Rusty Staub, Richie Zisk, Greg Luzinski, and Dave Kingman. The rationale is obvious. As these stars' defensive abilities waned, their bats remained valuable weapons. The DH was made for such a combination.

Indeed, there is now a preponderance of stars found in the A.L. at the expense of the N.L. Define a "star" hitter as one who has performed at a high level for many years. An operational definition would classify a hitter based upon his career totals of hits and/or home runs.

The following numbers demonstrates that there are far more star hitters in the A.L. than in the N.L. Of the 56 active players entering the 1993 season with either 1500 hits or 150 home runs, fully 38 (68%) are in the A.L. [An interesting article could explore the connection between this observation and the perception that the A.L. has overtaken the N.L. and is now the stronger league.]

More telling evidence is gleaned when we look at the breakdowns by career totals. The greater the cutoff in hits or homers, the greater the fraction of active players who exceed the cutoff are in the A.L. Is this coincidence? Are there explanations other than a "DH migration" hypothesis?

Active Players by Career Home Runs*

	AL	NL	Total
300+	4	3	7
200-299	12	5	17
150-199	12	5	17
100-149	10	10	20

* entering the 1993 season

Active Players by Career Hits*

	AL	NL	Total
2000+	11	4	15
1500-1999	15	7	22
1000-1499	14	12	26

* entering the 1993 season

Sure, some other things may be playing a role. A.L. ballparks are smaller on average than N.L. ballparks, so more home runs would be expected to be hit in the A.L. In addition, the DH rule itself allows more hitters to bat in the A.L. than in the N.L. These two explanations, it seems to me, are not sufficient to explain the almost 70/30 split in the above two classifications (back of the envelope calculations lead me to estimate that they could at most account for a 57/43 split). In any event, it is indisputable that many aging N.L. stars spend their final years DH'ing in the A.L.

Of course, there is an obscured aspect of DH migration that should not go unnoticed. Players who come up in the A.L. often remain in the A.L. throughout their career, shifting to the DH as their defensive abilities decline. These players sometimes change teams, but rarely change leagues. Active participants in this group include George Brett, Carlton Fisk, Harold Baines, Jose Canseco, Paul Molitor and others.

Thus, it would not be fair to argue that the "DH migration" phenomenon cannot fully explain the disparity since many A.L. stars have played their entire career in the A.L. Indeed, of the 58 players in my study, 33 have played their entire careers in one league with the split here being exactly two-to-one, 22 in A.L. and 11 in N.L.

Many before me have bemoaned the "1950's style" of baseball (e.g., no speed, wait for the three run homer) currently played in the A.L. while pointing to the N.L. as the "exciting" league with better fielding and more base running. I do not wish to contribute to that debate other than to accord the DH rule much of the credit/blame insofar as the premise is valid.

I don't know how much I buy the above argument, for it flies in the face of one of my pet beliefs. The N.L. dominated the A.L. for much of the 1960's and 1970's, attributable to its greater willingness to sign blacks and Hispanics. However, slowly but surely, the dominance has dissipated as players and coaches have intermingled over the years.

One of the most cited cases occurred in 1973, the very first year of the DH. Hal McRae was traded from the Reds to the Royals. McRae brought with him the winning style of play that made the Reds a powerhouse (and soon back to back World Champions). In the next 13 years, McRae led the Royals to 6 division championships and six other second place finishes. [Note that an earlier and perhaps better case may be made for Frank Robinson bringing the N.L. style to the Orioles with him in 1966.]

What conclusions do I draw? I'm not sure, but I hope the numbers that I have cited are not permanent reflections of the makeup of the two leagues. I believe in the sanctity of both major leagues, and have even learned to accept the DH. However, I adamantly refuse to accept the fact that a great many of the blossoming stars of the N.L. will migrate to the greener pastures of the DH-infested A.L. In some ways, the N.L. is becoming a testing ground (a "minor league"?) for the A.L.

Pundits are often heard to mutter that today's stars do not match up with the stars of the previous generations. Where, for example, are the next Willie Mays, Hank Aaron, or Frank Robinson? This type of reasoning (I am being kind) has more than an element of sophistry. It is not fair to compare the career totals of retired stars (after all, these three retired with 2001 home runs!) with the totals of today's stars.

Who knows how many home runs Jose Canseco or Mark McGwire will eventually hit? Today's game has a lot to offer. I stick up for it every chance I get. However, I do not want to see the day when virtually every great, proven offensive star plays in the A.L.

Rob Wood, 2101 California St., #224, Mountain View, CA 94040; home: (415) 961-6574; work: (415) 854-7101

****

MEASURING THE VARIABILITY IN PERFORMANCE

by Willie Runquist

Variability refers to differences in performance from one time to another. Obviously, variability (or its inverse, consistency) is an important aspect of baseball performance. However, despite its importance, the problem has to my knowledge been examined only once. In the *1992 Elias Baseball Analyst*, variability scores were given for several major league players as part of an article about Darryl Strawberry and the Los Angeles Dodgers.

Our primary purpose is to provide a procedure for measuring performance consistency. Consistency may mean different things in different contexts, and therefore

is dealt with in different ways. We will try to point out in each case exactly what is being measured.

In this first of a series of articles, we will examine the season by season variability in batting performance. We do not wish to become embroiled in the issue of how batting performance is best measured. Data will be presented for three measures which are generally poorly correlated and hence may be considered to measure different aspects of total performance. Those measures are batting average, isolated power (the ratio of extra bases to at bats), and run production (the ratio of the sum of runs scored and runs batted in to at bats).

Essentially, the problem may be exemplified as follows: in his six years with the California Angels, Wally Joyner had an overall batting average of .288 with seasonal values of .290, .285, .295, .282, .268, and .301. Variability is concerned with the differences among those seasonal values. The usual measure of variability among a set of values is the variance, or its square root, the standard deviation. These measures are based on the squared deviation of each value from the overall mean. Instead, however, we will use the average (mean) of the absolute deviations from the overall value, which we call V*. There are several advantages to this statistic, not the least of which is that the measure is easily computed and is directly and intuitively interpreted.

In the above example, the absolute differences between Joyner's seasonal batting averages and the overall average are easily seen to be .002, .003, .007, .006, .020, and .013. The mean of these values is .0085. It is usually convenient to multiply values by 1000 to eliminate decimals. Therefore, V* is equal to 8.5. On the average, Joyner deviated about 8 points from his overall batting average per year.

There were 38 major league players that had over 500 at bats in four of the five seasons from 1987 to 1991, and at least 400 at bats in the other season. We computed V* for each of those players for the five seasons of batting averages, isolated power ratios, and run production ratios.

Brett Butler has the most consistent batting average in the sample with a V* of 7, and Robin Yount is the most inconsistent with a V* of 28. For isolated power, Ozzie Guillen is the model of consistency with a V* of 6, and Mike Greenwell is the most inconsistent with a V* of 45. For run production, Joe Carter is the most consistent with a V* of 9, and Barry Bonds is the most inconsistent with a V* of 66.

Using only five seasons, one extreme value can have a profound effect on V*. A large value of V* may mean that the player had one highly unusual year amidst many years of consistent performance. That year may have been a "career season" such as Cal Ripken's .323 in 1991, or an "off year" such as Eddie Murray's .247 in 1989. Generally speaking, large values of V* are associated with players who had one discrepant season. For batting average the correlation between a player's largest discrepancy and V* was .82. Those discrepant seasons may result from an injury, from changing something about the game such as adding a rabbit to the ball, from a trade to a more or less friendly ball park, a change in the batting order, or even a prolonged run of good or bad luck.

V* is simply a descriptive measure of variability in performance, and does not in itself praise or slander the character of the player. At first glance there were some surprises in the rankings of these 38 players by V*. For example, Wade Boggs (who has one of the highest batting average V* at 21) has generally been considered to be one of the most consistent hitters of the modern era -- before 1992 anyway. Unfortunately, the term "consistency" is often used to refer to the overall quality of a player's performance. In interpreting V* you have to remember that a player who has a career average of .350 and hits .320 is just as variable as a .300 player that hits .270 or a .230 player who hits .260. Boggs had in fact been highly variable, but that variability was within a high overall level of performance.

There does not appear to be any relationship between V* and overall performance in the case of batting average (correlation of .04). For isolated power and run production, players with better overall performance tend to be more variable. The correlation between overall isolated power and its V* was .42, while for run production the correlation was .29. These indicate some association, but there are numerous exceptions. There is probably nothing very profound in these correlations since players who do not hit for many extra bases have less room to vary from year to year.

It is also important to note the differences in the rankings for the different performance measures. The rankings for batting average are unrelated to either those for isolated power or run production (r=.15 and .17).

The rankings for isolated power and run production are moderately correlated (r=.46). In other words, there is a tendency for players who are variable in power to also be variable in run production, but again there are many exceptions. This relationship should not be surprising, in view of the conventional wisdom that power hitters drive in more runs.

Since we used the same five years for each player, any systematic changes in playing conditions are the same for all players. This fact allows us to examine the deviation values and compute V* for the different years. For batting average, there was little difference between the years (16.6, 13.3, 16.1, 18.6, 15.9, respectively). For isolated power, 1987 produced a V* of 35.7 while the remaining years were 18.1, 23.4, 21.3, and 24.0. As 1987 was widely renowned as the "Year of the Homer", this difference should be expected. Sixteen of the players had their most deviant year in 1987; no other year affected more than six players.

For some players, the five years in question represent the first few years of their career. At the other extreme were players like Brett and Winfield who began their careers in the early 1970s. This fact allows us to examine the relationship between variability and experience by simply correlating V* with major league experience. Of the three statistics, only variability in batting average was related to experience. The correlation was only .28 which is high enough to be significant, but provides little predictability. The relationship, however, was not linear, so the correlation coefficient does not give a true indication of the extent to which the variables are associated.

The least variable players were those in the mid range of experience. The mean V* for players with 6-8 years of experience was 15.8, for those with 9-11 years it was 14.4, for those with 12-14 years it was 14.6, and for players with over 14 years of experience it was 19.8. The correlation ratio, which provides a more accurate indication of non-linear correlation, was .62. This verifies our belief that players are more variable at the beginning and end of their careers.

Similar results are obtained when age instead of major league experience is used as the covariate. Our use of the term experience does not imply causal status. Age and experience are inextricably confounded. In fact, with what we know about the aging process, it is much more likely that it is changes associated with advancing age that result in more variable performance. Just why it should affect batting average and not power hitting is not clear.

V* may in principle be used with any statistic. However, problems may arise when comparing players with widely differing numbers of at bats. Random variation will produce larger deviations with small samples of at bats, so that a player with few at bats is likely to have a larger value of V*. In such cases, we have found that dividing the deviation by the standard error of the statistic before computing V* will compensate for this effect. Essentially, it converts the deviation value to standard error units, so that a given deviation is considered more "important" when it is based on more at bats. This transformation, however, may only be used when a value may be assigned to each at bat. Statistics that consist of proportions (batting average, on-base average) or statistics that consist of true averages (slugging average, isolated power) are therefore suitable for this analysis, while other ratios (total average, runs created, production index) are not because there is no procedure available for estimating the standard error.

Willie Runquist, The Union Bay Oyster Chucker, P.O. Box 289, Union Bay, B.C. V0R-3B0 CANADA

****

THE STATISTICAL DEMISE OF THE "MOSS KLEIN" PITCHER

by Richard D. Adams

Clarifying the Quality Start Statistic: A "Quality Start" is defined as a starting pitcher going six innings and yielding three or fewer earned runs. This is a "silver lining" statistic. Rather than implying that a start was a quality performance, it implies that the start was not a garbage performance. After all, would you rather lead the league with the highest percentage of quality starts or with the lowest percentage of garbage starts?

Defining the "Moss Klein" Pitcher: A "Moss Klein" pitcher [named after a sportswriter who disparaged the quality start statistic] is defined as a starting pitcher who pitches six innings, gives up exactly three runs, and gets credit for a quality start. The purpose of this hypothetical pitcher is to cast aspersions upon the effectiveness of the

"Quality Start" statistic. The argument is that at the end of a season, a "Moss Klein" pitcher, with an admittedly unimpressive 4.50 ERA, would have an impressive 38 to 42 quality starts.

Adams' First Three Laws of Statistical Analysis: Before addressing this hypothetical pitcher, let me introduce you to Adams' first three laws of statistical analysis: (1) All non-uniform distributions can be expected to have a tail; (2) Fairy tales can come true in the tails of a distribution; and (3) Examples which are derived from the tails of a distribution and are not accompanied by an alternative theory, explanation, or metric should not be given credibility.

Analyzing the "Moss Klein" Pitcher: The following analysis is based on data from all major league games for the last nine years (1984-1992). Tied games and games of less than six innings have been ignored. The available data did not distinguish between earned runs and unearned runs. Therefore, the analysis is based on actual runs scored in the first six innings and Runs Allowed Average (RAA) is used as a surrogate for Earned Run Average (ERA). The detailed data is presented in Table 1 on the next page.

During the years 1984-1992: when a pitcher gave up exactly 3 runs in the first six innings, the pitcher's team won only 46% of the time. That's the bad news. Now for the good news. Just how bad is giving up 3 runs in the first six innings and what is the probability of doing so in each of 40 starts?

The mean number of runs scored in the first six innings of a major league game during this period was 2.95 with a standard deviation of 2.522. Thus, the 3 runs allowed by the "Moss Klein" pitcher is not significantly different from the mean scored. Next, the probability of scoring exactly three runs in the first six innings is .1466 (5645 + 37,826).

For a pitcher with 40 starts, the binomial probability of having more than 11 "Moss Klein" starts is statistically significant at the $\alpha = .025$ level. 14 or more such starts is statistically significant at the $\alpha = .001$ level.

Thus, the "Moss Klein" pitcher is an average pitcher with statistics taken from the "extreme" tail of the distribution. I would define this as a scenario so improbable that it is a perfect fit for Adams' third law of statistical analysis.

Additional Observations about the Quality Start Statistic: While we're here, let's consider the difference between quality starts and garbage starts over the last nine years.

1) 85% of all games won were quality starts (that means only 15% of all games won were garbage starts);

2) 46.5% of all games lost were quality starts (that means 53.5% of all games lost were garbage starts);

3) Teams having quality starts won 64.6% of their games (that translates into 104 wins in 162 games, a high probability pennant winning clip);

4) Teams having garbage starts won only 21.9% of their games (that's worse than the 1916 Philadelphia Athletics);

5) Pitchers having quality starts had a 2.21 RAA while pitchers having garbage starts had a 8.67 RAA;

6) A chi-square test of independence for Table 2 resulted in the following statistic: $X^2 = 817.8$ (df = 1) and $p < .00001$. Thus, the classification of the cells of the contingency table are not independent of each other (this translates directly into a statistically significant relationship between quality starts and winning baseball games at the $\alpha = .00001$ level).

Dick Adams, University of Baltimore, 1420 N. Charles Street, Baltimore, MD 21201-5779; home: (410) 465-6362, work: (410) 625-3310

TABLE 1:

1984-92 RUNS SCORED IN NON-TIED GAMES AT END OF SIX INNINGS							
Opponent Runs	0	1	2	3	4	5	6
Games won	4790 .253	4839 .509	3894 .715	2549 .850	1411 .924	749 .964	681 1.000
Games lost	1051 .056	2064 .165	2683 .307	2995 .465	2872 .617	2371 .742	4877 1.000
WL probability	0.820	0.701	0.592	0.460	0.329	0.240	0.123
WLP (runs or less)	0.820	0.756	0.700	0.646	0.600	0.565	0.500
WLP (runs or more)	0.500	0.442	0.370	0.291	0.219	0.165	0.123
Runs Allowed Ave	0.000	1.500	3.000	4.500	6.000	7.500	11.386
RAA (runs or less)	0.000	0.813	1.557	2.213	2.770	3.227	4.426
RAA (runs or more)	4.426	5.234	6.262	7.421	8.671	9.989	11.386

TABLE 2: QUALITY START CONTINGENCY TABLE

		WINNING TEAM HAD A QUALITY START					
		ACTUAL GAME COUNTS			PERCENTAGE OF TOTAL		
		YES	NO	TOTAL	YES	NO	TOTAL
LOSING TEAM HAD QUALITY START	YES	8173	620	8793	.4321	.0328	.4649
	NO	7879	2221	10120	.4186	.1174	.5351
	TOTAL	16072	2841	18913	.8498	.1502	1.0000