

BUCKING THE SYSTEM
OR, WHY THE NFL CAN'T FIND HAPPINESS WITH ITS PASSER RATINGS

By Bob Carroll

If you believe in your heart of hearts that Warren Moon is a better passer than Otto Graham, you're at one with the National Football League. Never mind that Graham is a card-carrying member of the Pro Football Hall of Fame and a quarterback who led the Cleveland Browns to seven league championships in ten seasons, while Moon is the oft-booped signal-caller for one of the NFL's least successful franchises.

According to the National Football League's Passer Rating System, Moon tossed for a 68.5 mark last season; Graham, in 1950 – a year his Cleveland Browns won the NFL Championship, could manage only a paltry 64.7. That makes it official; Warren is 3.8 better than "Automatic Otto."

Has George Orwell become an NFL flack? Is this reality or newspeak?

More! In the gospel according to the NFL, Dan Marino is the best passer ever. Until this year, Joe Montana was. A couple of other top ten performers: Danny White, the guy who made Dallas forget Roger Staubach, and Neil Lomax, whose success in St. Louis has made him a legend.

And it don't rain in Indianapolis in the summertime.

Well, it all depends, you say. Actually, it DOESN'T rain (or snow) inside the Hoosier Dome during any part of the calendar year, and Marino, Montana, White, and Lomax ARE good – maybe great – passers.

But, are they THAT good? The much-maligned NFL Way of Rating Passers places some present throwers at the top of the Hurler Heap and consigns such clutzes as Sid Luckman, Johnny Unitas, Y.A. Tittle, Joe Namath, and Bobby Layne to also-ran status.

If you only recently moved from lurching on pabulum to watching pro football, you might be inclined to believe that this is the true order of things. Children tend to doubt the world existed before they did.

But if you've been aware of the NFL for long enough to have seen any of the above-listed graybeards in action, you may be beating the walls of your cell over the injustice of it all. You've probably decided that today's NFL is a sinister plot by Pete Rozelle to discredit the heroes of your youth. The doctors prescribe warm baths and your children call you paranoid.

But are you?

To quote an early fan: "The fault lies not in our stars but in our rating system. "

There are folks out there who'd rather see the NFL's PRS canceled than see their taxes cut. They believe deep down that the passing of the passer rating system would bring universal peace and brotherhood, economic stability, and their favorite NFL quarterback to the top of the career list.

It seems that no one has much good to say about the NFL's way of ranking its passers, and that's too bad because the present system is certainly a "long bomb" ahead of every other way they tried to do it.

THE LESSONS OF HISTORY

For the first dozen years that the NFL existed, from 1920 through 1931, the league ignored the idea of statistics altogether. There weren't that many people showing up for games, and the league's leaders probably figured that no one cared about League Leaders.

THE COFFIN CORNER: Vol. 8, No. 9 (1986)

Then, in 1932, the NFL decided to start counting up its players' accomplishments. The decision to keep track of the numbers seems to have had little to do with Ultimate Truth, Revealed Verisimilitude, or the Reality of Sweat. It was just that the NFL decided they could sell a few more tickets if they could say that one of the players on the visiting team was the leading something-or-other.

The revelation that statistics can sell tickets is one of the great discoveries of the Twentieth Century, ranking just behind pop-top cans and ahead of penicillin. However, a codicil to the Law of Capitalistic Stats is that your numbers have to be believable. In other words, your stats have to somehow reflect the reality the fan feels in his gut before he'll ante up the price of admission.

Rushing, receiving, and scoring were easy: total yards, total passes caught, and total points scored. Those things reflected what fans saw and read. If the league said Beattie Feathers was the top rusher, fans could remember that he made a lot of long runs. Bronko Nagurski wasn't the leading rusher, but that was okay. Fans could believe that all his short gains didn't add up to Feathers' total. The stats reflected the perceived reality. Fans bought tickets to watch Feathers run – not very many tickets in those days, but a few. Enough to keep adding up rushing stats.

But passing was another kettle of fish. When a player hurled a pass, four different things might happen. The ball could bounce off the turf for an incompleteness, and that was bad. It could be caught for some yards, and that was good. It could be caught for a touchdown, and that was terrific. Or it could be intercepted by an opponent, and that was terrible. Fans tended to remember their passers by the number of terrific and terrible things they did. From the start, the perception of passers was more subjective than that of any other major stat.

From 1932 through 1937, the NFL ignored all the philosophical implications and simply added up the yards as though passing were just another kind of rushing. Anyone reading the stat lists was probably more than a little upset by that. After all, a lot of passers were throwing a lot of passes into the ground – or to opponents – and still ending up near the top of the league charts by completing a few long ones. Anyone throwing enough passes could eventually lead the league.

Green Bay's Arnie Herber threw often and deep. He was proud of his ability to hurl a football farther than some people could walk without packing a lunch. One time a couple of teammates bet him he couldn't throw the ball the length of the field if they allowed him the roll. Arnie chuckled and took the bet. Standing on one goal line, he unloaded toward the other. The ball arched high over midfield and came down at about the fifteen-yard-line. To Arnie's chagrin, the angle of fall was so great that the ball hit and bounced back toward him, just as his teammates figured it would.

Although he lost the bet, Arnie won three NFL passing championships between 1932 and 1936. But he seldom completed more than a third of his tosses. Nevertheless, when he connected, it was usually for a ton. Of course, having receivers who could run like Don Hutson and Johnny Blood didn't hurt his yardage any.

Nevertheless, long-throwing Arnie was generally perceived as the best passer around. His passing titles mirrored a reality that most fans could accept.

Then, in 1937, Sammy Baugh came to the NFL. He showed he could throw both long and short. And he could usually complete about half of his throws, no matter what length. Accuracy became the name of the game. Hence, the famous story where the coach says, "When the receiver breaks, Sam, hit him in the eye with the ball." And Baugh replies, "Which eye?"

Pass reality had changed.

So from 1938 through 1940, they determined the champ by completion percentage. And, after a couple of injury-plagued seasons, Baugh won the passing title in 1940. He was exceptional, but some of the league's best passers were getting hurt by this method of ranking. Now the shoe was on the other foot, so to speak. The advantage was with the occasional passer who had surprise on his side. The heavy-duty passers – what few there were in those years – suffered because, as their attempts increased, their completion percentage tended to decrease.

THE COFFIN CORNER: Vol. 8, No. 9 (1986)

Once more change was in order. From 1941 through 1948, the league went for a more sophisticated method by using an inverse ranking system in which the best mark in a category got one point, the second best mark two points, and so on. Once all the points were added up, the winner was the man with the lowest total. They tried to reward both the heavy-duty tossers and the low attempt-high efficiency throwers with their mixture of categories. For the busy throwers, they used total completions, yards gained, touchdowns, and interceptions. For the occasional hurler, percentages of completions and of interceptions. This went pretty well. People like Sammy Baugh, Cecil Isbell, and Bob Waterfield ended up at the top of the list. However, there were disadvantages. In 1949, they dropped number of interceptions from the list, as a category favoring the less active passers.

The biggest problem for the inverse rating system of the '40s was that no one could figure it out unless he had all the stats from all the passers. Even then, it took hours in those non-calculator days to compute who ranked where and add everything up. The average fan was lost. Moreover, by mixing totals and percentages, there was always the chance that some second or third team thrower would sneak into the top of the rankings. Finally, there was no logical way to figure a career mark. The NFL longed for a simpler system.

They thought they'd found it in 1950. They established the idea of a minimum number of attempts to qualify for the championship (100) and ranked everyone qualifying by average yards gained per pass. So, in that first year, Joe Geri who had generally awful stats but a high average gain finished third in the league. Nevertheless, the NFL stayed doggedly with this system until 1959, but six of the ten league passing champs threw fewer than 200 passes.

They knew something was wrong in 1957 when Tommy O'Connell, the Cleveland quarterback, finished on top. Tommy had taken over as Cleveland QB after Otto Graham retired, but no one had any trouble telling them apart. Paul Brown, the legendary Cleveland coach, was critical as all get out about Tommy's inadequacies and generally gave the impression that O'Connell would stay on as his quarterback only as long as a no one else applied. Reading the writing on the wall better than he ever read a defense, O'Connell, the "best" passer in the league, retired before the next season began.

Meanwhile, quarterbacks like Johnny Unitas and Bobby Layne won championship games, all-pro honors, but nary a passing title. Clearly, the best quarterbacks in the league were being sacked by the system.

Back to the drawing board. In 1960 and 1961, the NFL returned to the inverse ranking system, using six categories: total completions, total yards, total TD passes, percentage of completions, percentage of interceptions, and average gain per attempt. They also insisted that the qualifiers had to average ten attempts per game – 120 total in 1960, 140 in 1961.

This was fine by Milt Plum, the Clevelander who succeeded Tommy O'Connell at QB. The new system melded so well with his limited skills that he finished both years at the top of the standings. Plum's primary pass ability was to dump short, safe screens to Jim Brown and then watch the Hall of Fame fullback barrel for long gains. It made for great stats, but you either had to believe that Plum was the best passer in the world or that the system was short-sheeting the likes of Johnny Unitas, Y.A. Tittle, and Bart Starr.

In 1962, the league cut the categories down to four: percentage of completions, total touchdown passes, percentage of interceptions, and average gain per attempt. Despite some rumblings about including one total (touchdowns) in with three percentages, this system held through 1971. In 1972, they substituted percentage of touchdowns for total touchdowns. People like Tittle, Starr, and Sonny Jurgensen did well in this system. They should have; just about everyone agreed they were the best.

Unfortunately, there were still some real clinkers in this system. First of all, you still had to have every qualifiers' numbers to figure ONE qualifiers' rank. Worse, you couldn't really see what a passer was doing from year-to-year. A passer might finish on top one season, turn around and the next year finish fifth with the same kind of year because he was ranked against every other fulltime passer in the league. It looked like he slipped when in truth four others improved.

TO THE RESCUE RIDES THE SYSTEM

Ah, but relief was in sight! For three years, a special study committee had been looking into ways of solving the problems. This group, headed by Don R. Smith of the Pro Football Hall of Fame, Seymour Siwoff of the Elias Sports Bureau, and Don Weiss, NFL Public Relations Director, and including Jan Van Duser, Curt Mosher, Tom Grimes, Bill McGrane and Jack Horrigan spent three years studying the stats of all qualifying passers back to 1960. Eventually, they settled on a system that was Smith's brainchild.

First, they kept the Big Four categories: percentage of completions, percentage of touchdowns, percentage of interceptions, and average gain per pass attempt. But instead of ranking a passer's marks against every other passer, the wrinkle that was at the root of what looked like the Last of the Big Problems, they ranked him against a set scale.

To try to envision this, imagine yourself a school teacher with two classes of twenty students each. You give each class the same test.

One class is graded against itself – the old NFL system. The kids with the highest marks get "A's", the kids with the lowest marks get "F's", and the ones in the middle get "C's". It doesn't matter what the marks are; if the whole class is full of dunderheads and the highest mark is only a 36, that student will still get an "A". If, on the other hand, the class is made up of geniuses and the LOWEST mark is 96, that genius will fail! Boy! are the mommies going to scream! School boards being what they are – chicken – expect to be looking for another job in June.

You grade the second class against a previously determined scale. Let's say, A = 90 or better, B = 80 or better, etc. Now, imagine you've been teaching for twenty years and have given the same test every year. You've discovered that your best students always score above 90 and your worst can't make it to 60. All things considered, you'd have to say that this was the better way to rate the members of the class. It has fairness and continuity. Best of all, if your students can't pass the test, the test is wrong! Change the test and you get rehired.

This is what the present NFL system does. It rates its passers on a previously determined scale. And, if the passers aren't "passing" their test, the NFL changes the rules – er – test. More of that later; right now let's look at what the system-to-end-all-systems – the final option that first saw the light of day in 1973.

By studying thirteen years of qualifying passers, the committee determined that an average passer completed 50.0 percent of his passes, threw five touchdowns for every hundred passes, had eleven out of every 200 tosses intercepted, and gained seven yards per attempt on average. They awarded zero points for a really lousy job, one point for every average performance, and two for a record performance. There were plenty of increments. For example, here's part of the chart on completion percentage:

49.3	0.965	50.0	1.000
49.4	0.970	50.1	1.005
49.5	0.975	50.2	1.010
49.6	0.980	50.3	1.015
49.7	0.985	50.4	1.020
49.8	0.990	50.5	1.025
49.9	0.995	50.6	1.030

The total number of points (or fractions of points) earned by each passer in all four categories are then added together. That total is then matched to a scale that goes from 0 to 158.3. An "average" grade in all four categories will give a passer a 66.6 rating. Only a scattered few have ever scored better than 100.

According to NFL PR, the new system had four major pluses over all the other ways they'd tried before.

First, once a passer's rating had been established, whether for a season or a completed career, it was permanent. Thus, both career and season records were easily determined and more meaningful. In the

THE COFFIN CORNER: Vol. 8, No. 9 (1986)

previous system, ranking passers against other passers, a career rating bobbed up and down every year, depending on what every other active qualifying passer did.

Second, any individual passer's rating could be determined independently of any other passer. Now you could figure out what Fran Tarkenton was doing without first having to know the record of Roger Staubach (and every other passer with a qualifying number of passes).

Third, passers who didn't throw the required number of passes could still be rated for either a season or a career. Before, the only throwers who counted were those with X-number of passes; all the subs were nowhere.

And, finally, if you had all the tables to determine the points and convert the totals, computing the ratings was rapid and easy.

So everyone lived happily ever after, right?

Wrong!

THE SYSTEM UNDER SIEGE

The wonderful NFL Passer Rating System has turned out to be about as popular as the proverbial wet dog in a warm room.

From the start, some people screamed that you couldn't figure a passer's rating without a degree from M.I.T. That wasn't really fair, but it was true that you needed a copy of the NFL's little book of conversion tables. And most fans didn't have one. AND, it was easier to complain about it than to write to the NFL and ask for one.

Other people were terribly upset when their favorite quarterback finished well down in the rankings even though they KNEW he was the best. The answer to that, of course, is that "best" is a subjective term applied to QUARTER BACKS and includes many unmeasurables like ball-handling, leadership, play-calling, and courage. The system rated only how efficiently a PASSER threw the football. Even though most passers were quarterbacks and all quarterbacks were passers, the terms were not synonymous.

A scattered few felt that the numerical rating was too far divorced from any understandable concept (like total yards or average gain) to be meaningful. This, however, was only a matter of education. Once you've looked at a few tables of passer ratings, you begin to get the message that 88.7 is quite good and 37.4 gets a guy on the waiver wire.

There were still arguments against using percentages rather than totals to measure passers. Again, the point of the rating system was to measure efficiency, and for that percentages were proper.

The best counter-arguments came from statisticians who insisted the conversion tables gave too much weight to high completion percentages and low interception percentages but not enough to high touchdown percentages and outstanding gain averages. Such arguments may have been right even though they were wrong – or, at least, irrelevant.

What the statisticians overlooked was that the committee used averages from the period 1960-1972 to set the standards. The mid-points of 50.0% completions, 5.0% touchdowns, 5.5% interceptions, and 7.00 average gain were just what they purported to be – average. They were what a middling passer did during the period.

Moreover, the averages were not interconnected. Each figure was used as one-fourth of a passers' total rating. While it might be possible to "weight" the system toward touchdowns or yards, you'd play hell finding a consensus that would agree that a ten-yard touchdown pass for your side was more important than an interception that sets up a touchdown for an opponent. And, even if everyone could agree, you'd have to further complicate an already complicated system.

THE COFFIN CORNER: Vol. 8, No. 9 (1986)

For fourteen years, the NFL Passer Rating System has withstood all the challenges. It is now the "longest-used" of all the systems and seems likely to continue for some time into the future.

So, Pangloss was correct. It's the best of all possible worlds, right?

Wrong again!

THE PROBLEM WITH GREATS

And small wonder! The NFL has changed the test – er – rules. There are several factors that push current passers to the top of the heap and downgrade the heroes of the past. I have a sinking feeling that the NFL likes it that way too. After all, Marino's ascendancy can sell some tickets and boost some ratings. It's economically preferable to have the Greatest Passer of All Time still active than to have the GP of AT enshrined in the Hall of Fame. Ironically, Don Smith, the man who created the system that crowns Marino is the Vice President of the HOF.

But more important, he's a seeker of truth. If he really thought someone had a better system, he'd be first in line to petition for it. At last count, no one had so convinced him.

And, to tell you the truth, no one has convinced me either.

Year in and year out, the present rating system does a pretty good job of identifying the best and the worst passers. What small tinkering that might tune the accuracy slightly seem to just make things absurdly complicated. I can live happily with the yearly ratings.

If only they wouldn't try to rate the all-time greats!

It's only when I read that my boyhood idols now rank as bums that I feel my gorge rise. Remember how we started this? By showing how Warren Moon rated over Otto Graham.

Well, I never met Graham, but I grew up venerating his name. If he wasn't a god, he was at least a trusted friend. There's no way I can accept Moon over my ami!

Sorry, Charley, the statistics don't match the perceived reality.

In a nutshell, here's what happened. The NFL became very pass conscious after World War II. During the 1950's the emphasis was on long gains to the detriment, as it turned out, completion percentages and interception percentages.

In the 1960-1972 period, these things were more equalized. In theory then, the improved completion percentages and interception percentages canceled out the lessened touchdown percentages and lower average gains. The "average" of the four stats should have stayed the same.

But did they?

The average rating for all qualifiers from the 1960-1972 period was adjusted to read as 66.6. The rating for 1950 qualifiers, even after taking out some three low-ranked passers who threw more than 100 passes but fewer than ten per game, was 56.2 – more than ten points lower!

So what's happened since 1972? Well, pal, in 1985 the qualifiers averaged 78.3! More than eleven points OVER the treasured 66.6!

Here are the figures:

Average Qualifier Stat	1950	1960-72	1985
Completion Percentage	46.6	50.0	59.5
Touchdown Percentage	5.1	5.0	4.3
Interception Percentage	8.0	5.5	4.1

THE COFFIN CORNER: Vol. 8, No. 9 (1986)

Average Gain Per Pass	6.74	7.00	7.05
RATING	56.2	66.6	78.3

See, say the kiddies. It just proves that passers are better than ever. Just like our movies and our music. We are the world! Get out of our way, Old Timer!

At the risk of tripping over my beard, allow me to disagree. You're looking at an illusion, son. You're like the guy in the desert who thinks he sees a water hole a hundred yards away and then finds out it's a mirage. The water is a hundred MILES off. And, if you keep equating the mirage of newness with the truth of time, you're going to end up intellectually thirsty.

That applies to a lot more than football, but football is our subject for today. Tomorrow I'll tell you about music.

There were great passers in the '50s and there are great ones in the '80s. We probably should let it go at that, but being human – and I'll include you, my son – we inevitably try to compare across the eras. If we insist on doing it, we might as well do it right.

THE ADVANTAGES OF MODERNITY

If you are an NFL passer today, you have all the best of it over your father. Curiously, there is a son of a former quarterback in the NFL today, and Jeff Kemp will no doubt top his old man's stats before he's done.

First, there's a psychological change that puts a premium on ball control. Sometimes I think I'm watching a game with Milt Plum quarterbacking both teams. I have a hunch that some of today's dump-off specialists would have been benched in the bombs-away atmosphere of the 1950's. And remember Terry Bradshaw of more recent times? His rating usually wasn't so hot because he couldn't throw a screen pass worth a darn. How'd he make out?

It must be nice for a modern passer to know in a blow-out game that he can boost his rating by throwing little dump-offs. Otto Graham presided over lots of blow-outs, but he didn't know that years later a swing pass to fullback Marion Motley would help him more than a long heave to Dante Lavelli.

Then, today's tosser inherits 50 years of pass patterns, formations, offensive thinking, and clever ideas. It helps. But what is happening is the guys who created the methods are being statistically penalized by those who copied. Think of it like this: Columbus crossed the ocean in 1492. The next guy out could get across a little faster because he already knew the route. Did that make him a better sailor than Chris?

The new indoor stadiums and artificial turf help moderns too. They have many fewer games in which they have to throw in snow or mud. December in Cleveland is still the same, but December in Detroit is the Pontiac Dome. But of all the changes that Rozelle has wrought, none compare to the rule change of 1978. That was when they torpedoed the tough defensive backs by restricting their bump and run tactics after five yards. The effect was devastating to the defense. Receivers dance merrily through the secondary while cornerbacks play a frustrating game of "see-but-no-touchey." On every down, it's only a matter of time until one or more receiver breaks free.

Then, to make sure the passer had that time, they relaxed the rules on pass blocking to let the linemen use their hands. It was like the rulemakers told the passers, "We aren't sure you can shoot those fish in that barrel so you'd better use a hand grenade."

The upshot of all this is that 22 of the 29 qualifying passers in 1986 exceeded the old 66.6 rating.

ADJUSTING TO THE TRUTH

Well, that's all right. I'm not suggesting that they should revise the present passer rating system each year. Aside from it being a lot of work, you'd never be able to tell what the ratings were from week to

THE COFFIN CORNER: Vol. 8, No. 9 (1986)

week unless they revised them according to a new norm every Tuesday morning. Even to preserve the memory of Otto Graham's Greatness, that's more trouble than it's worth.

Besides, it's not necessary. Every baseball fan knows that batting averages were sky high between the two World Wars. When they read those .380's and .390's, they mentally subtract a few dozen points. We can do the same thing every year with Marino's and Montana's ratings. After all, it's only when those ratings are thought of as absolute numbers and then compared to other career ratings from the past, that we have a problem.

We have no difficulty with the order of finish in a given season or in the relative differences among the qualifiers. It's only the career marks!

Baseball statisticians have a simple way of dealing with the batting average disparity. They simply relate a player's BA in a given year to the league's BA for that year. It works fine. By using this relationship, they can demonstrate that a .300 average now is equal to .325 in the 1930's.

To equalize passer ratings and account for the advantages that today's throwers have, all we need to do is find out what the average qualifier did in a year, add or subtract the difference to 66.6, and apply it to a passers' rating. What we get is a readable "state-of-the-art" rating.

For example, Otto Graham's 1950 rating of 64.7 becomes 75.1, better than average but not one of his best. Warren Moon's 1986 rating of 68.5 becomes a less than average 56.8, suitable for the NFL's twentieth-ranked passer. These adjusted ratings – which should only be used in comparing career marks across generations – are more true than the real ratings in a sense. Graham in 1950 was about ten points better than average. Moon last year was about eleven points under the 1986 mid-point.

If we place these adjusted ratings beside a passers' yearly record, we get a quick and accurate record of how he rated against his contemporaries, something you can't do with a career record now.

The adjusted ratings also allow us to make realistic comparisons across generations. To get a passer's adjusted career rating, we simply multiply his adjusted rating each year by the number of passes he attempted that season. In Graham's 1950 case: $75.1 \times 253 = 19000.3$; for Moon in 1986: $56.8 \times 377 = 21413.6$. These point totals mean nothing in themselves because they are dependent on the number of passes thrown – the more passes, the higher the point total.

But, if we add up all the yearly point totals and divide by the total number of passes attempted, we will have a passer's adjusted average for his career. All figures will be in true relation to the ideal 66.6 midpoint, and once again we will have an accurate statistical picture of what really happened.

Marino and Graham will both be up there at the top, but today's lesser lights will not have the an unfair advantage. Nor will those guys whose bubble gum cards I once collected look like a bunch of stiffs.

In summary then, the NFL's Passer Rating System is alive and well in its yearly rankings, but it breaks down in career ratings because of circumstances beyond its control.

Let's fix it.