RATING THE EARLY BACKS

By Doug J. Jones

The Pro Football Hall of Fame is probably finished inducting any players from the early years of the NFL. This fact cannot stop us from debating who the greatest players really were. Anecdotes are nice but statistics provide authority. For lineman all we have are All-Pro votes. For ends we have All-Pro votes and receiving statistics. The backfield positions are more complex. What do the All-Pro votes and statistics tell us?

I have analyzed the consensus first team All-Pro selections for NFL backfields from 1927 to 1961. In 1927, the NFL consolidated its ranks to prevent the dilution of talent. 1961 represents the last year that a 4 man backfield was the standard on major All-Pro teams. The game underwent several major changes during this time period, and the roles of backfield positions changed as well.

I have determined that statistics appear to equate very well with All-Pro votes, and that the statistics that are used appear to be describable by era. Thus, All-Pro votes can be modeled.

The purpose of this is to legitimize ratings systems that I have created (and borrowed) as benchmarks. If we assume that the All-Pro selectors chose the best backs, then it follows that a ratings system that can predict these selections is a sound ratings system. In addition, by demonstrating the systematic nature of the selections we can look at backs with lower All-Pro vote totals in a different light. They may have been among the best – they just didn't fit the All-Pro paradigm of their era.

For each era, I created a statistical rule for All-Pro selections and created a hypothetical All-Pro backfield. The hypothetical teams were compared to the actual consensus team. For 1927 to 1931, I relied on the Green Bay Gazette Poll (the most respected All-Pro team of the day) alone. When the statistical system accurately predicted an All-Pro back, he was classified as a first order fit (1). Second (2) and third (3) order fits were runner ups and second runner ups respectively. An All-Pro back that could be reconciled by a statistically describable method other than the statistical system for the era was recorded as statistically explainable (S). All other All-Pro selections were rated as anomalies (A). Only offensive statistics are used.

Even during the two-way time period, backs appear to have been judged primarily on their offensive skills. One need only consider the 1927 All-Pro selections of the New York Yankees coach, Ralph Scott, to realize this; he divided his All-Pro team into a "clever attack" unit and a "power attack" unit. Quite possibly, with the relatively low scoring of the day, defense was a given. The true star was the back that could break a scoreless tie. Or, offense was simply more visible.

The Triple Threats: 1927 to 1931

To get a quantitative measure of the diverse skills of an early era back with only limited data available, I devised the Triple Threat Rating (TTR).

*TTR = ((Points by Touchdowns and Field Goals + (Touchdown passes * 3)) / Games*

Frankly, I was surprised to see how well the system worked. 15 of the 20 backs fit the system perfectly. The remaining 5 were second order fits.

1927	1928	1929	1930	1931
QB: Friedman (1)	Friedman (1)	Friedman (1)	QB: Friedman (1)	Clark (1)
HB: Driscoll (1)	Wilson (1)	Lewellen (1)	HB: Strong (1)	Blood (1)
HB: Lewellen (1)	Lewellen (1)	Plansky (1)	HB: Grange (2)	Grange (2)
FB: Nevers (2)	Diehl (2)	Nevers (1)	FB: Nevers (2)	Nevers (1)

Outside of Canton, just about anyone would recognize the 1927 All-Pro backfield as a representation of the best backs of the pre statistical era. Lewellen was one of the best punters of his era and also retired as the all time leader in touchdowns. Friedman was considered the best passer and an excellent runner. His arrogance may have kept him out of the Hall of Fame.

Triple Threats and Passers: 1932 to 1937

With more complete statistics available, I used rushing yards with a 3 yard bonus for each TTR point. I included championship games into the average if it helped the back. The most prolific passer (by yardage) was named All-Pro in seasons with poor rushing output (less than 4 backs with 500 yards rushing). The system for this era is a perfect fit for 18 of the 24 All-Pro backs.

QB	HB	HB	FB

1932	Clark (1)	Herber (1P)	Lumpkin (A)	Nagurski (1)
1933	Presnell (1)	Strong (1)	Battles (1)	Nagurski (2)
1934	Clark (1)	Strong (1)	Feathers (1)	Nagurski (2)
1935	Clark (1)	Danowski (1P)	Caddell (1)	Mikulik (A)
1936	Clark (1)	Leemans (S)	Battles (1)	Hinkle (3)
1937	Clark (1)	Baugh (1P)	Battles (1)	Hinkle (1)

Herber was probably the best passer in 1932, 1934, and 1936. He was only All-Pro in 1932 due to the system. Factor this with his career passing totals and his Hall of Fame status appears more reasonable. Herber's teammate, Blood, was also a specialist who usually didn't fit the All-Pro paradigm. He may have been the best receiver and defensive back of the early era, and he too is in Canton where he belongs. Although their style of play didn't win many All-Pro votes, it did win many games; their Packers were the best team of the era. I gave Leemans an "S" in 1936; he was named as a halfback but is usually listed as a fullback.

Non Functional Honor Teams: 1938 to 1946

The All-Pro paradigm of this era was to name the top two passers and the top two runners. For 7 of the 9 years, the All-Pro teams were so split. The two deviating years were 1939 and 1944. During both seasons, Baugh's backup, Filchock, was the NFL's top passer. Could it be that rather than name him All-Pro, most selectors decided to just pick another runner?

To rate the players, I looked at NFL passer rating (1960 - 1971 version) for passers (minimum 1,000 yards or 100 attempts). For runners, I used the rushing yardage with a bonus of 3 yards for points scored by touchdowns and field goals. Touchdown passes were no longer included; with full time passers, passing effectiveness (or ineffectiveness) was ignored. The system accurately predicted 28 of the 36 backfield positions.

	Passer	Passer	Runner	Runner
1938	Danowski (1)	Parker (A)	White (1)	Hinkle (A)
1939	Leemans (A)	Hall (1)	Osmanski (1)	Farkas (1)
1940	Baugh (1)	Parker (1)	White (1)	Drake (1)
1941	Isbell (1)	Luckman (1)	McAfee (1)	Manders (1)
1942	Baugh (1)	Luckman (2)	Dudley (1)	Famigelleti (3)
1943	Baugh (1)	Luckman (1)	Cuff (S)	Canadeo (S)
1944	Cuff (S)	Luckman (1)	Paschall (1)	Sinkwich(1)
1945	Baugh (1)	Waterfield (1)	Van Buren (1)	Akins (1)
1946	Waterfield (1)	Luckman (1)	Dudley (1)	Fritsch (1)

More and more, passers such as Baugh, Luckman, and Waterfield were the stars. The days of Dutch Clark's all out infantry attack were over. The glory days of the modern quarterback were beginning. But two quarterbacks? These players also played defense and special teams. No one has tried to reconcile an All-Pro defensive end to sub for Don Hutson (who played back on defense). After all, it only makes sense to have 5 defensive backs to match up against two passers. And Luckman hands off to Baugh…

Quarterbacks, Running Backs, Flanker Backs, and Kicker Backs: 1947 to1961

I have divided the backfield into four typologies pertaining to the function of a back. These types of backs are quarterbacks, running backs, flanker backs and kicker backs.

My rating system (QB2002) is an inverse ranking system that judges quarterbacks by their team's record, their passer rank (the NFL's 1960 - 1971 system – only quarterbacks with 1,500 yards or more were included in the ranking), and IMPACT_PTS.

*IMPACT_PTS = [(yardage by running and passing – sack yardage for team) / (total team yardage) *(points by touchdown and field goal for the team's offense other than touchdowns and touchdown passes by other quarterbacks on the team)]*

Consider this example: Bobby Layne ran and passed for 60 yards on an 80 yard touchdown drive. No matter how the touchdown was scored, Layne gets 4 IMPACT_PTS (60/80 * 6).

Running backs primarily rush. Therefore, I rated them by rushing yards alone. As simple as it sounds, this appears to be the most important stat.

Flanker backs also run with the ball, but many were just as effective at receiving. Flanker backs are ranked by yards from scrimmage and touchdowns (18 yard bonus). By the end of the study period, the proliferation of passing creates a caveat. Therefore, from 1957 to1961 at least 40% of a flanker back's yards must come from receiving or they are ranked as a running back.

No Kicker position was recognized by All-Pro selectors until 1965. A kicker back may not have been as prolific as running backs or flanker backs (although Walker led the NFL in adjusted yards from scrimmage in 1950), but his versatility was important to his team, and All-Pro selectors. To quantify the overall effectiveness of a kicker back, I have come up with the following criteria. To qualify, an individual must be the NFL's top scoring back, score at least 50 points from kicking, and 90 points overall. The kicker back must also have 400 yards from scrimmage with at least 200 from rushing. From 1957 to 1961, I added a kicker back to score over 1,000 adjusted yards from scrimmage (3 yard bonus for points by touchdown and field goal).

The All-Pro selections for this time period represent the offensive backfields. For the 1947 to 1956 era, 40 backs were selected All-Pro; 31 were first order selections and 3 were second order fits. 16 of the 20 backs were first order fits for the 1957 to 1961 era. The remaining 4 were second order fits.

The system for the 1947 to 1956 era
The top rated Quarterback
The top rated Running Back
The top rated Kicker Back (if applicable)
    If no one qualified as a Kicker Back, then an additional Running Back
The top rated Back in yards from scrimmage excluding players already selected.

	QB	RB	KB or RB	FLB
1947	Luckman (1)	Van Buren (1)	Harder (1)	Dudley (1)
1948	Baugh (2)	Van Buren (1)	Harder (1)	Trippi (1)
1949	Waterfield (A)	Van Buren (1)	Harder (1)	Canadeo (2)
1950	Lujack (A)	Motley (1)	Walker (1)	Geri (S)
1951	Graham (1)	Price (1)	Walker (1)	Jones (1)
1952	Graham (1)	Towler (1)	Price (1)	McElhenny (1)
1953	Graham (1)	Perry (1)	Walker (1)	McElhenny (A)
1954	Graham (1)	Perry (1)	Walker (1)	Matson (1)
1955	Graham (1)	Ameche (1)	Matson (A)	Gifford (A)
1956	Layne (2)	Casares (1)	Matson (1)	Gifford (1)

The system for the 1957 to 1961 era
The top rated Quarterback
The top rated Running Back
The top rated Kicker Back (if applicable)
If no one qualified as a Kicker Back, then an additional Flanker Back
The top rated Flanker Back excluding players already selected.

	QB	RB	KB or FLB	FLB
1957	Tittle (2)	Brown (1)	Matson (2)	Gifford (1)
1958	Unitas (1)	Brown (1)	Arnett (2)	Moore (1)
1959	Unitas (1)	Brown (1)	Gifford (1)	Moore (1)
1960	Van Brocklin (1)	Brown (1)	Hornung (1)	Moore (1)
1961	Jurgensen (2)	Brown (1)	Hornung (1)	Moore (1)

In a perfect world, hunger would be eliminated, global warming would not exist, and quarterbacks would not be judged upon their team's success. I find IMPACT a compelling rate; guys like Unitas, Layne, and Graham rate score well in this system. The rating must be good. It has been said that "rabid but unsophisticated fans" talk about running backs (The Hidden Game of Football). These results demonstrate the difficulties of comparing even contemporary backs.

Summary and Conclusions

A summary of the 5 eras is compiled below.

Era & Count	First Order Perfect Fit	Second Order Or Better	Third Order, Statistical Superlative, Or Better
1927-31 (20)	15 – 75.0%	20 –100.0%	20 – 100.0%
1932-37 (24)	18 – 75.0%	20 – 83.3%	22 – 91.7%
1938-46 (36)	28 – 77.8%	29 – 80.6%	33 – 91.7%
1947-56 (40)	31 – 76.5%	34 – 85.0%	35 – 87.5%
1957-61 (20)	16 – 80.0%	20 – 100.0%	20 – 100.0%

Total (140) 108 – 77.1% 123 – 87.9% 130 – 92.9%

All told, 77.1% of the consensus All-Pro backs of the time period could have been predicted by my system. This is a remarkably strong fit, and noteworthy for several reasons. Before 1932 no official statistics were available, and afterwards they were not fully reliable. For the first three eras, the backs played on defense and the system only measures offense. The data strongly infers that offensive statistics were a main measure of success even in this early era.

Since one cannot expect All-Pro selectors to feverishly calculate the best players by any system, let alone mine, the second order fits (87.9%) are important as well because they demonstrate that the general concept of the system is adhered to. Many of these players had overwhelming superiority in areas not measured by the system and/or had name recognition from having been voted All-Pro in a previous season. I have included some interesting examples:

In 1927, Ernie Nevers (4.33 TTR) beat out Jack McBride (4.85 TTR) to be named All-Pro at the fullback postion. Intangibles and mystique aside, the stats can still support Nevers. In 1927 he was involved in 90% of his team's touchdowns to 48.1% for McBride.

Nagurski won All-Pro honors as the top rated fullback in 1932, but also won honors as the number two rated fullback in 1933 and 1934. He had only 67.4% of Musick's rate in 1933 and 76.9% of Hansen's rate in 1934. But Nagurski's blocking and line-backing must have given him the edge. In 1934, he did everything but carry Beattie Feathers for 1,000 yards.

Tom "The Bomb" Tracy out-gained Jon Arnett in yards from scrimmage in 1958, but he also out-fumbled him 10 to 4. Arnett received the All-Pro honors.

The anomalies (fourth order fits or worse) are also worth examining. Of the 10 anomalies, 7 of the players were a first order fit All-Pro in at least one other season. Incidentally, all 7 of them are in the Hall of Fame.

Hopefully this research has initiated some new ideas to stir up old debates. Here are some final conclusions on rating the early backs that apply to comparing players in general…

Legitimize ratings systems against All-Pro selections

Most ratings systems are based upon theory. I argue that these theoretical systems should be grounded in perception. We have to assume that the best players were *usually* named to the All-Pro teams. Thus, the frequency by which a top rated player is named as an All-Pro can be used to rate a ratings system. If ratings system A accurately ranks the consensus All-Pro quarterback as number one 7 times in a 20 year time period whereas ratings system B does so 13 times in the same period, one can reasonably argue that B is a better system. This same method can also be used on a wider scale to see how often the top 3 quarterbacks in each conference are named to the Pro Bowl.

Scrutinize All Pro selections against era paradigms

All-Pro votes are not the final word on a player's relative greatness or mediocrity. Because All-Pro selectors appear to follow voting patterns by era we can surmise that skills are emphasized differently between eras. Successful play, however, is timeless. Special consideration needs to be afforded to players (Blood & Herber come to mind) whose skills fall outside of the paradigm. First the paradigm must be identified. Then a ratings system needs to be developed (or borrowed from another era) to rate players that fall outside the paradigm. Finally, these players can be ranked, their statistics benchmarked, and theoretical *all-pro* votes can be interpolated.

Compare players based on functionality

In addition to being wary of comparing Cliff Battles to Cory Dillon, we must also be careful about comparing Jim Brown to Lenny Moore. Moore should be compared to backs who had similar roles on their team, i.e. Gifford, McElhenny, and Matson. Brown should be compared to players such as Ameche, Casares, and Taylor.

Finally, always develop a system in which your favorite player rates the highest

There are lies, damned lies, and statistics. Remember not to get too carried away. Besides, if you can't statistically *prove* that your guy is better, you can always yell louder.

Sources and Notes:

For player statistics I used The Football Encyclopedia – *Professional NFL Football From 1892 To Present* By Cohen and Neft (1991 edition). For All-Pro teams I used Total Football by Carrol, Gershman, Neft, and Thorn (1997 edition). The Coffin Corner article "All Pro Addenda" (Vol V no. 2) was also used.

Consensus NFL All-Pro backfields were created by counting how many first team selections each player had. Second team selections were used as tie breakers. I used the following All-Pro teams: The Green Bay Press Gazette, The Associated Press, Professional Football Writers Association, International News Service, Official NFL Team, Pro Football Illustrated, The Sporting News, and The United Press.