

Friday, September 4, 2009

Updated: December 7, 11:28 AM ET

A guide to ESPN's SPI ratings

By Nate Silver
Special to ESPN.com

Objective

The SPI rating is designed to provide the best possible objective representation of a team's current overall skill level. In particular, the SPI ratings are intended to be forward-looking: They measure a team's relative likelihood of victory if a competitive match were to be held tomorrow. This concept may differ somewhat from a retrospective or backward-looking ratings system. The SPI ratings are not trying to reward or punish teams based on their past results; rather, they are trying to predict which teams will have the most success going forward.

The challenge in preparing an international soccer ratings system is that there is relatively little reliable data to go by, as compared with other sports. If a particular international team is not engaged in a major competition, such as the World Cup, it may play only a handful of meaningful matches each year. Compare that to a 162-game season in baseball, an 82-game season in basketball, or a 16-game season in American football. Many of these games, moreover, may be against teams of inferior quality, or may feature marginal lineups as many of a team's star players are engaged in club competition instead. For that reason, it is important to be somewhat expansive about the amount of data that we use in a soccer ratings system. Things like margin of victory and home-field advantage, which are ignored by some other ratings systems, play a fairly large role in SPI. More distinctively, SPI blends ratings from club competition with those from international play, providing for a more robust assessment of the level of talent on a particular team.

Soccer is a rich, wonderful and unpredictable sport, and it would be quite a shame if a single number could tell us everything that we needed to know about a soccer team. SPI does not. It merely reflects the relatively limited statistical information that is available in international soccer, and does so in a way that is as fair and accurate as possible. In other words, SPI is designed to serve as a general guideline -- as a starting point for debates about team quality. It is not intended to be the ending point or to settle all arguments.

Basic approach

The SPI proceeds in four major steps:

Calculate competitiveness coefficients for all games in database

Derive match-based ratings for all international and club teams

Derive player-based ratings for all games in which detailed data is available

Combine team and player data into a composite rating based on current rosters; use to predict future results.

Each step is described in some detail below.

Step 1 -- Competitiveness coefficients

One of the difficulties in evaluating international soccer is that the seriousness with which a team treats a particular game can vary significantly from match to match. It's as though on some occasions, the New York Yankees are the New York Yankees, and on others, the Yankees have been replaced by the team's Triple-A affiliate, the Scranton-Wilkes Barre Yankees. The latter sort of match is not likely to be very informative as to how strong the New York Yankees really are, if they were to play the Boston Red Sox in a critical game that day.

Soccer Power Index

Who's the No. 1 team according to Nate Silver? View the [SPI Rankings](#) in full.

SPI's approach to this problem is to calculate a competitiveness coefficient for each match. The goal of the competitiveness coefficient is to measure how much of its "A" lineup -- the lineup that a team would use if a World Cup match were played on that day -- a team uses for each game.

SPI FAQ

Need more information about SPI? The [Soccer Power Index FAQ](#) has all the answers.

The competitiveness coefficient is determined by evaluating each player in the lineup on a particular day, and seeing how often they played in other competitions we know to be important, such as the World Cup, European Championships and Confederations Cup. Each player in the lineup is scored from 0 to 1 depending on the fraction of the possible minutes they played in such contests. A "fudge factor" is used to account for the fact that, for instance, some players may be new to the lineup because they have been injured, or because they are young players who have recently elevated their games to the point that they belong on the "A" team. Essentially, if a player has played half of the possible minutes in important competitions near in time to that date, they will receive full credit, but the numbers are scaled down proportionately from there.

These preliminary ratings are then averaged across each member of a team's lineup that day, weighted by the number of minutes each player played. These are referred to as the team competitiveness coefficients (TCC). The team competitiveness coefficients are then multiplied together and then multiplied again by a constant, which yields the combined competitiveness coefficient or CCC.

$$CCC = TCC (\text{Home Team}) \times TCC (\text{Visiting Team}) \times 1.27$$

So, for example, if Brazil is playing Colombia in a friendly match, and Brazil has a TCC of .25 (indicating that about one-fourth of their players are from the "A" lineup) and Colombia has one of .67, the CCC for that game will be:

$$\begin{aligned} CCC &= .25 \times .67 \times 1.27 \\ CCC &= .213 \end{aligned}$$

This multiplicative process ensures that both teams must be taking a game seriously before it receives a particularly high weight. If, as in the example above, Colombia is taking the game fairly seriously but Brazil is not, the game will not receive a high CCC.

TCCs are bounded at a minimum of .10. This means that the minimum CCC is:

$$.10 \times .10 \times 1.27 = 0.0127$$

By contrast, the maximum CCC -- the one used automatically, for example, for all World Cup matches, is 1.27. This means that there is potentially as much as a 100-fold difference between the weight that the SPI gives to a World Cup game and the one it gives to a friendly match or minor international competition that the teams are not taking very seriously.

In some cases, detail on each team's lineup is not available for a particular match. In these cases (and these cases only), a default CCC is used depending on the competition type. The default CCCs are based on an empirical analysis of lineup composition from that competition in games for which we do have rosters available. The default CCCs are as follows:

World Cup 1.27

World Cup -- Interconfederation Playoff 1.09

European Championship 1.00

World Cup Qualifying -- Europe 0.82

Confederations Cup 0.79

World Cup Qualifying -- South America 0.68

World Cup Qualifying -- Africa 0.68

World Cup Qualifying -- Asia 0.68

Africa Cup of Nations 0.57

World Cup Qualifying -- Oceania 0.57

World Cup Qualifying -- CONCACAF 0.55

Qualifying -- European Championship 0.47

AFC Asian Cup 0.46

Oceania Nations Cup 0.46

Qualifying -- Africa Cup of Nations 0.32

Copa America 0.25

Gold Cup 0.24

Qualifying -- Gold Cup 0.24

Qualifying -- AFC Asian Cup 0.24

Friendly Match 0.24

Relative to the FIFA Rankings, the SPI tends to rate more highly the World Cup, European Championship, Confederations Cup and most World Cup qualifying (especially on continents like Europe and South America, where it is difficult to qualify; less so for continents like North America, where the qualifying process is more forgiving). Other competitions -- particularly minor continental championships like the Gold Cup -- are weighted lower than in FIFA, as are friendlies.

Keep in mind, however, that these default ratings are just a default -- they are overridden wherever specific lineups are available. So, for example, if a knockout stage match in Copa America is being treated quite seriously by both teams, it will receive a high weight, even though by default Copa America receives a low weight since teams often aren't fielding their "A" lineups.

Step 2 -- Match-based ratings

The goal of the match-based ratings is to develop an offensive (OFF) and defensive (DEF) rating for any given team at any given moment in time, which in turn reflects their goal-scoring ability and goal-prevention ability, respectively.

The first step in calculating the match-based ratings is to evaluate individual games based on the number of goals a team scores and allows, relative to the quality of competition. Specifically, we arrive at figure known as Adjusted Goals Scored (AGS) and Adjusted Goals Allowed (AGA) total for each team for each match. These are calculated as follows ...

$$AGS = ((GS - OPP_DEF) / (\text{MAX}(0.25, OPP_DEF * 0.424 + 0.548))) * (AVG_BASE * 0.424 + 0.548) + AVG_BASE$$

$$AGA = ((GA - OPP_OFF) / (\text{MAX}(0.25, OPP_OFF * 0.424 + 0.548))) * (AVG_BASE * 0.424 + 0.548) + AVG_BASE$$

... where GS and GA is the raw number of goals that a team scores and allows in a particular game, OPP_OFF and OPP_DEF are the opponent's offensive and defensive strength ratings, and AVG_BASE is a constant indicating the average number of goals scored per game in international competition (about 1.37 goals per team per game).

The Adjusted Goals Scored and Adjusted Goals Allowed figures can differ quite substantially from the raw figures, depending on the quality of competition. For example, in April 2001, Australia scored a 31-0 win over American Samoa, a team that routinely allows double-digit goals to its opponents and almost never

scores. Australia's AGS and AGA figures for this game are 3.92 and 1.36, respectively, meaning that it's treated as no more than about a 4-1 win. By contrast, when Paraguay defeated Brazil 2-1 on July 14, 2004, they received an AGS and AGA of 3.62 and 0.29. This one-goal win, therefore, is treated as the equivalent of a 3-0 or 4-0 victory.

An additional adjustment is performed for home-field advantage, which is worth about 0.57 goals in international soccer (this is very significant; home-field advantage is worth two to three times as much in international soccer as it is in the NFL). If a game is played on a neutral site, the home-field penalty/away-team bonus is split evenly between the two teams. The goals scored and goals allowed figures are also adjusted for game length. A team that wins a game in a penalty shootout is considered to have scored an additional $\frac{1}{2}$ goal.

One characteristic of the AGS and AGA figures is that a team may wind up with a positive score even if they have lost the game (that is, their AGS for that game is higher than their AGA), or a negative score even if they have won the game. For instance, if a team loses 3-2 at Brazil or Spain, it will usually finish with a net positive rating for that game, since most teams will be beaten much more badly than 3-2 on the road against such opponents. Conversely, if it wins 2-1 at home against San Marino, it will generally lose some credit, since most teams should be capable of a much more persuasive victory.

Bear in mind that the aim of SPI is to be predictive. If a team has only defeated San Marino by one goal, they get a win in the standings column nevertheless. But, we have found, goals scored and goals allowed, when adjusted in this fashion, are a much better predictor of future results than wins and losses alone. A team that beats San Marino only 2-1 at home will usually run into a lot of trouble unless it improves its form in future matches. Conversely, a 3-2 loss on the road against Brazil really can be considered a "moral victory" more often than not and bodes well for a team's future competitions.

Once a team's AGS and AGA are calculated for each game, they are combined by means of a weighted average. The average is weighted based on two factors: the combined competitiveness coefficient (as described in Part 1) and a recentness factor.

The recentness factor gives less weight to games that were played further back in time. This is used by establishing a cut-off point at a given point in time and calculating the recentness factor in a linear fashion from there. For example, if the cut-off point is four years, that means a game played exactly four years ago will be given no weight, a game played three years ago will be given one-quarter weight, a game played two years ago will be given one-half weight, a game played one year ago will be given three-quarters weight, and a game played yesterday will be given almost full weight. In addition, a bonus of up to 25 percent is given to games played within the past 100 days to reflect a team's most recent form.

The cut-off point is dependent, however, on the frequency with which a team plays competitive matches. For teams like Spain that engage in competitive matches relatively frequently, the cut-off factor may be as little as three and a half years (less than the four years used for the FIFA rankings). For teams that play less frequently and for which little reliable data is available -- say, New Zealand -- the cut-off point may be considerably further back in time -- SPI may look as much as eight years back or longer. This is intended to strike a balance between the relative lack of competitive matches in international soccer as compared with other sports, and the fact games further back in time are self-evidently less indicative of a team's current talent level.

Once we have calculated this recentness factor, in addition to the competitiveness coefficient, we can "roll up" (take a weighted average) the AGS and AGA figures that a team receives in each match to produce overall OFF and DEF ratings. These ratings can be calculated relative to the current game date, or relative to any other point in time.

There is one significant issue, however, that we have somewhat glossed over. A team's OFF and DEF ratings are based in large part on its adjusted goals scored (AGS) and adjusted goals allowed (AGA) figures for individual games. But AGS and AGA are in turn determined, in part, by referencing the OFF and DEF factors of a team's opponents.

To place this seeming paradox in plain English: Suppose that Mexico beats the United States 2-1 in a qualifying match. How much credit should we give Mexico for that win? It depends on how good the United

States is. But how good is the United States? It depends on how good its opponents were. And how good were the United States' opponents? It depends on how good their opponents' opponents were. This "loop" can continue indefinitely. In mathematical terms, the problem is that we are trying to solve for the independent and dependent variables at once.

Fortunately, there is a shortcut -- a process known as iteration. This is the same process used for many of the more popular college football and college basketball rankings. We start with an initial estimate of opponent quality -- for example, how many goals the opponent scored and allowed over all games in the database. Then this estimate is continually refined -- made more and more accurate -- each time we run through another iteration or "loop" and incorporate more information. After about 20 to 30 iterations, this estimate becomes stable down to several decimal places, and we are able to calculate reliable OFF and DEF factors for any given team on any given date.

Note that OFF and DEF have a relatively specific interpretation. OFF reflects the number of goals we'd expect a team to score against an "average" opponent ("average" in this context means a team that ranks in the 50s or 60s worldwide -- a team like Canada or Lithuania), and DEF reflects the number of goals we'd expect them to concede against such an opponent.

In addition to calculating match-based ratings for international teams, we must also do so for club teams as a precursor to Step 3. This process operates in an identical fashion to the international team ratings as described above, with a few very minor exceptions. First, a lower home-field advantage constant is used, since the travel distances are shorter in club competition and home field has historically made somewhat less difference (although it still matters greatly relative to other sports). Second, there is no "competitiveness coefficient" -- all games are assumed to be equally important, with the exception that matches in the UEFA Champions' League are given a 50 percent bonus. Third, since club competition, unlike international competition, features clearly demarcated "seasons" that tend to be associated with relatively heavy roster turnover, an additional penalty is applied -- equivalent to an extra months' worth of downtime -- for the gap between each season when the recentness factors are computed.

It is also necessary to develop an adjustment factor for these club-team ratings, since the average level of play in the "Big 4" leagues and the Champions' League (the competitions used in our program) is considerably higher than that between two international teams of average quality. The adjustment factor was calculated based on comparing the game-level ratings (see Step 3) of players in international and club games and ensuring that, on average, a player will receive about the same rating when he plays for his club team as he does when he plays for his international side. Some rough equivalencies between club and international teams, based on results through July 2009, are as follows:

Brazil ⇔ FC Barcelona

Germany ⇔ Chelsea

USA ⇔ FC Porto

Sweden ⇔ Tottenham Hotspur

Bolivia ⇔ Sunderland

Tanzania ⇔ Derby County

Step 3 -- Player-based ratings

The OFF and DEF ratings we obtained for each team in Step 2 is only half of the SPI. The other half are the player-based ratings -- an assessment of the quality of the particular players in an international team's lineup based on their performance in both club and international play.

The club leagues used for the player-based ratings are the "Big 4" European leagues -- England, Spain, Italy and Germany -- plus the UEFA Champions' League. These leagues are home to more than 90 of the world's

100 best players, according to a recent analysis by the soccer Web Site and magazine FourFourTwo. We should note, however, that the mere presence of a team's players in one of the Big 4 leagues does not alone give them any credit. Rather, his international team may either gain or lose credit based on the performance of that player for his club team. The ratings are very carefully designed such that a team will not be penalized (nor rewarded) if its players make their home in club leagues other than the "Big 4."

The player-based ratings are calculated by evaluating individual games for which we have detailed data. The starting point is the adjusted goals scored (AGS) and adjusted goals allowed (AGA) figures that we have calculated for a particular game. From the AGS we subtract the number of goals scored. With the AGA, we subtract from the number of goals allowed. That results in a plus-minus rating for each game.

For example, in their 2002 World Cup quarterfinal match against England, which they won 2-1, Brazil's AGS is 3.81 and their AGA is 0.29. If we subtract the international average of 1.37 goals scored per game, we wind up with an OFF rating for that game of +2.44. If we subtract the AGA, from the average goals allowed, we get a DEF rating of +1.08.

The key purpose of the player ratings algorithm, then, is to take a team's plus/minus rating for a particular game and assign it out to their individual players. That is, each player will receive his own personal OFF and DEF rating for each game.

The first and foremost requirement is that the sum of the ratings for all individual players on a team must equal their team's rating for that game; soccer is too much of a team sport to make any other assumption. So in this game against England, the OFF ratings for Brazil's starting 11 (plus substitutes) must total plus-2.44, and their DEF ratings must total plus-1.08. This is an inviolable property of our ratings.

Allocating credit/blame to individual players is a multi-stage process, but the basic components are as follows:

Primary (direct) credit for goals scored. If a player scores, then we're going to give him some credit for that. Specifically, we assign him half (50 percent) of the credit for his goal. Less credit (only 20 percent of the total) is assigned if a player scores on a penalty. If there are any own goals, we also pin the blame on the guilty players at this point.

Secondary (indirect) credit for goals scored (and allowed). The other 50 percent of the credit for scoring is assigned to the teammates that were on the field with a player at the time of his goal -- this is equivalent to a plus-minus rating in a sport like hockey. Forwards and midfielders are assigned proportionately more credit for assisting with scoring than defensive players. Conversely, when a team allows a goal, the players who are on the field at the time the goal is conceded receive a hit to their defensive ratings. The goalkeeper, of course, receives the most substantial penalty, followed by the defenders, the midfield and the strikers.

Bookings. No surprise here, but getting a red card puts your team at a huge handicap. In fact, we've found that when a player gets sent off, it decreases his team's scoring output by around 0.3 goals per 90 minutes, and increases the other team's scoring by 0.5 goals per 90 minutes. Therefore, a player receives a substantial penalty for being dismissed, assessed against both his OFF and DEF ratings. The magnitude of the penalty depends in part on when the booking occurs -- getting sent off in the 1st minute is much more damaging than getting sent off in the 88th minute. A player is also punished for cautions, but the penalty is much less substantial.

Residual ratings. Just as we give credit to players for scoring, we also have to penalize them for failing to score. Essentially, we assign a player a very small penalty to his OFF rating for every minute he's on the field while his team fails to score, and a very small amount of credit to his DEF rating for every minute he's on the field when his opponents do not score. If a team gets shut out then, for example, the OFF rating for the strikers and the midfielders will be substantially negative because of this residual rating. Conversely, this is the principal way in which defenders and particularly keepers (and to a lesser extent midfielders) get credit for preventing goals.

Note that an individual's rating may be negative for a game, even when a team's rating is positive, or vice versa. For instance, if a player gets sent off in the 10th minute, his rating will almost certainly be negative even if his team wins 2-0 (we do, however, give some "extra credit" to his teammates for managing such a

good result while down a man). Another common situation is that a team might win a high-scoring game -- say, 4-3. In this case, the overall ratings for the strikers and midfielders will usually be in positive territory for that game, but the defenders and goalkeeper will not generally be rated highly.

Below is an example of how the ratings were apportioned for the Brazil-versus-England match described above.

Brazil vs. England 2002

Player	MIN	Pos	OFF	DEF	TOTAL
Rivaldo	90	F	+0.83	+0.03	+0.86
Ronaldinho	90	M	+0.77	+0.04	+0.81
Gilberto Silva	90	M	+0.18	+0.08	+0.26
Kleber	90	M	+0.18	+0.08	+0.26
Ronaldo	70	F	+0.20	+0.02	+0.22
Roberto Carlos	90	D	+0.06	+0.13	+0.19
Cafu	90	D	+0.06	+0.13	+0.19
Edmílson	90	D	+0.06	+0.13	+0.19
Lucio	90	D	+0.06	+0.13	+0.19
Roque Junior	90	D	+0.06	+0.13	+0.19
Marcos	90	G	+0.01	+0.16	+0.16
Edilson	21	F	-0.02	+0.03	+0.01
Brazil			+2.44	+1.08	+3.52

Rivaldo and Ronaldinho, as the goal-scorers in this game, naturally get the best ratings. Although ordinarily a midfielder who scores a goal will be rated a bit higher than a striker who does the same, Ronaldinho's are a bit worse than Rivaldo's because he also received a caution. The non-scoring midfielders, getting credit for their solid two-way play, are next on the list, just slightly trailed by the defenders. But everyone is well into positive territory, except for the one substitute, Edilson, who gets a relatively poor rating because both of Brazil's goals were scored before he took the field.

Note that the OFF and DEF ratings add up to plus-2.44 and plus-1.08, respectively, which was Brazil's overall rating for this game. Since the ratings for individual players ultimately stem from the team's rating for this game, one implication of this is that a player will receive more credit for a goal scored against a tougher opponent, a keeper will receive more credit for a clean sheet against a tougher opponent, and so forth.

By running through this process for all games in the database, we can create an overall rating for a player, which can be expressed on either a cumulative or -- as is more useful for our purposes -- a per-90-minute basis, which we call OFF90 and DEF90. (More technically, a per-96 minute basis, since we assume three minutes of stoppage time at the end of each half). This process employs a weighted average much like the one that takes place in Step 2, where games are weighted based on a competitiveness coefficient and a recentness factor. Both international and club games are included in the average. The competitiveness coefficients are the same as described in Step 2 for international play, and are fixed at .36 for club play (exception: .54 for Champions' League games). The recentness factor uses a cut-off point of four years, meaning that a game played four years ago will be given no weight, a game played two years ago will be given one-half weight, a game played one year ago will be given three-quarters weight, and a game played yesterday will be given almost full weight. As a result of this, we have a single OFF90 and DEF90 rating for each player in the database.

The OFF90 and DEF90 ratings for individual players are designed such that they can be recombined to provide another assessment of strength for a particular team. For example, here are the combined OFF90 and DEF90 ratings for a lineup recently used by Spain in its game against Estonia.

Spain vs. Estonia

Player	Pos	OFF90	DEF90
Iker Casillas	G	-0.00	+0.10
Carlos Marchena	D	-0.01	+0.06
Joan Capdevila	D	+0.06	+0.08
Raul Albiol	D	+0.06	+0.04
Gerard Pique	D	+0.08	+0.10
Xabi Alonso	M	+0.14	+0.05
Marcos Senna	M	+0.09	+0.04
Cesc Fabregas	M	+0.15	+0.05
David Silva	M	+0.17	+0.04
Fernando Torres	F	+0.35	+0.02
David Villa	F	+0.38	+0.01
Subtotal		+1.47	+0.58
x-1 for Def Rating		+1.47	-0.58
+ League Average		+1.37	+1.37
Player-Based Ratings		2.84	0.79

Any set of ratings can be combined in this way. Theoretically, they could even be used to assess the impact of particular players. For example, if David Villa (OFF90 rating of plus-0.38) were injured and replaced by Juan Manuel Mata in the lineup (OFF90 rating of plus-0.12, not shown here), we would estimate that this would reduce Spain's scoring output by about 0.26 goals per game, the difference between the ratings for Villa and Mata.

What we are more interested in, however, is looking at the quality of a team's "A" lineup -- the one that we expect them to use in the World Cup and other major competitions. In order to determine this lineup, we look at which players have been playing in highly rated games and competitions, using competitiveness coefficients through a process parallel to that used in Step 1. If a player shares his spot in the starting lineup, or is frequently substituted out during the game, that position will be split among that player and the players relieving him. Once these lineups are determined, we are able to calculate player-based OFF and DEF ratings for each international team.

Step 4 -- Composite ratings

The final step actually consists of two substeps. The first substep is reasonably straightforward. Thus far, we have generated two sets of OFF and DEF ratings for each international team: one based on team-level results in international play (Step 2), and the other based on the performance of individual players in both

international and club play (Step 3). We then need to figure out how much emphasis is given to each of these sets of ratings.

This is determined by comparing two things:

The number of competitive international matches a team has played recently.

The number of minutes this team's "A" lineup has played in recent club-level competition in the Big 4 leagues (plus the UEFA Champions' League).

The more data we have of the second type and the less data that we have of the first type, the more weight is placed on the player-based ratings (Step 3) as opposed to the game-based ratings (Step 2). For instance, for a team like England, for whom virtually all the regulars play in the Big 4 leagues but which has a somewhat limited set of recent international play because the team failed to qualify for Euro 2008, about 70 percent of the weight is placed on the player-level ratings. For a team like Russia, which plays very frequently internationally but which has only a few players in the Big 4 leagues, only about 15 percent of the combined rating is based on the player-based numbers. Some teams may have none of their players in the Big 4 leagues or the Champions' League at all; in these cases, 100 percent of the weight is placed on the game-based ratings and the player-based ratings are irrelevant. The weighting scheme is designed such that, for a typical "major" international team, about half the weight will be placed on the game-based ratings and the other half on the player-based ratings, although this fraction can vary substantially from team to team.

Once the OFF and DEF ratings have been combined into a composite rating, there is one last step, which is creating an overall rating, RATE, that scales from 0 to 100 and reflects a team's overall strength.

One feature of the OFF and DEF ratings as we've designed them is that they can be combined to estimate the win probabilities for any given game between any two given teams. These probability estimates were derived using something known as a multinomial logit model -- essentially, we've gone back into time and examined what happened when two teams of a given strength rating faced off against one another. For example, based on their current OFF and DEF ratings, this formula estimates that Spain would defeat the United States 61 percent of the time on a neutral field, draw 27 percent of the time and lose the remaining 12 percent of the time. Another example: Germany would defeat Switzerland 54 percent of the time, draw 29 percent of the time and lose 17 percent of the time.

In order to calculate RATE for a particular team, we have them play a round-robin against all other teams in the world using this formula, and then add up the percentage of the possible points (three for a win, one for a draw) that each team would score in such a round-robin. So let's create an example case. For instance, if Team A were to play such a round-robin, our ratings might predict that it would win about 88 percent of its games (scoring three points), draw 9 percent (scoring one point) and lose 3 percent (scoring no points). Team A's overall rating would then be ...

$$(.88 \times 3) + (0.09 \times 1) + (.3 \times 0)$$

$$(1.00 \times 3)$$

... which evaluates to 0.91, which we express as "91" without the decimal place.

SPI ratings run from a theoretical minimum of 0 to a theoretical maximum of 100. A team with a rating of 100 would be a lock to beat every other national team, while a team with a rating of 0 would be guaranteed to lose to every other national team.

As a general guideline, the following terms can be used to describe national teams:

- 85+: Elite
- 80-84: Very strong
- 75-79: Strong

- 70-74: Good
- 60-69: Competitive
- 50-59: Marginal
- 25-49: Weak
- 0-24: Very weak

Although RATE is the best overall measure of team quality, we would note that the formula to predict win probabilities is not linear. For example, against elite competition -- such as most of the field in South Africa - defense (DEF) tends to be more important than offense (OFF), and teams with stronger DEF ratings are slightly stronger than their RATE suggests. Conversely, against weak competition, teams with stronger DEF ratings are slightly more likely to be upset than those with stronger OFF ratings, and are slightly weaker than their RATE suggests.

Nate Silver is a renowned statistical analyst who was named one of "The World's 100 Most Influential People" by Time Magazine in 2009. He gained acclaim for outperforming the polls in the 2008 U.S. presidential elections and created baseball's popular predictive system, PECOTA.
