Thomas M. Gray, III
Bachelor of Arts

The Rhetoric of the Bowl Championship Series: An Analysis of the Persuasive Methods Underlying an "Objective" System

By Thom Gray
University of Tennessee Chancellor's Honors Program
2007 Graduate in English (Rhetoric and Writing)

Introduction

Since I began my formal education in 1990, I have been caught in the middle of the "technological revolution." I have watched the computer grow from a tool for playing *The Oregon Trail* between lessons to a vital part of every student's life from middle school to the university. I have watched the Internet encourage worldwide communication among diverse cultures and help us spread and gather information faster than ever before.

I have also watched computers make my entire generation quite intellectually lazy.

The Internet in particular has made two large contributions to this trend. The first is that my generation is learning to argue in extremely short form. We tend to substitute one-minute blog posts and blurbs on message boards for well-considered and persuasive communication. To any professor of English and especially of rhetoric, the problems with this are fairly obvious even if they do not directly affect classroom learning.

The Internet's second contribution to intellectual laziness, however, is less obvious, and the fact that it is a potentially negative development seems to go unnoticed. Because quantitative data is so readily available and its analysis has become so automatic, my generation has learned to value numbers and figures far more than thoughts expressed in written or verbal forms. On the Internet, human problems are reduced to stat plots, and entire issues are debated using nothing but charts and graphs. The danger, as I see it, is that my generation is taking numerical arguments at face value far too often, failing to see the underlying assumptions – and the rhetoric – which makes those numbers significant in the first place.

Nowhere is the "computerization" of rhetorical principles more obvious to me than in the computer polls used in the Bowl Championship Series system in college football. The system purports to be a relatively "objective" evaluation of the best teams in football and is thus used to assign the top two teams to the National Championship game. People who disagree with its conclusions are often dismissed as "whiners" or as biased towards one team or another.

What no one seems to question openly when critiquing the BCS or with any other mathematical system is the dubiousness of a claim to "objectivity" made by any system that is written by and for humans.

While the BCS is somewhat of a light-hearted example of "computerized rhetoric," I hope that my analysis of its underlying rhetorical assumptions raises a few questions in the reader's mind about the rhetoric underlying all manner of computer analyses in our world today. I further hope that the Rhetoric faculty here at UTK can find ways to engage students in similar explorations of other purportedly "objective" or "mathematical" systems in the classroom, so that the importance and elegance of verbal rhetoric does not get lost in a maze of cold computer models.

The Rhetoric of the Bowl Championship Series: An Analysis of the Persuasive Methods Underlying an "Objective" Mathematical System

By Thom Gray

I. How College Football Got the Bowl Championship Series

The central question of every field, every art, is the same: what, or who, is "best?" The more important question is often, how do we determine what is "best?" Plato argued that a quorum of wise men should be able to arrive at the best ideas and decisions through deductive reasoning and argumentation. Essentially, when competing ideas are pitted against one another in an "intellectual arena" of sorts, one will ultimately dominate the others.

However, in a pure democracy like early Athens or in a democratic republic like America, the question of "best" is more often a question of convincing others of what (or in the case of an election, who) is the best. Doing so requires identifying and meeting individuals' or groups' criteria.

And therein lies the usefulness, function, and necessity of rhetoric in civic discussion: a person educated in and aware of rhetorical principles and strategies can find those means of persuasion and tailor his or her arguments to match an audience's needs. This is the central usefulness of rhetoric to the educated individual: the ability to make one's thoughts relevant to others.

All of the above is self-evident in most fields, especially law, politics, and philosophy. What is not so self-evident is the place rhetoric has in sports.

After all, sports are an objective activity, aren't they? Two teams meet on a field, and one wins. Some leagues like Major League Baseball or the NBA even stretch the contests to five or seven games to "check their answers," so to speak. By the end of the

game, there is no need to persuade anyone of a point of view. The answer as to who is the "best" between two teams can typically be settled on the field, a much more literal arena than the metaphorical "arena of ideas," and a much more effective one.

Or is it?

Consider the case of football in the National Collegiate Athletic Association (NCAA). In the Football Bowl Subdivision (also called Division I-A) alone, there are 119 teams all vying for the same National Championship, the same title of "the best team." In order to truly determine who is the "best" among these teams based on a team's record alone, every team would have to play 118 opponents – 6786 games every season (Colley 6).

But because football is an extremely physical game, the regular season for college football is only 14 weeks long, and every team takes at least one of those weeks off. Thus, a Division I-A team can only play a maximum of 13 of the other 118 Division I-A teams during the regular season – about 11.7 percent of the entire Division. A perfect record against those teams would still mean that a team had not been properly judged against 88.3% of the other teams. Therefore, if college football is supposed to award the National Championship to the "best" team out of all the Division I-A teams, then declaring a National Champion based on the teams' regular season records alone would be inaccurate at best and unfair at worst.

The dilemma for college football, then, is that even an undefeated team cannot truly claim to be the "best," nor can a team with very few losses truly be out of contention for the title. Even a playoff system would often force the arbitrary exclusion

of some teams with the exact same overall record as others, simply because a playoff system requires an even number of teams.

So how does the college football world solve this dilemma? In the past, they used the same method as a classical republic: a measured vote among an educated few.

In 1936, the sports writers of the Associated Press began ranking teams among themselves via a poll, in which teams would be granted points based on how individual writers ranked them. The team with the most points in the final poll of the season would be declared the AP's National Champion. Polls of this kind had existed all over the country for several decades, but the results of the AP poll gained much greater exposure – and a degree of national prestige – because of the AP's national distribution (Billingsley 1).

The situation became more complicated, however, in 1950, when United Press International (UPI) began publishing a poll taken among a selection of Division I-A coaches. As one might expect, it did not take long for the two polls to disagree on which team was "best." In 1954, three teams went undefeated in Division I-A college football: Ohio State, UCLA, and Oklahoma. The coaches picked UCLA as the National Champion, while the AP writers chose Oklahoma (Billingsley 1).

The two polls would disagree nine more times between 1954 and 1997 ("NCAA College Football National Champions" 1). Many of the splits originated from varying results in the "Bowl games," invitational games supported by different college football conferences, which would pit one conference's champion against another's champion, or against another team that was invited to play because of its achievements or popularity. Though one Bowl would often try to invite the #1 and #2 teams in the country, they often

failed to produce such a true "National Championship" game because of conference loyalties and the need to sell tickets, both of which favored popular teams over "better" teams. As a result, the question of who was "best" was usually left solely to a vote, without an objective way to judge which team was superior. Thus, the National Championship was often decided by subjective means.

Naturally, a vote about who was "best" in an activity which is by its very nature tied to objective means of judging teams – actual head-to-head contests – was often dissatisfying. Critics claimed that the polls were subject to the individual biases of the sports writers and the coaches. They argued that coaches would favor their own teams and teams in their conferences, that writers would be subject to regional biases, and that tradition and self-interest could sway both of the major polls substantially.

Also suspect was the fact that coaches and writers, who were busy with their own teams on Saturdays, were ranking teams that they could ostensibly never even watch on television or in person. This became a particular problem when multiple teams went undefeated or finished with a single loss and coaches and writers were left with no single objective way to decide which of the tied teams was truly "best."

Another complaint about the human polls was that the writers and coaches were well aware of each others' polls, and furthermore of each other as individuals. There was therefore a risk of a false consensus among the two polls and among the participants of each poll, not only because of a sort of group-think effect, but because often the very writers and coaches who voted in the polls could persuade other voters of their opinions – that is, they could use rhetoric on them. According to critics, this decreased the objectivity and legitimacy of the polls in determining a true champion, especially when

several teams were clustered near the top of the rankings. They essentially argued that as a means of determining a champion, rhetoric was inferior to deductive or objective means – that is, playing a definitive National Championship game.

During the 1990s, the administrations of the four major Bowl games – the Orange Bowl, the Sugar Bowl, the Fiesta Bowl, and the Rose Bowl – had a series of loose alliances in which they attempted to pair the top two teams in the nation in a single game every year, thus providing a somewhat definitive answer to who was "the best" ("BCS Chronology" 1). However, that process was hampered by the Bowls' insistence on retaining their original ties to individual conferences: if two undefeated conference champions were courted by two different Bowls, they would naturally never decide the Championship on the field. Furthermore, the "two best teams" and the final National Champion were still subject to the same supposedly arbitrary votes that had been criticized for several decades.

When Michigan and Nebraska both finished undefeated during the 1997 regular season and won their conferences' respective bowl games, the two major polls once again split on the issue of the National Championship: the AP poll chose 12-0 Michigan as their National Champion, while the coaches' poll (now conducted by *USA Today*) chose 13-0 Nebraska ("NCAA College Football National Champions" 1).

Frustration with the two-poll system and the loose Bowl alliance finally hit a breaking point, and the pollsters, NCAA executives, and the executives for the various bowl games all came together to create a united system – a sort of "new constitution" for college football. The result was the Bowl Championship Series system.

II. What is the Bowl Championship Series?

The Bowl Championship Series (BCS) started in 1998, when the four major Bowls decided to host the top two teams in college football in a single game every season ("BCS Chronology" 1). The four bowls would host the National Championship game on a rotating basis, and when a Bowl was slated to host the National Championship game, it would forego its traditional conference ties in favor of the two best teams regardless of conference. The other three Bowls would continue hosting the champions from the conferences they had traditionally represented and would pair them against three "at large" teams from the BCS' member conferences. Among the teams which did not play for the National Title, the Southeastern Conference (SEC) Champion would earn a bid to the Sugar Bowl, the Big 12 champion would play in the Fiesta Bowl, the Atlantic Coast Conference (ACC) or Big East champion would play in the Orange Bowl, and the Pacific 10 (Pac-10) or Big Ten champion, or both, would play in the Rose Bowl. The key was that each of those six conferences could guarantee its champion a major Bowl Game every year, while the two best teams overall, regardless of their conferences, would play for the National Title.

Of course, the central question still remained: how would the Bowls decide who the top two teams were?

The answer was to create something akin to a combination of a republic and a technocracy. The BCS would rely on the two major nationwide polls – the AP writers' poll and the *USA Today* coaches' poll – but would also judge the teams with purely mathematical computer rankings. The idea was to balance the assumed expertise of *USA*

Today's coaches and the AP's sports writers with computerized rankings based on pure data, theoretically "objective" measures of teams' talent. Thus, the BCS could claim that it matched the top two teams against each other every year, and that their status as the "top two teams" could be verified objectively. After all, how could computer programs suffer from conference loyalty, regional biases, or any of the other persuasive pitfalls that allegedly plagued the "human" polls?

The first year of the BCS system, the 1998 regular season, was largely considered a success, at least as far as determining a National Champion was concerned. Undefeated Tennessee (12-0), the Southeastern Conference champion, met ACC champion Florida State (11-1) in the Fiesta Bowl, and Tennessee emerged victorious as the lone undefeated team in the nation ("Bowl Championship Series Final 1998-1999" 1).

But two seasons later, following the 2000 regular season, controversy arose when one-loss Florida State was chosen to play against undefeated Oklahoma for the National Championship. The reason that choice was controversial was because Florida State's one loss was to Miami (Fla.), another one-loss team that was ranked #2 in both human polls at the end of the season ("Bowl Championship Series Final 2000-2001" 1). In the human pollsters' minds, Miami had proven objectively that they were a superior team to Florida State by defeating them on the field. But because both teams merely appeared as sets of numbers to the computer polls, the computers still ranked Florida State ahead of Miami based on equally objective numerical factors, namely the difficulty of their schedule and their margin of victory over their opponents.

The controversy deepened when Miami won its Bowl game handily, but Florida State lost to Oklahoma 13-2 in the Orange Bowl ("Bowl Championship Series All-Time

Results" 1). Miami finished with one loss and Florida State finished with two, seemingly verifying the human voters' assertion that Miami was a better team than Florida State.

The next year, after the 2001 regular season, a similar situation occurred when the computers chose one-loss Nebraska to play undefeated Miami (Fla.) in the Rose Bowl ("Bowl Championship Series Final 2001-2002" 1). Again, the #2 team from both human polls, the University of Oregon, was ruled out by the computers in favor of Nebraska, whose one loss had come at the end of their season against Colorado. That loss had forced Nebraska out of the Big 12 conference championship. The computer rankings at the time were specifically designed to exclude conference names and to weight conference games equally with other games in their calculations, to the computers still ranked Nebraska as the second-best team based on the raw numbers. The human voters, however, traditionally believed that a team should at least win its own conference in order to earn a chance at a National Title, and that a team that lost near the end of the regular season could not be superior to a team that played and won its best games at the end of the year. Oregon had done both by winning the Pac-10 overall, while Nebraska had not.

Once again, the human polls were somewhat vindicated when Nebraska lost the National Title 37-14, while Oregon defeated Colorado, the team that had already defeated Nebraska ("Bowl Championship Series All-Time Results" 1). The "objective test" on the field had determined a significantly different order than the computers' "objective" mathematic analyses.

In an even more confusing scenario, no teams finished undefeated in Division I-A college football in 2003, forcing both the human polls and the computers to choose the

top two teams among six one-loss teams, none of whom had faced each other in the regular season: the University of Oklahoma, Louisiana State University (LSU), the University of Southern California (USC), Miami University (AKA Miami of Ohio), Boise State University, and Texas Christian University (TCU). Miami of Ohio, Boise State, and TCU were all essentially eliminated from contention simply because none of them were members of the "automatic bid" conferences in the BCS. Oklahoma was the #1 team and undefeated until the last weak of the year, when they lost the Big 12 conference championship to Kansas State. As a result, Oklahoma lost their automatic bid to the BCS and fell to #3 in both human polls as a result ("Bowl Championship Series Final 2003-2004" 1). But the computers still chose Oklahoma to play against LSU in the Sugar Bowl. Again, neither Oklahoma's late-year loss nor losing their conference championship seemed to matter to the computers, while it was a major criterion for the human voters.

LSU defeated Oklahoma 21-14. Meanwhile, USC won the Rose Bowl against Michigan ("Bowl Championship Series All-Time Results" 1). The result was that LSU and USC both finished with one loss in BCS member conferences, but never met, while Oklahoma got to play for the National Championship, but finished with two losses and no conference title.

The final outcome of this confusing scenario in which human and computer criteria clashed badly was exactly what the BCS system was designed to prevent: a split championship in the polls. LSU finished #1 in the coaches' poll, while USC finished #1 in the AP writers' poll.

During the next season, the 2004 regular season, the BCS faced an even bigger controversy when five teams finished undefeated and won their respective conferences: USC, Oklahoma, Auburn University, the University of Utah, and Boise State ("Bowl Championship Series Final 2004-2005" 1). Again, Utah and Boise State were essentially eliminated from contention because they were not affiliated with any of the Bowl-affiliated conferences.

This time, it was the human pollsters who ended up choosing the wrong National Title matchup. The computers ranked Auburn #2 behind USC because the computers estimated their schedule to be tougher than Oklahoma's. However, Auburn was excluded from the National Championship game because USC and Oklahoma had been the top two teams in the preseason human polls ("2004 College Football Rankings" 1), making it nearly impossible for Auburn to jump either team in the final polls.

USC handily defeated Oklahoma in the Orange Bowl 55-19, while Auburn went on to win the Sugar Bowl and finish the season undefeated. Utah became the first team outside of the BCS conferences to earn an "at large" bid in the BCS and won the Fiesta Bowl against Pittsburgh 35-7 ("Bowl Championship Series All-Time Results" 1). After the Bowl games, USC finished #1 in every poll while Auburn finished #2 – but the two teams would never meet each other or Utah to determine which undefeated team was truly the best ("2004 College Football Rankings" 17). The controversy was so heated that the Associated Press refused to be a part of the system anymore, forcing the BCS to replace the AP writers' poll with a poll taken among a panel of college football experts by Harris Interactive ("BCS Chronology" 1).

And to complicate matters even further, the computer average had ranked Oklahoma above USC and Auburn before the National Championship game was even played, taking even more legitimacy away from the computers' evaluations of the teams ("Bowl Championship Series Final 2004-2005" 1). Somehow, the computers had accurately picked the #2 team in the country, but had been entirely wrong about who was #1.

The BCS was learning the hard way that achieving an objective result in a competitive world was not as easy as it seemed. The irony was that in many cases, the supposedly objective measures of who was the "best" team – the computer rankings – were the ones which were skewing the system out of sync with apparent football reality – the actual results of the Bowl games. What was it about the computer rankings which made them so fallible, so "human?"

The answer: just like other humans, the computer polls' were designed by people who were – and still are – subject to the same rhetoric as the writers and coaches and who judged teams using the same or similar data. As a result, their "objective" computer systems were not purely mathematical in nature, but responded to the same rhetorical principles that swayed their human counterparts. That rhetoric was simply being transformed into mathematical descriptions and evaluations instead of verbal ones. The computer rankings did not eliminate the rhetoric from the BCS system, but merely shifted its focus towards pure *logos*.

As the brief history of the BCS computer polls demonstrates, human beings designed the computer polls to satisfy human desires and to solve human debates by re-

interpreting the same data used to persuade humans. The results have often resembled the very same human controversies the BCS was intended to eliminate.

III. Minds Behind the Machines: How the BCS Computerizes Rhetoric

The BCS was supposed to match up the two best teams in college football in the same Bowl game every year. The problem was that before 1998, the primary methods for determining the "best two teams" were ultimately opinion polls. In the eyes of the college football world, these polls were subject to the individual and collective biases of the people involved. The BCS's solution was to judge teams not just by experts' opinions, but by mathematical rankings executed by computer programs. The assumption was that computers were supposedly not subject to any biases, and thus could theoretically measure a team's ability more objectively than humans could.

While this assumption might seem logical, it becomes immediately suspect when one considers that the BCS did not use a single mathematical formula to determine the contenders for championships, but instead started with three *different* computer formulas. During its second year, 1999, the BCS even added five more formulas for a total of eight. Since 2002, the BCS has incorporated six computer polls with six different formulas for determining the "best" teams ("BCS Chronology" 1).

If the computers were supposed to evaluate the teams "objectively" – to provide a single answer to the question of, "who's best?" – then why were there multiple computer polls, and why were those polls changed so frequently? What the BCS seemed to gloss over in their claim to objectivity was that behind every "objective" computer was a

subjective human being, made of the same flesh and bone as the voters in the human polls and subject to the same persuasion.

The BCS currently uses six computer ranking systems designed by six people from quite different fields: Dr. Wesley Colley, a mathematics professor at the University of Alabama at Huntsville; Jeff Anderson and Chris Hester, writers for the *Seattle Times*; Richard Billingsley, president of the College Football Research Center; Jeff Sagarin, a sports statistician for *USA Today*; Kenneth Massey, a mathematics professor at Carson-Newman; and Dr. Peter Wolfe, a professor at the University of California at Los Angeles.

As noted above, not all of these individuals have been involved with the BCD since its inception. Of the six current computer polls in the BCS, only Jeff Sagarin's poll and the *Seattle Times* poll by Anderson and Hester have been part of the system since the BCS was formed in 1997 (the other original poll, the *New York Times* poll, was dropped in 2001). Massey's poll and Billingsley's were added in 1999, and Wolfe's and Colley's programs were added in 2001 ("BCS Chronology" 1).

There is a case to be made that these polls were added at different times to lend an increasingly "academic" *ethos* to the BCS system. The three original computer polls were all written and conducted by sports writers, which made them more comfortable options since the Associated Press' sports writers had long been an authority in the debate over the National Championship. Billingsley, whose poll was added in the first expansion of the BCS, is a sort of "superfan" with a wealth of football knowledge but no academic credentials to speak of; thus, he was not only accessible to the college sports world, but also authoritative within it. Massey, on the other hand, was a doctoral student at Virginia Tech when his model was introduced into the BCS; this gave him ties to both a traditional

college football powerhouse and to the loftier world of academia (Rovell 1). Wolfe and Colley added even more intellectual firepower to the system when they joined the BCS in 2001: both have PhDs in mathematics with specialties in computer modeling, and both have produced long mathematical justifications for their formulas.

But behind the degrees and the football pedigrees, the reasoning for these six formulas' inclusion in the BCS rankings lies in the formulas themselves, and more importantly in the philosophies underlying those formulas. Some of the differences are ultimately minor ones. For instance, Peter Wolfe's program ranks and considers all four-year college football programs in the United States (Wolfe 1), while Colley's relies only on Division I-A schools (Colley 1). This reflects a difference in these individuals' presumptions and beliefs: Colley values I-A schools more and thus rewards BCS teams for playing and defeating more I-A programs, while Wolfe seems to believe that every win is worth consideration when evaluating teams.

But in some of the computer polls' other criteria– or more accurately, their designers' – we find a wealth of debate about which criteria should ultimate decide who is "best" in college football, a wealth of competing criteria for "the best team," much of which is driven by the many fiascos which have already plagued the BCS in its short lifespan.

The most substantial criteria of – and the most controversial differences among – the BCS computer polls involve the following:

1.) Strength of schedule and conference strength

2.) Preseason rankings

3.) Margin and time of victory

What is most fascinating about these criteria is that they embody and even reiterate some of the very same problematic "biases" in the human polls that the BCS was supposed to counteract. The logic behind the computers' major differing criteria can ultimately be traced to some appeal to *ethos, pathos,* or adjusted *logos* which also plays a part in the human polls, further calling into question the theoretical "objectivity" of the computer rankings.

IV. "It's Who You Play": Strength of Schedule and the Ethos of Conference Strength

One of the stated reasons that college football uses polls and not raw records to determine its top teams is the shear number of teams in Division I-A and the natural disparity among those teams. What complicates the situation even more is that the teams are also divided into conferences based on region, size, and tradition. This means that a large part of each team's schedule is comprised of the same few teams every year. This makes teams' schedules even less representative of the entire Division than random schedules might be. The computer pollsters' solution is not only to rank teams by their wins and losses, but also to adjust those raw numbers based on the strength of the other schools that a team plays. This calculation is known as a team's "strength of schedule."

This is more of a challenge than one might think, because this means that the computers are compensating for unbalanced schedules by ranking a team's results against other teams who also play unbalanced schedules. Therefore, a team cannot merely be judged by its opponents' combined wins and losses, either, since that is not a much more accurate reflection of a team's strength than its own combined wins and losses.

The *Seattle Times* system (Anderson and Hester) and Massey's rankings solve this problem by factoring in opponents' "conference strength," the calculated total talent within a team's conference. This allows the computers to judge the majority of an opponent's competition all at once, since most teams play a majority of their schedules within their geographic or traditional conference. Both systems calculate a conference's strength based on the teams' wins and losses against non-conference opponents (wins and losses within the conference would naturally cancel each other out). The *Seattle Times* then adjusts its conference strength rating to reflect the strength of a conference's non-conference opponents (Anderson 1). Massey's ratings, however, factor in the ratings for the teams within the conference instead, then adjust the average rating within the conference for the "parity" (top-to-bottom consistency) within the conference (Massey 1). Thus, Anderson and Hester reward conferences that schedule tough non-conference games consistently, while Massey's system promotes conferences that field better teams within their own conference.

Ranking teams based on their conferences, though, presents an interesting problem. The computer rankings are supposed to eliminate traditional biases and geographical attachments that exist in the human polls – a team's football *ethos*. But one of the most evident traditional biases is conference loyalty or the stigmas attached to certain conferences. Rewarding a team because of its conference actually *emphasizes* the conference-associated *ethos* often used by human voters, one of the very "subjective" measures that the computer polls were designed to counteract.

Furthermore, since the BCS system already rewards teams from certain conferences with its six automatic bids for conference champions, the polls that rate

teams based on conference strength give a sort of "double-bonus" to teams that already benefit from strong traditional ties to the major conferences. This has manifested itself in years like 2004, when undefeated teams from Utah and Boise State found themselves automatically out of contention for a National Championship because they did not participate in major conferences. An even stronger example of this conundrum was the 2006 season, when Boise State was not considered for the title game even though they were one of only two undefeated teams in Division I-A; in fact, they won a Bowl game with an at-large bid and finished the year as the only undefeated team in Division I-A ("2006-2007 College Football Season" 1). The *ethos* that the computer rankings are supposed to omit is apparently inherent in the BCS' associations, and some of the computers simply reward teams for playing in major conferences even more.

As an alternative to using conference strength to rate teams' schedules, two of the six BCS polls – Wolfe's and Colley's – instead calculate a team's strength of schedule by taking into consideration its individual opponent's strength of schedule (Wolfe 1, Colley 6). A win over a two-loss team with the fifth-strongest schedule in the country, for instance, is worth more in these systems than a win over a two-loss team with the fiftieth-strongest schedule. The effect of this rating is cumulative throughout the season: the most accurate reflection of a team's opponents' strength of schedule is produced at the very end of the season when every team has played all its games, and when, after all a team's ranking counts the most.

History has shown, however, that this method can create very counterintuitive results. Consider, for instance, the 2000 season. Florida State and Miami were both one-loss teams, but Miami had defeated Florida State in a head-to-head contest. The human

polls took this into consideration and named Miami as the #2 team behind undefeated Oklahoma. The computers instead selected Florida State because of Florida State's strength of opposition rating. The math is tricky, but the logic is this:

Florida State lost to Miami, a one-loss team, which was a "good loss." Miami had lost to Washington, which was also a one-loss team; that loss was also a "good loss." But Washington had lost to Oregon, a two-loss team, which made their loss a somewhat "bad loss." By proximity, Florida State had the "best loss," followed by Miami, followed by Washington. Thus, the head-to-head victory for Miami was irrelevant to the computer formulas: Miami was not "Miami" to the computers, but a team one fewer spot away from a loss to a two-loss team.

By removing the *ethos* of the name "Miami" from the evaluation, the computers also removed the best objective measure of a team's superiority: a head-to-head match-up. The goal here was to remove a rhetorical element, but removing that element did not increase the polls' objective correctness.

Another problem with this method is that, as previously mentioned, it does not fully take into consideration a team's strength of schedule until the very end of the season. As even a causal observer of sports might notice, a team's raw talent and skill level can change drastically over the course of a season because of injuries, changes in personnel, coaching changes, and the like. One team might defeat the best team in the nation in September, but that "best team" may suffer injuries to star players that drop them to the 20^{th} by the end of the season. In that case, the team that defeated the "best team" will not get as much credit for the victory at the end of the season, because the

computers never see that opponent as the "best team," but as the 20th-best team. Human polls can adjust for this shift in *ethos*; the computer polls typically do not.

Richard Billingsley's rankings attempt to adjust for a team's changing inter-organizational dynamics as the season progresses by rewarding a team for defeating a "rank" and not the team that occupies that spot on any given Saturday. Billingsley's computers calculate a team's opponent's strength on the day of a game and credit the team with a win or a loss over that position, not just that team. For instance, a win over a #1 team is a win over a #1 team even if the top spot in Billingsley's poll is occupied by someone else at the end of a season. In this way, Billingsley's poll removes makes each team's post-season *ethos* irrelevant and emphasizes a team's *ethos* from week to week (Billingsley 1).

The problem here, of course, is that in order for Billingsley's method to be an accurate measure of an opponent's strength, his rankings must also be accurate on the day of the game, just as the other computer polls assume that their own methods are sufficient to rank a team's opponents or conferences. This is a particularly troublesome problem at the beginning of the season, when fans may already demand an evaluation of the teams, but the computers lack the data necessary to provide one.

This brings us to another hot-button issue with the human polls that some of the computer polls purport to address: preseason rankings.

V. "It's Where You Start": The Ethos of the Preseason Rankings

Humans are often unwilling to wait for data to accumulate before forming opinions. This peculiarity even extends to sports, where accumulating data – that is, watching the games – is supposedly the best part of fandom

In the case of college football, fans and experts alike want to debate the merits of the teams before the season ever begins. The result is that several outlets, including the Associated Press, *USA Today*'s coaches' poll, and the Harris poll all release preseason polls for college football. Teams that start atop these polls are favored for the National Championship game before the season even begins. Teams ranked lower in the preseason poll have an uphill climb to the top two spots and, according to some critics, have to earn "style points" to impress the voters even if they end up with the same record as the top preseason teams.

Preseason polls became a major problem in the BCS in 2004, when USC, Oklahoma, and Auburn all finished the regular season undefeated. According to the computers, Auburn had a better strength of schedule, which earned them an average ranking of #2 in the computer polls behind #1 USC ("Bowl Championship Series Final 2004-2005" 1). But the human polls favored USC and Oklahoma because they had been the top two teams in the preseason rankings ("2004 College Football Rankings" 1). In the postseason, USC defeated Oklahoma handily for the National Title, and Auburn won their bowl game and finished undefeated overall ("2004 College Football Rankings" 17). The computers' evaluation of Auburn – and many fans' – was vindicated by the postseason results.

The preseason rankings are one of the most prominent examples of how a team's preexisting *ethos* directly affects their ranking in the BCS. Teams from "big name"

programs with storied histories, such as Notre Dame, tend to fare better in these preseason polls regardless of their recent successes or failures. Newer "powerhouse" teams like Boise State, however, often start off low or even unranked in the preseason polls, virtually shutting those teams out of a National Championship game before the season even begins.

Four of the BCS computer polls – Sagarin's, the *Seattle Times*', Wolfe's, and Colley's – do not start with any preseason rankings whatsoever. Instead, they rank teams as the season progresses and become more accurate as they incorporate more data from more games (Sagarin 1, Anderson 1, Wolfe 1, Colley 3). The problem is that these polls can thus be very chaotic and inaccurate or even impossible to calculate early in the college football season. Anderson and Hester, for instance, do not even release their first rankings until the fifth week of the season (Anderson 1), and the cumulative average for the BCS computer polls is not released until the sixth week of each season. This can result in minor controversies when the early computer polls are too divergent from the human polls. The human polls may artificially align themselves with the computer polls and skew the entire system, just as the human polls allegedly created artificial consensuses with each other before the BCS implemented its current system.

These four polls also operate on the assumption that previous success really isn't a qualification for a National Champion. For instance, they do not reward programs for factors as season-to-season consistency, building great teams over time, or other measures of long-term success. Writers and coaches alike might argue that teams who were successful in previous seasons really do deserve have a legitimate advantage in the current season; for instance, a one-loss team that returns all its starting players the next

season probably has an advantage over a teams with new starters at key positions. In this way, the historical *ethos* of a program can in fact be a very logical way to consider their chances of success before the season begins.

To account for such advantages, Billingsley and Massey both carry over a team's ranking from the previous year and use it as a starting point for their polls. Billingsley's rankings before the first week of the season are the exact same as his rankings at the end of the previous season. This is especially important in Billingsley's strength of schedule ratings, as he uses his own rankings to determine teams' strengths of schedule from week to week. Billingsley considers a win against a highly-ranked team to be important no matter where that team is ranked at the end of the season (Billingsley 1).

Unlike Billingsley, Massey does not transfer his rankings directly from the end of one season to the beginning of the next. Instead, he uses a "weighted average" of the previous year's ratings – numerical calculations of a team's strength – to estimate a team's strength at the beginning of the next season. Successful teams gain a slight but not overpowering advantage from year to year, while unsuccessful teams are not overly handicapped because of their prior problems (Massey 1).

Both of these methods for ranking teams in the preseason are purely numerical. They do not adjust for the same factors as the preseason human polls, since those polls often judge a team by its human composition from season to season (that is, the players and coaches who leave, return, or are added to the team). The preseason computer polls also do not take into account a team's entire history as human polls allegedly do; instead, only a team's most recent success is rewarded. But the use of even recent historical data to rank teams early in the season is still rooted in *ethos* and not *logos*, since preseason

rankings inherently give teams an advantage in the rankings before they earn that advantage on the field – before the concrete facts show evidence of their superiority for that season.

Again, the computer polls disagree about one of the persuasive concerns they were all intended to eliminate. Two of the computer polls even add their own spin on preseason polling by using strictly mathematical criteria. But whether preseason rankings involve numbers or judgments about personnel, they are still rooted in teams' historical *ethos* and not their on-field *logos*. The assumption that a team's players and coaches will perform the same way is equivalent to the assumption that a team's statistical performances will remain comparable. Both are subjective until they are judged alongside teams' actual performances.

As a third example of a controversy within the BCS computer polls demonstrates, sometimes even raw game scores have a subjective element. In the past, the BCS has even tweaked the math in its computer polls to reflect a completely subjective value – here, sportsmanship – in its supposedly objective system, and has sacrificed a measure of objectivity to do so.

VI. "It's How Much You Score": The Pathos of Margin of Victory

Unbalanced schedules and changes in teams' composition are only two of the many problems with a "wins only" determination of the best teams in college football. Another is that while winning is the major objective in all sports, not all wins are equal, even against equal teams. If, for instance, Auburn beats South Carolina 23-17, and

Tennessee beats South Carolina 49-3, it would be fair to say that Tennessee is better than Auburn, especially if those results are consistent with Tennessee's other victories. Tennessee's higher score would indicate a stronger offense, and South Carolina's lower score versus Tennessee would also indicate that Tennessee has a better defense than Auburn.

Because humans naturally cannot view every single football game, these *prima facie* evaluations of final scores factor somewhat heavily into the human polls. If, for example, Auburn beat South Carolina by six points while Tennessee beat Kentucky by 35 points in the same week, and South Carolina and Kentucky were roughly equivalent teams, then Tennessee would probably be ranked ahead of Auburn in the human polls if Tennessee and Auburn had the same overall record. At the end of the season, Tennessee might be considered better than Auburn if they had equivalent records, but Tennessee had beaten its opponents by an average of 14 points while Auburn had only won by an average margin of three points.

However, ranking teams based on scoring margin has two key downsides in human polls, downsides which carry over to computer polls. The first is that final scores do not always reflect how close a game was or how dominant one team was over the other. For instance, if a team scores 35 points against their opponent in the first half, they may substitute their starting players for backups in the second half to avoid injuries, since it is unlikely that the other team will be able to come back and win the game. Likewise, a game could be hotly contested throughout and feature a score of 28-24 in the final seconds of the game; the leading team may then score an incidental touchdown and kick an extra point as time expires. In this example, the final score of each game would be 35-

24, but one victor was clearly more dominant over its opponent than the other. However, human pollsters cannot possibly watch every single game, and computers do not watch the games at all. Thus, the victories look comparable on paper and in equations when they were not quite so comparable on the field.

The second problem with margin of victory is related to the first, as teams tried to manipulate the occasional discrepancy between the final score and the winning team's actual skill. Before the BCS went into effect, teams were sometimes accused of leaving their starting players in games that were well in hand and using aggressive offensive play-calling to "run up the score." This resulted in runaway victories against weaker teams which looked more impressive to human pollsters than they actually were.

When the BCS first formed, it incorporated the margin of victory into its computer rankings. Many people felt that this worsened the problem of runaway scores by removing subjective interpretations of the final scores, which were sometimes appropriate for the biggest "victories." Human pollsters may have been able to see how abnormal a score of 54-0 was in a football game, for instance, but computers could only interpret raw numbers and would reward teams even when their scoring margins were "too big."

But the issue with running up the score is not just an issue of skill per se. After all, if a team is capable of scoring an extremely high number of points while holding its opponents to very few, that team probably quite good even if the opposing team is relatively weak. According to critics, the problem with a team's deliberately running up the score on its opponent is that it demonstrates bad sportsmanship. The willingness to embarrass an opponent is an especially unfavorable attribute in a champion in any sports.

This is especially true in football, where forcing the other team to continue playing at normal speed and intensity can lead to serious injuries, and where "running the clock out" at the end of a game is usually a safer option.

Margin of victory was one of the factors blamed for the BCS controversies in 2000 and 2001, since Florida State and Nebraska allegedly had several "runaway victories" which inflated their computer rankings. In 2002, the BCS informed its computer pollsters that in order to continue participating in the BCS computer rankings, they would have to remove the margin of victory from consideration in their polls ("BCS Chronology" 1). There was significant disagreement over this decision. Two computer pollsters, Dave Rothman and Herman Matthews, refused to change their formulas, and their rankings were removed from the BCS' calculations. By contrast, Kenneth Massey and Richard Billingsley simply changed their formulas to match the BCS' new standards (Lindsay 1).

If margin of victory was such a controversial statistic, then why were two pollsters willing to be excluded from the BCS in order to keep it? Though he does not incorporate point spreads into his BCS rankings, Jeff Sagarin does rank teams separately based on a "pure points" system, and he explains his logic quite succinctly: "Pure Points…is the best single predictor of future games" (Sagarin 1).

Other pollsters in the current BCS seem to recognize this fact as well. Even though they changed the way they consider margins of victory, both Massey and Billingsley still use game scores in their formulas. A team's total points are calculated as its "offensive power" and the points it allows are recorded as "defensive power." Massey weights teams' rankings based on their cumulative offensive and defensive powers in

order to predict games (Massey 1). Billingsley uses only a team's defensive performance from the current season to predict winners ahead of time (Billingsley 1).

So the BCS leaves us with an odd paradox regarding the margin of victory. Pundits seem to agree that teams should not be rewarded for their margins of victory as pure numbers. They also agree that teams' ability to score and to prevent the other team from scoring is one of the best ways to evaluate their performances. But the subjective ramification of the margin of victory – the potential for bad sportsmanship – is still considered unsavory enough to omit that statistic from the BCS. In other words, one of the strongest objective criteria for evaluating teams must be thrown out because of an ultimately emotional appeal to sportsmanship.

Thus, a critical *pathos* appeal directly affects the computer rankings. Humans do not allow the computers to consider a strong objective evaluation because of a subjective value, thus illustrating once again the critical flaw in assuming that the BCS computers can be "objective" at all.

VII. So Who Really *Is* Best in the BCS?

So what really *is* the best way to determine a National Champion? Which poll, computer or human, is most consistently correct as it tries to predict which team will ultimately win the biggest game of all? Can we objectively determine a National Champion before a National Championship game is played? Can we objectively determine a National Champion even *after* a National Championship game is played?

The short answer is that based on the failures of the current polls and models, we cannot. A survey of the individual computer rankings over time shows that they are consistently inconsistent, alternately picking unforeseen underdogs as winners of the National Title game and selecting teams for the National Title game who proved to be miserably inferior to their opponents. The computer polls compare neither favorably nor unfavorably to the human polls in that regard: they succeed and fail at similar rates. As discussed in Section II, the BCS system, which combines several different evaluations of the teams, has not been any more consistent than the individual polls thus far.

It seems that the BCS' supposedly objective method fails to produce an objective champion more often than it succeeds.

The problem with the computerized elements of the BCS is that they purport to remove the rhetoric from a naturally rhetorical exercise in a search for theoretical "objectivity." Instead, the system mere isolates the *logos* within a larger rhetorical debate. In the process, the BCS may exclude equally legitimate criteria for championship teams, criteria which simply fall into other rhetorical categories: the *ethos* of a team's history of success, the *pathos* of sportsmanship and fair play, even the teams' individual names and identities. They are forgetting that champions can not always, if ever, be reduced to numbers.

The result of the BCS' methodology is not objective at all, but simply a much drier – and in many cases, less accurate – form of rhetoric. Even the computers' evaluations of the teams still rely on *logos*, and the computers' definitions of the proper *logos* are just as divergent and subjective as their human designers' and editors'.

All this demonstrates that on some level, "objectivity" is not only nearly impossible to achieve in competitive realms like sports, but also unsatisfying for the competitors and their supporters. History demonstrates that even when the biggest games are played out and concrete, "official" winners rise above the din, humans simply cannot always agree on which team is the "best." What they can do is persuade others that their criteria for the best teams are the best criteria, or that their teams' credentials fit other individuals' own criteria. Where mathematics cannot declare a champion, rhetoric can create a champion in the hearts and minds of fans and experts alike.

Perhaps in the end, the true champion in college football is the team that makes the best case for itself on *and* off the field. Perhaps rhetoric is the *best* way to determine a championship. And perhaps the best teams will always differentiate themselves from one another through rhetoric more than through any other means. We should not be uncomfortable with this, but should embrace it as one of the things that keeps sports interesting and engaging.

Appendix: The "Other" Polls

While this study only considers the polls that are and have been part of the actual BCS formula, there are literally dozens of different individuals and agencies that rank college football teams with opinion polls and computer formulas. Kenneth Massey, one of the BCS computer pollsters, keeps track of over 100 of these polls on his College Football Rankings Comparison. Many of the other computer polls are conducted by mathematicians like Massey and claim to be supported by sound academic research.

It would have been a monumental and redundant task to construct a comprehensive comparison of every existing electronic ranking system for college football, which is why I decided to focus on the most relevant computer polls, the ones used in the actual BCS formua. But it is nonetheless instructive to note that these dissenting voices exist outside the BCS, and that their results sometimes vary greatly from the BCS' "consensus vote" even in relatively uncontroversial seasons.

For instance, after the 2006-2007 season, the BCS polls all agreed that the winner of the National Championship game, one-loss Florida, was the "true" National Champion. But among the polls on Massey's site alone one can find six different National Champions: Florida; Ohio State, the team that Florida defeated for the BCS title; Louisville, the one-loss winner of the Orange Bowl; USC, the two-loss winner of the Rose Bowl; LSU, the two-loss winner of the Sugar Bowl; and Boise State, the lone undefeated winner of the Fiesta Bowl. Perhaps more strikingly, one can also find polls that rank Florida as low as 16^{th} and Boise State as low as 22^{nd}. One poll even ranks USC 55^{th} – a far cry from their BCS rank of 5^{th}.

Several statisticians and computer experts test the various college football polls for accuracy versus game results or harmony with the final "consensus" rankings. I had hoped early in this process to identify a few polls that were consistently atop these "rankings of the rankings" and thus to identify the prototype for rhetoric in college football – the "perfect argument" for a National Champion.

After significant digging, I realized that there was apparently no such thing. Even the weekly BCS results consistently rank in the 20s in accuracy rankings.

Instead, I found that just as the polls had vastly different criteria and results, they also suffered varying accuracy problems from week to week and from year to year. I am glad I did the research, though, because I feel like my final thesis, and my realization about the vital multiplicity of persuasive criteria among humans and computers alike, was far more satisfying and useful.

I have included a few good sources for comparisons among college various football polls here. I will warn the reader that while the math these experts use is sound, interesting, and well-respected, their choices in website design leave something to be desired.

Resources for Comparisons Among the Rankings

Kenneth Massey's College Football Ranking Comparison:

Kenneth Massey lists how 112 different human and computer polls (number varies from week to week) rank all 119 Division I-A teams and tabulates the mean and median rankings of each team and the standard deviations among the polls.

http://www.masseyratings.com/cf/compare.htm

Eugene Potemkin's College Football Ranking of Rankings Systems:

Assuming that a higher-ranked team should beat a lower-ranked team in the same poll, Dr. Eugene Potemkin, a Russian mathematician, ranks the individual polls for accuracy over the season.

http://rsport.netorn.ru/cf/rankrank.htm

Wobus Sports' Football Ranking Correlations Summary:

Assuming that a consensus among Massey's 102 polls is close to a real ranking of the teams, John Wobus, a computer technician at Cornell University, tests the polls' accuracy from week to week.

http://www.vaporia.com/sports/fb-corr.html

WORKS CITED

"2004 College Football Rankings." Date unknown. ESPN.com. 1 May 2007. http://sports.espn.go.com/ncf/rankingsindex?pollId=null&weekNumber=1&seasonYear=2004

"2006-2007 College Football Season Final BCS Standings." 3 Dec 2006. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/bcs_standings_120306.html

Anderson, Jeff and Chris Hester. *The Anderson and Hester-Seattle Times College Football Computer Rankings.* 2007. AndersonSports.com. 1 May 2007. http://www.andersonsports.com/football/ACF_frnk.html

"BCS Chronology." 1 March 2007. BCSFootball.com. 1 May 2007. http://www.bcsfootball.org/bcsfb/history

"Bowl Championship Series All-Time Results." 1 March 2007. BCSFootball.com. 1 May 2007. http://www.bcsfootball.org/bcsfb/results

Billingsley, Richard. "The BCS Made the Right Move." 23 June 2002. ESPN.com. 1 May 2007. http://espn.go.com/ncf/s/2002/0622/1397882.html

Billingsley, Richard. "Understanding the Rating System." 1 March 2007. College Football Research Center. 1 May 2007. http://www.cfrc.com/Archives/In_Search_of.htm

"Bowl Championship Series Final 1998-1999 College Football Season BCS Standings." 7 Dec 1998. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/1998_archive_bcs.html

"Bowl Championship Series Final 2000-2001 College Football Season BCS Standings." 3 Dec 2000. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/2000_archive_bcs.html

"Bowl Championship Series Final 2001-2002 College Football Season BCS Standings." 8 Dec 2001. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/2001_archive_bcs.html

"Bowl Championship Series Final 2003-2004 College Football Season BCS Standings." 7 Dec 2003. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/2003_archive_bcs.html

"Bowl Championship Series Final 2004-2005 College Football Season BCS Standings." 4 Dec 2004. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/2004_archive_bcs.html

Colley, Wesley. *Colley's Bias Free Football Ranking Method: The Colley Matrix Explained.* Date unknown. Colley's Bias Free Matrix Ratings. 1 May 2007. http://www.colleyrankings.com/matrate.pdf

Lindsay, John. "BCS Drops Margin of Victory from Rankings Formula." 22 June 2002. ESPN.com. 1 May 2007. http://espn.go.com/ncf/news/2002/0621/1397538.html

Massey, Kenneth. "Massey Ratings Description." 15 August 2000. *Massey Ratings*. 1 May 2007. http://www.masseyratings.com/theory/massey.htm

Matthews, Herman. "About ExpertPicks." 1 May 2007. ExpertPicks.com. 1 May 2007. http://www.expertpicks.com/aboutherman.asp

"NCAA College Football National Champions." 1 May 2007. CollegeFootballPoll.com. 1 May 2007. http://www.collegefootballpoll.com/champions_national.html

Rovell, Darren. "The People Behind the BCS Rankings." 27 Nov 2000. ESPN.com. 1 May 2007. http://espn.go.com/ncf/s/2000/1127/906539.html

Sagarin, Jeff. "Jeff Sagarin NCAA Football Ratings." 24 Jan 2007. *USAToday.com.* 1 May 2007. http://www.usatoday.com/sports/sagarin/fbt06.htm

Wolfe, Peter. "About These Ratings." Date unknown. University of California – Los Angeles. 1 May 2007. http://prwolfe.bol.ucla.edu/cfootball/descrip.htm

Zillgitt, Jeff. "A Look Behind the BCS Computer Ratings." 20 October 1999. *USAToday.com.* 1 May 2007. http://www.usatoday.com/sports/comment/jzcol21.htm