Should Professional Boxing Change Its Scoring System? 
A Comparison of Current and Proposed Methods*

David J. Algranati†
Daniel L. Cork‡

June 7, 2000

CarnegieMellon
H. John Heinz III School of Public Policy & Management
Working Paper #2000–06

Abstract

In the aftermath of the controversial draw verdict in the first bout between Evander Holyfield and Lennox Lewis, numerous suggestions have been advanced to reform the process of scoring professional boxing contests. This paper compares the status quo scoring system in boxing, the 10-Point Must System, to New Jersey’s 10-Point Majority System and the “consensus scoring” technique currently being considered by the New York State Legislature. We compare the three scoring systems on theoretical grounds and by comparing the results when the systems are applied to every world title fight sanctioned by the sport’s four major sanctioning bodies between 1986 and 1999 that was decided on the judges’ scorecards.

*The authors are particularly indebted to Bob Daniel for assistance in obtaining the record books from which our data were collected and for useful comments throughout our analysis. The authors also thank Bob Yalen for clarification of some of the scorecards used in this analysis and the Nevada State Athletic Commission for providing the official scorecards of the Erik Morales-March Antonio Barrera (2000) fight.
†dja@stat.cmu.edu; Department of Statistics and Heinz School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA 15213
‡cork@stat.cmu.edu; Department of Statistics and Heinz School of Public Policy and Management, Carnegie Mellon University, Pittsburgh, PA 15213
Executive Summary

In their May 2000 report, the Boxing Task Force of the National Association of Attorneys General (NAAG) recommended that state boxing commissions evaluate and consider alternative scoring systems. The two alternatives specifically mentioned in this report are New Jersey’s 10-Point Majority System and a consensus scoring technique created by Dr. Ralph S. Levine of Stanford University; this latter system is currently under consideration in the New York State Legislature.

This paper examines and compares the status quo method of scoring boxing matches, the 10-Point Must System, with the Majority System and Levine’s method (hereafter, the “NY System”). We analyze the systems both conceptually and by comparing their results when applied to virtually every world title fight that went to judges’ decision between 1986 and 1999. Our key findings include:

- The current handling of referees’ point deductions under the Majority System is flawed—a one-point deduction can effectively count as a zero- or two-point deduction—but the mistake can easily be corrected. See Section 3.3.1.

- Despite the fact that their procedures are different, the Majority and NY Systems are very similar in practice. In any round where each judge’s round scores are 10-9 or 10-10, they produce the same “consensus” score. Furthermore, after correcting the Majority System’s point deduction problem, the Majority and NY systems identify the same winner in every world title fight that went to decision over the last 14 years. See Sections 3.1 and 4.2.1

- Contrary to claims made in the system’s Utilization Manual, the Majority System does not always award the bout to the fighter with the most total points on the scorecards. In fact, it is possible that a unanimous decision for Boxer A under the Must System—where every judge has more points for Boxer A than for Boxer B—can be ruled a victory for Boxer B by the Majority and NY Systems. See Sections 3.3.2 and 4.2.4.

- The Majority and NY Systems’ practice of crafting a single score for every round of boxing oversimplifies the nature of scoring professional boxing matches. These consensus scoring methods seem predicated on the assumption that judges disagree on round scores only if one judge is incompetent or dishonest, or had their view obstructed. They fail to recognize that many rounds in boxing are closely fought and do not have one, unique, unequivocally correct score. The Majority and NY Systems insistence that a single score be manufactured in close rounds can lead to decisions which disagree with the majority opinion of the judges. See Sections 3.3.3, 3.3.4, and 4.2.4.

- The assertions that the Majority and NY Systems are more immune to the performance of one “bad” judge than the Must system are unfounded. Unlike the Must System, the application of the Majority and NY Systems can lead to situations in which the score of one minority judge can dictate the outcome of the bout. The Majority and NY Systems arrive at a verdict which is at odds with the verdict of the majority of the judges in 6.5% of the world title fights that went to decision between 1986 and 1999. See Sections 3.3.3 and 4.2.4.

We believe that this work constitutes the evaluation and consideration of alternative scoring systems called for by the NAAG Boxing Task Force. If state commissions face a choice between these three systems, we conclude that the 10-Point Must System remains the superior method for scoring professional boxing matches, for a variety of theoretical and empirical reasons.
1 Introduction

Not all professional boxing matches end with a dramatic knockout, or with one fighter deemed unable to continue. Instead, fights are frequently decided on the scorecards of three official judges seated at ringside. The judges score the fight on specific criteria based on what they see in the ring. However, observers in the public and media sometimes arrive at opposite conclusions; they do not necessarily follow the same scoring system mandated for use by official judges, and their opinions may also be influenced by personal passions for or against certain fighters as well as professional reputations of the fighters and their trainers—not to mention thousands or millions of dollars wagered on the bout. Despite the vital importance of the process for judging professional boxing matches to the conduct and integrity of the sport, the topic has received surprisingly little rigorous analysis.

Instead, boxing scoring most often flares up as a heated topic in the immediate aftermath of a bout that is deemed by the public or media to be a “bad” decision. Recent attention to boxing scoring has centered almost exclusively around the first fight between Evander Holyfield and Lennox Lewis in New York City on March 13, 1999 (hereafter referred to as Holyfield-Lewis I). The bout was of particular importance because, for the first time since December 1992, boxing would have an undisputed heavyweight champion of the world. But the excitement and anticipation leading up to the bout quickly turned into anger and frustration. Judge Eugenia Williams scored the bout 115-113 for Holyfield, judge Larry O’Connell scored the bout even at 115-115, and judge Stanley Christodolou scored the bout 116-113 for Lewis. Thus, the highly anticipated fight to crown a new undisputed champion resulted in an indecisive split decision draw—despite popular sentiment that Lewis had comfortably beaten Holyfield.

On the heels of this widely criticized decision, many proposals to reform the boxing scoring system were issued. One such reform—the 20-Point Must System—was adopted by the World Boxing Union (WBU), a minor boxing sanctioning body, only nine days following the Holyfield-Lewis draw. The 20-Point Must System mandates that the boxer winning a round earns twenty points, and the loser nineteen or less; by contrast, boxing’s standard 10-Point Must System awards ten points to the winner, and nine or less to the loser. In theory, the system gives judges more latitude to rate a boxer’s performance in each round. However, the new system’s debut—in an April 1, 1999, bout between Robert Norton and Jacob Mofokeng—showed that the system served only to add 10 points to each fighter’s score for each round. Indeed, the WBU press release (1999) trumpeting the new reform bluntly implied what typical scores would look like under the reformed system when it noted that “WBU judges will change from a ten point must system (10-9) to a 20 point system (20-19)”—literally a simple addition of 10 points for each boxer. Largely an aesthetic change, the 20-Point Must System has not been adopted by any major sanctioning body or state boxing commission.

Another proposed reform that gained currency after Holyfield-Lewis I was “open scoring,” a procedure under which judges’ scores (in some form) would be made known to the audience during the bout. In the wake of Holyfield-Lewis I, open scoring was advocated by powerful boxing promoter Bob Arum and popular welterweight champion Oscar de la Hoya; open scoring was written into proposed legislation at the state level (in a New York boxing reform package) and federal level (H.R. 1240, 106th Congress, introduced by Rep. Traficant of Ohio). The Washington, D.C., Boxing and

---

1 In boxing, an “undisputed” champion is one who is recognized by boxing’s three major sanctioning bodies, the World Boxing Association (WBA), the World Boxing Council (WBC) and the International Boxing Federation (IBF). The heavyweight titles had been fragmented since former champion Riddick Bowe gave up the WBC title in 1992 (Saraceno, 1992).
Wrestling Commission adopted open scoring for a card of three title bouts on April 24, 1999. But—as has happened various other times in boxing history when open scoring is proposed—support for the procedure decayed rapidly in the months following Holyfield-Lewis I. In July, 1999, the Association of Boxing Commissions (ABC) heard comments from former North American Boxing Federation president Mario Latraverse on the performance of open scoring in the Washington, D.C., fight card. As reported in the ABC’s newsletter (1999):

[Latraverse] observed that judges who were not scoring the same as other judges, tried to make those rounds up later in the fight so as to have the same score as the other judges. Contestants appeared to give up after they knew there was no chance to win the fight. The crowd severely booed the odd judge and security problems for the judge developed.

Subsequently, at the same meeting, the ABC unanimously voted to maintain their position of opposing open scoring. Open scoring lost further ground when New York Governor George Pataki vetoed the boxing reform bills (Unattributed, 1999) and was dropped by the Washington, D.C., Boxing and Wrestling Commission; meanwhile, the Traficant-sponsored bill in Congress remains mired in subcommittee. The arguments against open scoring are for the most part in line with Latraverse’s observations: judges’ scoring will be unduly impacted (either by bringing themselves in line with other judges or with the opinion of the audience), the competitiveness of the bouts will be adversely affected (fighters will fight defensively to protect leads, or fighters far behind in points will stop trying), personal safety of judges may be at risk if they disagree with the crowd, and the drama of the announcement of the decision at the end of the fight will be severely diminished.

One proposed scoring reform that remains viable and is still advanced as a potential solution is the use of “consensus” scoring methods. Such methods use the same basic observations from the judges that the 10-Point Must System does, but they combine that information differently in order to arrive at a decision. The three judges’ scores for a particular round are used to create one “consensus” score for that round; if the bout goes to decision, the boxer with the most points on the “consensus” scorecard wins. The most notable and publicized feature of consensus scoring is that, had it been used in Holyfield-Lewis I, Lennox Lewis would have been declared the winner.

Variants of consensus scoring have been adopted by one major boxing state, may be written into law by another, and have been recommended for use by a prestigious national task force. Commissioner Larry Hazzard, Sr., of the New Jersey State Athletic Control Board, has created a consensus scoring method—the 10-Point Majority System—that is now used to decide title fights in New Jersey. The New York State Legislature is currently considering a consensus scoring method for bouts in its state (NY State Assembly Bill A10094; NY State Senate Bill S7801), and Governor Pataki has come out in support of consensus scoring (Office of the Governor of New York, 1999). In May 2000, a Boxing Task Force empaneled by the National Association of Attorneys General and chaired by New York Attorney General Elliot Spitzer, stopped short of fully endorsing consensus scoring but recommended that “state commissions evaluate and consider such alternative [scoring] systems” (Boxing Task Force, 2000).

2The Association of Boxing Commissions (ABC) is a voluntary organization consisting of the boxing or athletic commissions of the various states, as well as the tribal organizations that schedule boxing matches on several Indian reservations. The current chairman of the ABC is Greg Sirb, the executive director of Pennsylvania’s Athletic Commission.
In this paper, we provide a thorough examination of the properties and merits of two major consensus scoring methods relative to the existing 10-Point Must System. We believe that our analysis of the proposals—both conceptual and empirical—will be valuable to state commissions in deciding whether to adopt or experiment with consensus scoring. In Section 2, we describe the scoring systems under consideration in greater detail and outline the claims made in support of each of them. In Section 3, we compare the conceptual properties of the systems; more tellingly, in Section 4, we analyze the systems’ performances on virtually every world title fight that went to judges’ decision between 1986 and 1999. We close in Section 5 with a discussion of findings and recommendations.

2 Boxing Scoring Systems

This paper compares three professional boxing scoring systems, the 10-Point Must System, the 10-Point Majority System, and the consensus scoring technique currently under consideration in New York (to avoid any confusion as to what “consensus” scoring means, this latter system will be called the “NY System” for the purposes of this paper). What these three systems have in common is that they all require the same basic observations from judges. Where they differ is how they take these observations from the judges and determine the outcome of the bout. This section briefly describes the judging process, and then how the three scoring systems operate.

2.1 Scoring a Professional Boxing Match

When any of the three scoring systems examined in this paper are used, the scoring of a boxing match is done by three judges. Typically these judges are seated on three separate edges of the boxing ring. It is also possible—but now very rare—to have two judges at ringside with the third score being provided by the referee, who is in the ring with the boxers.

In determining which boxer wins a particular round, judges are instructed to consider four criteria. One of these criteria is the number of clean punches landed by each boxer, where clean punches are those thrown with a clenched fist that land above the waist on the stomach, chest, and the front or side of the head. Another criterion is defense, the avoidance of punches via blocking, bobbing and weaving (movement of the head and waist), and good footwork (Kaczmarek, 1996). Boxers are also to be evaluated for their effective aggressiveness—whether the fighter forces the action of the bout by pursuing his opponent and initiating action, while landing punches. A final criterion, ring generalship, is described by professional boxing judge Tom Kaczmarek (1996) as “[applying] to the fighter who uses skills beyond his punching power to control the action in the ring;” this includes the use of feints and footwork, as well as the ability to make the opponent fight a style he would prefer not to fight. A judge’s determination of how well a boxer has satisfied these criteria must be based only on what that judge observes and believes; for instance, they do not see the unofficial statistics on punches thrown and landed that are recorded by commercial parties and frequently cited by on-air broadcasters. Kaczmarek (1996) considers effective aggressiveness and clean punches landed as one category, and deems this composite category the most important scoring criterion; ring generalship is the next most important criterion to Kaczmarek, followed by defense.
After considering the relative performances of both fighters based on the scoring criteria, a judge must assign a score for a particular round. Under the three scoring systems considered in this paper, the winner of the round receives ten points, and the loser of the round receives nine points or less. In the event that the judge decides that each boxer is equally deserving—that there is no “loser” of the round—the judge has the discretion of scoring the round “even,” assigning ten points to each fighter.

On the surface, the ability to assign the loser of a round “nine points or less” seems to give individual judges wide latitude in their scoring; in practice, this is not the case. Kaczmarek (1996) suggests that a round won by a “close’ to a ‘definite’ margin” should be scored 10-9. Rounds are almost always scored 10-9, unless a fighter is knocked down during the round, in which case the round would be scored 10-8. Were a fighter to be knocked down by his opponent two or three times, the corresponding score would be 10-7 and 10-6, respectively. Judges vary in their propensity to score rounds even (10-10); additionally, if one fighter thoroughly dominates his opponent but does not score a knockdown, some judges will score the round 10-8.

Another factor which impacts scoring is that the referee is empowered with the ability to penalize fighters for rules infractions; such infractions include blows below the belt, hitting after a round bell rings, and—infamously—biting an opponent. These penalties are usually one point off the offending fighter’s score in the round in which the infraction occurred. Typically, point deductions will only occur for repeated infractions; the referee will usually warn the offending fighter unless the first infraction is particularly egregious.

If a bout ends in knockout or technical knockout, the judges’ scores (and hence the scoring system employed) play no role in the outcome. If a bout is stopped because a fighter is unable to continue due to an accidental foul (e.g., a fighter is badly cut from an accidental headbutt), the scorecards are used to determine the winner of the bout if the bout has lasted longer than a certain number of rounds (four, under the unified rules of the ABC). Otherwise, if the contest does not last longer than this threshold amount of rounds, the bout is ruled a technical draw.

2.2 The 10-Point Must System

The 10-Point Must System is the “status quo” scoring system in boxing. It is, and has been for many years, the predominant method to score professional boxing matches. The 10-Point Must System is currently the scoring system used by every state commission except New Jersey’s. The World Boxing Council (WBC), one of boxing’s four major sanctioning bodies, adopted the 10-Point Must System for all of its title bouts in 1968 (Kaczmarek, 1996).

Under the Ten-Point Must System, the total amount of points for each fighter on each of the judges’ scorecards is tallied. If Judge 1 has more points for Boxer A than for Boxer B, Judge 1’s scorecard is treated as one vote for Boxer A winning the bout. If Judge 1’s scorecard is “even” (the same amount of points for each fighter), this is treated as a vote for a draw. Table 1 enumerates the different type of decisions that are possible under the 10-Point Must System. In this table, “A” means a judge has more points for Boxer A on his scorecard than for Boxer B, “E” means a judge has the same amount of points for Boxer A and Boxer B, and “B” means a judge has more points for Boxer B on his scorecard than for Boxer A.

---

3Kaczmarek (1996) describes 10-6 rounds as “very rare.” Typically, if one fighter in a bout is sufficiently dominant to knock his opponent down three times in a round, the referee stops the bout (regardless of whether an official “Three Knockdown Rule” is in effect).
Example Scorecard

<table>
<thead>
<tr>
<th>Decision Type</th>
<th>Judge 1</th>
<th>Judge 2</th>
<th>Judge 3</th>
<th>Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unanimous Decision</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Majority Decision</td>
<td>A</td>
<td>A</td>
<td>E</td>
<td>A</td>
</tr>
<tr>
<td>Split Decision</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>Split Decision Draw</td>
<td>A</td>
<td>E</td>
<td>B</td>
<td>Draw</td>
</tr>
<tr>
<td>Majority Decision Draw</td>
<td>A</td>
<td>E</td>
<td>E</td>
<td>Draw</td>
</tr>
<tr>
<td>Unanimous Decision Draw</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>Draw</td>
</tr>
</tbody>
</table>

### TABLE 1: Decision Types Under 10-Point Must System

#### 2.3 The 10-Point Majority System

The 10-Point Majority Scoring System was created by New Jersey State Athletic Control Board (NJACB) Commissioner Larry Hazzard, Sr. (1999). It is currently the method used to score professional boxing matches in New Jersey.

This system maintains a single, “master” scorecard. For each round, the score on the master scorecard is arrived at by applying a procedure to the three judges’ scores for that round. At the end of the bout, the boxer with the most points on the master scorecard wins the bout. If each boxer has the same number of points at the end of the bout, the bout is declared a draw.

Following the procedures as described in the NJACB Utilization Manual for the system, the majority scorecard entry for a given round is calculated as follows:

1. Total the points for each fighter, across judges.

2. If one fighter has more points than the other, then:
   
   (a) **Award Points to Winner:** Award 10 points to the fighter with the greater number of points.

   (b) **Award Points to Loser:**
     
     i. If at least two judges have the same exact round score (for both the winning and losing fighter), then award the loser the number of points he received on those common scorecards.

     ii. If there is no common round score among the three judges, then compute the average of the three scores for the loser of the round. The master scorecard records the whole-number portion of the average of their scores.

3. If both fighters have the same amount of points, the master scorecard records the whole-number portion of average of their scores (e.g. if the total points for each fighter in the round was 29, their average is 9.667, and the master scorecard would record the round 9-9).

For example, if the three judges’ round scores were 10-9, 10-9, and 10-8, the master scorecard would record 10-9; 30 is greater than 28 so the first fighter gets 10 points, while the first two judges’ scores of 10-9 give the loser 9 points. If the three judges’ scores were 10-8, 10-10, and 9-10, the master scorecard would record 10-9; 29 is greater than 28 so the first fighter gets 10 points, and the loser gets
the integer part of $\frac{8+10+10}{3}$ (which is 9) because there is no common round score. If the judges’ scores were 10-9, 10-10, and 9-10, the master scorecard would record 9-9.

In the case of point deductions by the referee, the point deductions are taken before the procedure is applied (Hazzard Sr., 2000).

### 2.3.1 Claimed Benefits of 10-Point Majority System

Several claims have been made about virtues of the 10-Point Majority System relative to the 10-Point Must System. In the NJACB Utilization Manual, Commissioner Hazzard argues that the Majority System is a “fairer method of scoring bouts,” that it will “minimize the number of bouts ending in draws,” and that it “award[s] the victory to the boxer with the most accumulated points from all three (3) of the judges” (Hazzard Sr., 1999, original emphasis). In an interview, professional boxing judge Daniel Van de Wiele noted that the New Jersey system “has helped to find the correct winner in several cases” and that it may “reduce the bad decisions” in boxing (Freitag, 2000). In his critique of consensus scoring, Kimball (1999) advances the argument made by Majority System supporters that the system “supposedly . . . lessens the impact one incompetent judge can have on a fight.”

### 2.4 The NY System

The scoring system currently under consideration by the New York State Legislature (referred to here as the “NY System”) is a consensus scoring technique developed by Dr. Ralph Levine of Stanford University. Like the 10-Point Majority System, the NY System constructs a master scorecard; the fighter with the most points on this master scorecard wins the bout if it goes to a decision, while an equal number of points on the scorecard results in a draw.

The NY System differs from the 10-Point Majority System in the procedure it uses to take the three judges’ scores from a round to arrive at the score to record on the master scorecard. The 10-Point Majority System decides upon a winner and loser based on the totals the individual judges’ scorecards, and then calculates winners’ and losers’ points separately. By contrast, the NY System operates on the three judges’ round scores as units. Specifically, the System drops the judge’s score which most favors Boxer A, as well as the score which most favors Boxer B. The remaining score is the one recorded on the master scorecard. In cases where two (or three) judges have the same round score, this procedure would lead to that score being recorded on the master scorecard.

For example, if the three judges’ round scores were 10-9, 10-9, and 10-8, the master scorecard would record 10-9 since two judges are in agreement. If the three judges’ scores were 10-8, 10-10, and 9-10, the master scorecard would record 10-10; the first judge’s score is most favorable to Boxer A and the third judge’s score is most favorable to Boxer B, leaving the second judge’s even score as the consensus score. This differs from the Majority System’s awarding the round 10-9. Finally, if the judges’ scores were 10-9, 10-10, 9-10, the master scorecard would record 10-10; the 10-9 and 9-10 scores would be dropped.

The NY System is invariant with regards to referees’ point deductions; it does not matter whether the deduction is taken before or after the procedure is applied. Because the point deduction is applied on every judge’s scorecard, the resulting score will be the same in both cases.

---

4Van de Wiele stresses that having “three competent and honest judges” scoring the fight is still “the most important thing,” regardless of the system used to combing the judges’ scores to reach a decision (Freitag, 2000).
<table>
<thead>
<tr>
<th>Judge 1</th>
<th>Judge 2</th>
<th>Judge 3</th>
<th>Majority</th>
<th>NY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 2: All Possible 10-9/10-10 Round Scores Under Majority and NY Systems

### 2.4.1 Claimed Benefits of the NY System

The Sponsor’s Memorandum in support of the consensus scoring bill before the New York State Legislature claims that the system corrects for individual judges miscalling a round:

*Under the 10-Point Must System,* if one judge “miscalls” just one round — such as if his or her view is blocked from key action, as one of the Lewis-Holyfield judges admitted was the case — the judge may end up awarding the fight to the lesser fighter in a close match. . . Under the consensus scoring system, one judge miscalling a round will not affect the outcome of a fight.

The memo also claims that the adoption of this scoring system will “greatly enhance the integrity of boxing matches held in New York State,” and that this scoring method “will increase the likelihood that the superior boxer wins.”

### 3 Conceptual Evaluations of Systems

This section will go through the some of the claims made about the Majority System and the NY System and evaluate their merits, conceptually.

#### 3.1 Similarity between Majority System and NY System

It is important to realize that although the Majority System and NY System are two different systems, they often arrive at the same result. The round score 10-9 is the most common round score, by far, in professional boxing. If judges only score rounds 10-9 (reflecting the opinion that one fighter outboxed the other) or 10-10 (even), five inclusive scenarios result; these are shown in Table 2. Under each of these, the Majority and NY Systems produce the same score. The only exception is the round with round scores 10-9, 10-10, and 9-10. But this round only differs in the number of points assigned (9-9 in the Majority System, 10-10 in the NY System); both scoring systems agree that it is an even round. Inasmuch as 10-9 is the most common score recorded by boxing judges, Table 2 describes the most frequent three-judge combinations of scores in boxing.

---

5Of course, Table 2 presents illustrative cases and smooths over the various permutations of judges and winning boxers that result in the same decision type. For instance, the first line of the table could show scores 9-10, 9-10, and 9-10, but the point—that both the Majority and NY Systems record a one-point victory for the second boxer—remains unchanged.
The two scoring systems do not always agree on all rounds. For example, a round with scores 10-8, 10-10, and 9-10 would be scored 10-9 under the Majority System, but 10-10 under the NY System. Still, while the two systems are not completely equal, they are very similar in that they do agree on the scoring of the rounds which, *a priori*, seem to occur most frequently in boxing matches.

### 3.2 Ease of Application

In theory, applying an algorithm like those used to generate majority scorecards should not be difficult. However, there have been occasions in boxing when the simple totaling of scorecards under the 10-Point Must System has been done incorrectly. Examples of this include the first bout between Azumah Nelson and Jesse James Leija for the WBC superfeatherweight title, the Jose Luis Lopez versus Ike Quartey WBA welterweight title bout, and the James Thunder versus Franco Wanyama nontitle heavyweight contest.\(^6\) In all three of these cases, the scorecards were totalled incorrectly and, initially, the incorrect verdict was announced. Though the calculation errors by the ringside officials were, in all likelihood, honest mistakes, they do increase the perception of impropriety in professional boxing.

Accordingly, it is legitimate to worry that the complexity of consensus scoring methods relative to simple totaling of judges’ cards—particularly the more complex 10-Point Majority System—may make such errors in calculation more frequent (Kimball, 1999). At the end of a fight that goes to the scorecards, the ringside clerk is under pressure to generate an accurate score quickly while a large live audience and larger television audience waits; if simple arithmetic errors have caused problems in the past, errors would likely seep into a complicated point assignment algorithm that takes simple summation as a first step.

### 3.3 Claims about the Alternative Systems

Some of the specific claims listed in Section 2 that have been made in support of the Majority and NY Systems can be addressed conceptually or theoretically. Such conceptual analysis relies on hypothetical cases and counterexamples which are very interesting and illustrative; we will turn to an empirical analysis of how often such counterexamples occur in practice later in the paper.

#### 3.3.1 The Majority System is “Fairer”

There are many ways in which to assess the “fairness” of a system. It seems that this statement is predicated upon the assertion that the system will give the bout to the boxer with the most total points; that claim will be discussed in Section 3.3.2.

One respect in which the Majority System may be decidedly unfair is the way in which it handles point deductions. Intuitively, if the referee deducts one point from Boxer A, then Boxer A should have one point less than he would have earned had there been no deduction. Under the Majority System, this is not always the case. For example, consider a round were two judges score the round 10-9 for Boxer A, and one judge scores the round 10-9 for Boxer B. The Majority System would score this round 10-9 for Boxer A. Now, if there is a point deduction to A during the round, the current rules of

---

\(^6\)Nelson was initially announced as the winner in his September 1993 bout with Leija, but correct calculation led the fight to be called a draw; similarly, Quartey was initially announced as the winner but the outcome was actually a draw. The 1995 Thunder-Wanyama miscalculation actually reversed the announced decision; Thunder was declared the winner, but the correctly-added scorecards favored Wanyama.
the Majority System hold that point deductions are taken before the procedure is applied. Thus, after the deduction, two judges would have scores of 9-9, and one would have a score of 10-8, for Boxer B. Boxer B then has 28 total points as compared to 26 for A; B gets 10 points for the round. Since two judges have the round scored 9-9, Boxer A receives 9 points for the round. The master scorecard records the round as 10-9 for Boxer B. Since the round would have been 10-9 for Boxer A had there been no deduction, this is effectively a two point difference (Boxer B gained a point while Boxer A lost a point).

The above example is a case when a single point deduction has the effect of deducting two points. We can also construct an example where a single point deduction is effectively no deduction at all. Consider a round where all three judges score the round 10-8 for Boxer A; the score recorded on the master scorecard is 10-8. Now, consider the same scores with a point deduction on Boxer A. Since the point deduction is taken before the procedure is applied, the score recorded on the master scorecard is still 10-8 (Boxer A has more total points [27 versus 24], so he is the winner of the round and receives 10 points, while Boxer B receives 8 points). This situation poses clear problems. Not only is it normatively bad (if there is a point deduction, it should be reflected in the score), but if a boxer had reason to believe he is likely to win a round 10-8, he may be encouraged to be reckless or to foul, since there is no effect of a point deduction. While fighters never know what score they will receive from the three judges, it is fairly common to score a round in which a knockdown is scored as 10-8. So, if Boxer A scores a knockdown early in the round, he can be reasonably confident that as long as he maintains his advantage, he will earn 10-8 scores from all three judges. Consequently, he may be tempted to foul, thinking that even if the referee decides to deduct a point, it will have no ramification on his score in the bout.

While it seems clear that the way point deductions are currently handled under the Majority System is unfair, this problem can be easily remedied. Simply taking deductions after the procedure has been applied will eliminate the problems the Majority System has with point deductions.

### 3.3.2 Majority System Gives Victory to Boxer with Most Points

In the NJACB 10-Point Majority System Utilization Manual, Commissioner Hazzard argues that “sports fans are accustomed to acknowledging the winner of an athletic contest as the individual or team accumulating the most points at the end of a contest. This is the most basic and fair concept of selecting winners in athletic competition” (1999). Hazzard further raises an example where one judge scores the bout 100-90 for Boxer A, and two judges go on to score the bout 96-94 for Boxer B. Hence, summing across the scorecards, Boxer A has 288 points and Boxer B has 282; since “Boxer A has more accumulated points, than Boxer B,” says Hazzard (1999), Boxer A “should be determined the winner of the bout” (1999). As Hazzard correctly points out, the 10-Point Must System rules this fight a split decision victory for Boxer B. This notion that the Majority System gives the victory to the fighter with the most total points is a rationale for the claimed fairness of the system relative to the 10-Point Must System.

The argument is deficient on at least two conceptual fronts, though. First, the argument glosses over the distinction between objective and subjective scoring in sports. Scoring in sports like basketball, baseball, football, and hockey is objective. At the margins and in extreme cases, there may be subjective elements to scoring in these sports, such as rulings pertaining to whether or not a player is out of bounds. But, ultimately, baskets, runs, touchdowns, or goals are all objective distinctions, and virtually all observers know when one has occurred and when one has not. If, at the end of four quarters of
basketball, the Chicago Bulls had 100 points and the Utah Jazz had 94 points, but the National Basketball Association declared the Jazz the winner of the contest, there would likely be outrage in the entire sporting community. In basketball, the team with more points always wins.

Such is not the case in boxing—a subjective scoring sport where points reflect the opinion of the judges rather than how many times an objective goal was accomplished. Awarding the contest to the individual or team with the most points is neither the most basic nor the most fair “concept of selecting winners in athletic competition.” In sports where scoring is subjective, the norm is not to merely add up all the judges’ scores to determine the winner. For example, in diving, the highest and lowest scores and then the total of the remaining scores is executed (FINA, 2000); similar practices hold in other sports like gymnastics.

The second problem with the argument that the Majority System gives the victory gives the victory to the boxer with the highest point total is that it does not always do so. Even the manual’s hypothetical example—with marginal totals 100-90 for Boxer A and 96-94 (twice) for Boxer B—can be constructed so that Boxer A does not win. Table 3 illustrates that Hazzard’s claim that the Majority System gives the victory to the fighter with the most points is not always true. This may not be bothersome to most people, since giving the win to the individual with the most total points is not a common goal in subjectively scored competitions.

What may bother people, though, is what Table 4 reveals. In the hypothetical bout described by Table 4, not only does Boxer A have the most total points, he has the most total points on each of the three judges cards. That is to say, under the Must System, this is a unanimous decision for Boxer A. However, under the Majority System (and the NY System), this bout is a victory for Boxer B.

Hence, the Majority System does not always give the victory to the fighter with the most points. Certainly, if one wanted to always identify the winner of a fight as the fighter who has attained the most points, there is a very simple means of doing so: merely add the points up at the end of the bout, and award victory to the fighter with the most points. This would always achieve the goal of giving the bout to the fighter with the most points, and is much more straightforward and less convoluted than the Majority System. However, such a summation strategy is likely not a good method of determining the winner in light of the fact that boxing is a subjectively scored sport.

<table>
<thead>
<tr>
<th>Round</th>
<th>Judge 1 A</th>
<th>Judge 1 B</th>
<th>Judge 2 A</th>
<th>Judge 2 B</th>
<th>Judge 3 A</th>
<th>Judge 3 B</th>
<th>Majority A</th>
<th>Majority B</th>
<th>NY A</th>
<th>NY B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>90</td>
<td>94</td>
<td>96</td>
<td>96</td>
<td>94</td>
<td>96</td>
<td>96</td>
<td>94</td>
<td>96</td>
</tr>
</tbody>
</table>

Table 3: Counterexample to Hazzard’s Claim
Consensus Scoring Reduces Influence of “Bad” Judges

Claims that consensus scoring techniques like the Majority System decrease the influence of “bad” judges seems to be predicated on two points: first, that two “good” judges calling the round correctly will offset one “bad” judge miscalling the round on the master scorecard and, second, that the specific example off how such systems resolve Holyfield-Lewis I clearly demonstrates how the system attenuates the influence of a “bad” judge.

The practice of discarding a minority (“bad”) judge’s score in a round is discussed further in Section 3.3.4. This section will focus on Holyfield-Lewis I, the example most commonly cited in recent literature when discussing the Majority System or the NY System. The scores for Holyfield-Lewis I are given in Table 5.

Table 5 illustrates that under the Majority System and the NY System, Lennox Lewis would have won the first Holyfield-Lewis bout, with a score of 115-113. Under the Must System, O’Connell’s even scorecard counts as one vote for a draw, Christodolou’s scorecard counts as one vote for Lewis, and Williams’ scorecard counts as one vote Holyfield. The result, under the Must System, is a split decision draw. Judge Williams’ scorecard has come under particular scrutiny since her scorecard favored Holyfield (whereas popular perception was that Lewis won), and for her controversial decision to score the fifth round for Holyfield. That specific round found Holyfield in a defensive posture during most of the round while withstanding attack from Lewis, who was comparatively much more active. The Majority and NY Systems score round five as 10-9, Lewis, as an artifact of the 10-9 scores from O’Connell and Christodolou.

Due to her scoring of the first Holyfield-Lewis bout, Eugenia Williams is now perceived as a “bad” judge; suspicions that the fight was fixed have centered around Williams’ financial situation and her affiliation with a sanctioning body whose title Holyfield defended in the March 1999 fight. The basic reason why this example is used to justify and bolster the Majority and NY Systems is that it would seem that these systems attenuate the influence of the “bad” judge, Williams.
TABLE 5: Scores for Holyfield-Lewis I. Holyfield's scores are in the columns marked “H”, Lewis's in the columns marked “L”.

<table>
<thead>
<tr>
<th>Round</th>
<th>O'Connell</th>
<th>Christodolou</th>
<th>Williams</th>
<th>Majority</th>
<th>NY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H L</td>
<td>H L</td>
<td>H L</td>
<td>H L</td>
<td>H L</td>
</tr>
<tr>
<td>1</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>2</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>3</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
</tr>
<tr>
<td>4</td>
<td>9 10</td>
<td>9 10</td>
<td>10 9</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>5</td>
<td>9 10</td>
<td>9 10</td>
<td>10 9</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>6</td>
<td>10 9</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>7</td>
<td>10 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>8</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
</tr>
<tr>
<td>9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
</tr>
<tr>
<td>10</td>
<td>10 10</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
</tr>
<tr>
<td>11</td>
<td>10 9</td>
<td>10 10</td>
<td>10 9</td>
<td>10 9</td>
<td>10 9</td>
</tr>
<tr>
<td>12</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
<td>9 10</td>
</tr>
<tr>
<td>Total</td>
<td>115 115</td>
<td>113 116</td>
<td>115 113</td>
<td>113 115</td>
<td>113 115</td>
</tr>
</tbody>
</table>

Actually, though, the influence of a single judge on the outcome of the bout under the Majority and NY Systems is not entirely clear. Suppose that O’Connell’s and Christodolou’s scorecards remain unchanged but that Williams had blatantly favored Holyfield by awarding every round 10-9 to Holyfield. Her score would then be 120-108 Holyfield rather than 115-113 Holyfield. Under the Majority and NY Systems, Holyfield would then win rounds 3, 6, 7, 8, 9, 10, and 11, while Lewis would win rounds 1, 2, 4, 5, and 12. Round 7 would be scored even. The result would be 114-113 for Holyfield under the Majority System, and 115-114 for Holyfield under the NY System. In this case, the “bad” score of 120-108 for Holyfield forces the Majority and NY Systems to favor Holyfield. In contrast, under the status quo Must System, the verdict would be the same: a split decision draw.

Williams’ score need not be this extreme or blatant to have the Majority and NY Systems score the bout for Holyfield. Again, keep O’Connell’s and Christodolou’s scorecards the same. Now, if Williams simply scored rounds six and seven for Holyfield instead of rounds four and five, her score would be 115-113, just as it was in the actual fight. But, with that simple change, the Majority System would score the bout 114-113 Holyfield and the NY System would score the bout 115-114 Holyfield. Again, though, the Must System would rule the fight a split decision draw.

Under the 10-Point Must System, each judge has one vote. The magnitude by which they deem a boxer the winner—whether 115-113, 120-108, or 120-84—does not matter; only the direction of their decision is consequential. In essence, the Must System already has a method to deal with “bad” judges: they only get one of three votes, while a boxer needs two votes to win a bout. Under the Must System, if two judges score a bout for Boxer A, the third judge cannot deny the victory from Boxer A, no matter how lopsided their score for Boxer B. This is not the case for the Majority or NY Systems. If Judge 1 and Judge 2 both score the bout for Boxer A, but disagree on some of the round scores, it

---

7 The Majority and NY Systems have different totals due to the fact that the Majority System would score round 10 as 9-9 and the NY System would score round 10 as 10-10. See Table 2.
is possible that Judge 3 can “sway” the Majority or NY System master card to be scored for Boxer B, if Judge 3 scores the rounds on which Judges 1 and 2 disagree in favor of Boxer B. Only in the event that Judges 1 and 2 score each round exactly the same does a “rogue” Judge 3’s score not matter at all; under the Majority and NY Systems, Judges 1 and 2 will define the score for the master scorecard by virtue of their agreement.

3.3.4 NY System Increases Likelihood that the Superior Boxer Wins

Table 4 illustrates a case where Judges 1, 2, and 3 all score the bout for Boxer A (they each individually identify Boxer A as the superior boxer), but the NY System awards the bout to Boxer B. This occurs because of the lack of agreement among the judges in rounds one through seven. It is reasonable to question why there would be such disagreement. The sponsor’s memorandum filed in support of the New York Assembly Bill that would mandate the use of consensus scoring speaks about judges “miscalling rounds:

Under the consensus [NY] scoring system, one judge miscalling a round will not affect the outcome of a fight. For example, if the miscall of the round is caused by chance, it is unlikely that another judge will make the same mistake in the same round. If the miscall resulted from a judge’s view of some key action being blocked, consensus scoring should correct the call for the round because the other two judges, whose views were presumably unobstructed, would overrule the judge who could not see. When the consensus [NY] scoring method is used, the miscalled rounds do not matter, except in the unlikely event that all of the judges miscall the same round.

So, it seems reasonable to assume that the explanation of the scoring in Table 4 from those who believe the above quote would be that in rounds one through seven, one of the judges miscalled the round. The underlying assumption in the above quote is that the only reason there would be disagreement among boxing judges in a round is if at least one of them miscalled the round.

In boxing, there are many rounds in which there is a clear winner—where Boxer A is the aggressor, exhibits better defense, throws and lands more punches as well as the stronger, more effective blows. However, not all rounds are quite so clear. Boxing matches ideally pit fighters who each have some modicum of skill, stamina, and training; hence, rounds can be fought vigorously by both fighters with a large amount of “give-and-take” between them. Perhaps one fighter lands more total punches, but the other fighter clearly lands much more effective punches. Different judges may weigh the four scoring criteria in slightly different ways; they may also vary in their willingness (and perceived ability) to score even rounds. In short, there are some rounds in professional boxing on which three honest, fully-trained, attentive, and competent boxing judges can still legitimately disagree as to a winner and loser. In such close rounds, there is not one, unique, unequivocally correct score for the the round, yet that is what consensus scoring methods like the Majority System and the NY System seek to manufacture for each round.

If the only reasons why judges disagree on scoring a particular round are that one judge lacked concentration, was dishonest, or had their view obstructed, then combining three scores into one (in many cases, “discarding” the odd score) might be appropriate. However, this is not always the case. In rounds with very close, competitive action, lack of complete agreement among the three judges likely reflect the closeness of the round, not that one judge “miscalled” the round due to his view being obstructed or other illegitimate reasons. Collins (1999) defends the judging of the 1999 fight
between champion Oscar de la Hoya and challenger Ike Quartey, saying that the decision “should not be considered a robbery, or even a bad call. The truth of the matter is that it was an extremely difficult fight to score, a fight that challenged the three ringside judges almost as much as it did the combatants.” Collins further notes that the three ringside judges returned common scores on only three rounds; “other than that, who won which round was far from obvious.”

It is much more likely that disagreements among judges in these bouts reflect the closeness and competitiveness of the rounds between two boxers with different styles, rather than obstructed views or incompetence.

Marc Ratner, the executive director of the Nevada State Athletic Commission, has stated that “if the judges are trained right and philosophically right, to take the minority score away is not the best way to do business. I think that is wrong” (Feour, 2000b). We agree with Ratner, in that we believe that coercing the three judges’ scores into one oversimplifies the scoring process in close rounds. The premise that there is always an unequivocally correct score for each round of boxing is at odds with the subjective nature of the sport. Judges are obliged to consider all the feints, foot movement, head movement, aggression, and punches exhibited by both fighters in three minutes of boxing and then, instantaneously, quantify their impressions of that round to reflect the boxers’ performances; disagreement among the judges on particular rounds is not necessarily a harbinger of corruption or other problems but may speak to the closeness and competitiveness (and, arguably, the quality) of a boxing match.

4 Performance of Scoring Systems on Actual Bouts

While the conceptual counterexamples and findings in the previous section are interesting, it could be argued that the most important characteristic of these systems is how they perform on actual fights. Seeing the methods in action can shed more light on some of the claims made about the scoring systems. For example, a claim such as the “Majority System minimizes draws” is difficult to evaluate conceptually. It is clear that draws are possible in both systems—the construction of a consensus, master scorecard does not preclude both boxers having the same number of points on said card—but it is difficult to have an inkling as to which system would result in fewer draws. Another reason to see the scoring systems in action is that conceptual examples rely on constructing counterexamples like Table 4—a reversal of a unanimous decision—but if such cases never happen in practice, then they may not be considered severe problems with consensus scoring methods. Then, it could be argued that the claimed benefits of those systems outweigh the very rare theoretical mistakes allowed under the systems.

One way to evaluate the performance of these systems on fights is to implement the systems (as has been done in New Jersey), and, after a period of time (several years), evaluate the results of its performance. An alternative approach is to look back at fights which have already happened, and evaluate how the systems would have performed in past fights. This approach makes the assumption that the assignment of judges to the fight would have been the same, had a different scoring system been in use. It also makes the assumption that the judges would have scored the rounds the same.

Of the nine rounds where agreement was not unanimous, all three judges returned different scores in one round and one judge dissented from his peers in each of the remaining eight rounds. If one were to represent the de la Hoya-Quartey fight in the manner used by consensus scoring advocates—where the idea is to correct for the influence of one “rogue” judge—one would have to account for the fact that each of the three ringside judges in this fight was a “rogue” dissenter in at least one of those eight rounds.
way, regardless of the scoring system in use. These assumptions are exactly the same as those implicitly
made by observers who have applied the Majority and NY Systems to Holyfield-Lewis I in order to
gauge the quality of their system; Holyfield-Lewis I is the only real fight that the Utilization Manual
for the 10-Point Majority System or the sponsor's memorandum for the bill in the New York State
Legislature uses to illustrate and bolster their case. We are comfortable making these assumptions, and
have chosen a retrospective analysis of past fights as a means to evaluate the scoring systems.

Our approach has limited precedent in boxing literature. Trunzo (1999) went beyond relying
solely on Holyfield-Lewis I in a *Ring Magazine* article that applied a consensus scoring method to sev-
eral handpicked “bad” fights: these included Holyfield-Lewis I, Pernell Whitaker-Julio Cesar Chavez,
Oscar de la Hoya-Ike Quartey, Oscar de la Hoya-Pernell Whitaker, Lennox Lewis-Ray Mercer, George
Foreman-Shannon Briggs, and James Toney-Dave Tiberi. Trunzo finds that, in addition to awarding
Holyfield-Lewis I to Lewis, the consensus scoring system he uses would award the Whitaker-Chavez
bout to Pernell Whitaker. Under the Must System, the bout was a highly controversial majority de-
cision draw; like Holyfield-Lewis I, many people felt that one fighter (Whitaker) had won the bout.
On all other bouts Trunzo examined, the verdict was the same under both the Must System and the
“consensus” scoring system he applied.

4.1 The Data

In order to compare the three scoring systems, we have chosen to apply each system to all the World
Boxing Association (WBA), World Boxing Council (WBC), and International Boxing Federation
(IBF) world title fights that went to decision from 1986-1999 and all the WBO world title fights
that went to decision from 1990-1999. Our data, which includes vacant and “interim” world title
bouts, has a total of 956 title fights which went to decision. A world title fight is one in which one
or more of the sanctioning bodies’ championship belts for a particular weight class are at stake in the
fight. We choose to work with world title fights for two reasons, the first of which is that they are the
most prestigious fights in boxing; they are at once the fights that we might reasonably expect to have
the best-trained judges at ringside and the fights which are potentially most vulnerable to the “rogue”
judging that consensus scoring seeks to ameliorate. The second reason is more practical; the scorecards
for world title fights are readily available in published form, without having to compile scorecards from
various state commissions. Specifically, our data are drawn from the 1986–1999 record books pub-
lished annually by Fight Fax Inc. (1987–2000, annual), which include the scorecards for every world
title fight for a particular year.

---

9Trunzo (1999) describes the system he applies as “If at least two of the three judges see the round in favor of one
fighter, that fighter wins the round by the appropriate margin.” Trunzo never addresses what to do in situations where all
three judges disagree. However, when at least two judges agree, the system he applies is similar to the NY system.

10Although the WBO existed prior to 1990, their world title fights were not recorded in our data source until that year.

11We were unable to obtain round-by-round scorecards for six fights: Zapata-Dias, WBA Flyweight, 12/7/86; Kalambay-
Barkley, WBA Middleweight, 10/23/87; Roman-Konadu, WBC Super Flyweight, 11/7/89; Ruiz-Velasco, WBO Junior
Bantamweight, Dominguez-Kobozev, WBC Interim Cruiserweight, 10/24/95; Wright-Bingham, WBO Jr. Middleweight,
11/9/96. Additionally, we did not use the scorecard from the first fight between Ricardo Lopez and Rosendo Alvarez,
3/7/98, due to the questionable nature of those cards (no point deduction was taken on one of the judges cards, despite
the rules requiring such (Sulaiman, 1998)).
Table 6: Comparison of basic outcome of 956 world title fights under the Majority System and the NY System.

<table>
<thead>
<tr>
<th></th>
<th>NY System Decision</th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Champion Wins</td>
<td>Challenger Wins</td>
<td>Draw</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majority System</td>
<td></td>
<td>645</td>
<td>0</td>
<td>3</td>
<td>648</td>
</tr>
<tr>
<td>Decision</td>
<td></td>
<td>0</td>
<td>264</td>
<td>1</td>
<td>265</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>39</td>
<td>43</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>648</td>
<td>265</td>
<td>43</td>
<td>956</td>
</tr>
</tbody>
</table>

Besides the larger number of fights, we believe there is another crucial reason why our dataset is a better set of fights on which to evaluate the scoring systems than a more restricted set of “bad decisions,” such as that analyzed by Trunzo (1999). The latter approach induces a “selection bias” on the results. That is to say, if we only look at fights which are felt to have been ruled on inappropriately by the Must System, then applying an alternative system cannot lead to any worse results. What is missing by selecting fights as Trunzo did is that there is no way to see how alternative systems would perform on fights which are believed to have been scored correctly by the 10-Point Must System. Such an analysis would miss cases like Table 4, wherein the Majority and NY Systems would reverse a unanimous decision—seemingly turning a “good” decision on which all judges agree that Boxer A won the bout into a “bad” decision by awarding the contest to Boxer B.

In the analyses that follow, we refer to one boxer as the “champion” and one as the “challenger.” In standard title fights, where an incumbent champion defends his title against a rival, the labels are assigned in the obvious manner. But there is not always a designated champion. In the case of vacant titles (where neither fighter has a title going into the bout), we label the higher-ranked boxer as the “champion.” In the case of unification bouts (where both fighters have titles), one fighter was arbitrarily identified as the “champion” and the other as the “challenger.”

4.2 Results

The three scoring systems were applied to all 956 WBA, WBC, IBF and WBO world title fights in the data set and the results are discussed in the following sections.

4.2.1 Comparison of Majority and NY Systems

As discussed in Section 3.1 and illustrated in Table 2, there is a large degree of similarity between the Majority and NY scoring systems. While they do not always yield the same score for every round, one way to evaluate how different they are is to see how often they identify a different winner. Table 6 summarizes the results of the comparison between the two systems. The Majority System verdict is on the rows, and the NY System verdict is on the columns of the crosstabulation.

Table 6 reveals that the Majority and NY Systems differ on only eight out of 956 fights (0.8%). Recall that Section 3.3.1 details how, occasionally, the Majority System’s practice of deducting points before applying the procedure can lead to unfair outcomes, such as a one-point deduction counting

---

12The sanctioning bodies—the WBA, WBC, IBF, and WBO—maintain their own separate sets of boxer rankings, by weight class, which are used to prioritize boxers in line for a championship.
4.2.2 Comparison of Must System to Majority/NY Systems

In Table 8, we compare the outcomes under the Must System to outcomes under the NY System (and, equivalently for these fights, the corrected Majority System). The outcomes are tabulated by Must System decision type (unanimous, majority, or split) and by winner (champion, challenger, or draw).

The first striking feature of Table 8 is that the Must System and Majority/NY System very frequently identify the same winner; the two systems agree on 91.0% (870 out of 956) of the fights in the dataset. This high agreement speaks to the fact that there is a clear winner in many boxing matches, and that the particular manner of combining the judges’ scorecards to declare a winner is inconsequential. The performance of the competing systems on the remaining 9% of fights—still a non-trivial quantity—is the crux of the debate.

4.2.3 Draws Under the Majority/NY Systems

Table 8 also illustrates that there are fewer draws under the Majority System (43) than under the Must System (47), but not by a wide margin. This provides weak, partial evidence for the claim that consensus scoring minimizes draws. Inasmuch as there are 14 years of data used in this study, the 0.4% reduction in draws is roughly equivalent to one fewer draw under the Majority and NY systems every 3.5 years of title fights.
<table>
<thead>
<tr>
<th>Must System Outcome</th>
<th>Champion Wins</th>
<th>Challenger Wins</th>
<th>Draw</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Champion Wins</td>
<td>624</td>
<td>12</td>
<td>17</td>
<td>653</td>
</tr>
<tr>
<td><strong>Unanimous</strong></td>
<td>504</td>
<td>1</td>
<td>5</td>
<td>510</td>
</tr>
<tr>
<td><strong>Majority</strong></td>
<td>43</td>
<td>2</td>
<td>4</td>
<td>49</td>
</tr>
<tr>
<td><strong>Split</strong></td>
<td>77</td>
<td>9</td>
<td>8</td>
<td>94</td>
</tr>
<tr>
<td>Must System</td>
<td>Challenger Wins</td>
<td>6</td>
<td>235</td>
<td>15</td>
</tr>
<tr>
<td><strong>Unanimous</strong></td>
<td>2</td>
<td>157</td>
<td>1</td>
<td>160</td>
</tr>
<tr>
<td><strong>Majority</strong></td>
<td>0</td>
<td>22</td>
<td>4</td>
<td>26</td>
</tr>
<tr>
<td><strong>Split</strong></td>
<td>4</td>
<td>56</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Decision</td>
<td>Draw</td>
<td>18</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td><strong>Unanimous</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Majority</strong></td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td><strong>Split</strong></td>
<td>10</td>
<td>14</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>648</td>
<td>265</td>
<td>43</td>
<td>956</td>
</tr>
</tbody>
</table>

Table 8: Comparison of Must System by Decision Type with Majority/NY Systems

But Table 8 also reveals an interesting phenomenon about the difference in draws. Specifically, the two systems only agree on which fight is a “draw” 11 times. The Majority/NY Systems convert 36 draws under the Must System into fights with identified winners, but they also take 32 fights in which the Must system identifies a winner and convert them into draws. Six unanimous decisions are made into draws, as are eight majority decisions.

4.2.4 Reversal of Majority Opinions

The claims that the Majority system is “fairer” or reduces the amount of bad decision and that the NY system “increases the likelihood that the superior boxer wins” is brought into question by this analysis. There are nine unanimous decisions in which the boxer each judge identifies as the winner is not granted the victory under the Majority/NY Systems; in three of these cases, the hypothetical counterexample of Table 4 becomes reality as the unanimous victory is reversed. There are two majority decisions where no judge’s total is for a “Boxer B” but, Boxer B is awarded the decision under the Majority/NY systems. There are 62 fights (6.5%) where at least two judges agree on the outcome of the bout, but the Majority/NY systems results in a verdict different than the majority of the judges. If one believes that agreement between two or three of the official ringside judges is evidence that these judges are correctly perceiving that the superior boxer is winning the fight while, perhaps, one “rogue” judge submits erroneous scores, then it is somewhat disturbing that the Majority/NY Systems reject that agreement in 6.5% of these title fights.
4.2.5 Awarding Victory to the Fighter with the Most Points

One claim made in support of the Majority System is that it awards victory to the fighter with the most points. Tables 3 and 4 reveal that the claim is not strictly true, conceptually. But our dataset allows us to test this claim empirically. We define a “summation" scoring system where a fighter wins if he has the most total points on the three judges' scorecards, and the fight is a draw if the total points are equal. We find that the Must System agrees with the summation system on 897 of 956 bouts (93.8%), while the Majority/NY Systems agree with the summation system on 884 of 956 of the fights (92.5%). So, not only does the Majority System fail in its objective of awarding the fight to the fighter with the most points, it actually accomplishes this objective less often then the Must system. As we emphasized in Section 3.3.2, we do not feel that the summation strategy of giving the fight to the fighter with the most points is a good scoring method in subjectively scored sports; however, it is interesting to see that the Majority System falls short of one of its professed goals.

4.2.6 “Bad Decisions,” Old and New

Our primary contribution to the literature in this paper is Table 8, the results of applying the competing systems to every one of 956 world title fights that went to decision. But, inevitably, the argument over consensus scoring in boxing turns to specific examples of “bad decisions,” which might or might not be resolved if a new scoring system were implemented. Our dataset gives us a rich source from which to select questionable decisions and see what difference consensus scoring might have made.

The Majority and NY Systems definitely have success stories in terms of resolving “bad decisions;” it is on the strength of the argument that they would have bestowed victory on Lennox Lewis in Holyfield-Lewis I that has fueled much recent discussion of the methods. But other fights stand out as well. Veteran USA Today boxing writer Jon Saraceno (1999) lists the following bouts as “the three worst draws I’ve seen first hand:” Ray Leonard-Tommy Hearns II, Whitaker-Chavez, and Holyfield-Lewis I. All three of Saraceno’s “worst draws” are resolved by the Majority/NY Systems in a fashion which the public would have approved; they score the fights for Hearns (113-112 on both systems), Whitaker (115-112 on the Majority system, 116-113 on the NY system) and Lewis (115-113 on both systems), respectively. These three high-profile “bad decisions” are cases where either the Majority or NY Systems could have fixed—changed from an unsatisfying draw to a decision that would be “correct” in the idea of popular and media observers.

It is important to realize, though, that the Majority/NY systems do not resolve all controversial draws in a fashion consistent with the majority of public opinion. Consider the first Azumah Nelson-Jeff Fenech bout, held in 1991, which ended in a split decision draw with judges’ scores of 115-113 Fenech, 116-112 Nelson, and 114-114. Los Angeles Times writers Earl Gustkey and Allan Malamud both disagreed with the decision, as Gustkey had the bout 118-110 Fenech, and Malamud had Fenech “by four points” (Gustkey, 1991; Malamud, 1991). The president of the Australian National Boxing Federation, Peter Burchall, protested the outcome of the first Nelson-Fenech fight, saying, “we feel an unfair decision was rendered” (Unattributed, 1991a). In response to the decision and the ensuing controversy, WBC President Jose Sulaiman ordered a rematch between Nelson and Fenech (Unattributed, 1991b). However, had the Majority or NY System been in use for this fight, there would have likely been even more uproar following the contest: both systems would have scored the bout 115-113 for Nelson.
Another example of a draw which is not resolved in a way consistent with public opinion is the 1997 Lonnie Bradley-Otis Grant match. The official scores were 115-113 Grant, 115-113 Bradley, and 114-114. The Majority and NY systems score this bout 115-113 Bradley. Yet, at the time the draw verdict was announced following the fight, Bradley seemed to be both surprised and happy to be given a draw. Indeed, the public sentiment in this bout was that Grant deserved the victory (Todd, 1997; Stubbs, 1997).

Many high-profile fights in our dataset are bouts which, in the eyes of some, were “bad decisions:” these include Michael Spinks-Larry Holmes II, Marvin Hagler-Ray Leonard, Jose Luis Ramirez-Pernell Whitaker I, James Toney-Dave Tiberi, Evander Holyfield-Michael Moorer I, Julio Cesar Chavez-Frankie Randall II, George Foreman-Axel Schulz, Francois Botha-Axel Schulz, Pernell Whitaker-Oscar de la Hoya, and Oscar de la Hoya-Felix Trinidad. Anyone who was disgruntled with any of these listed decisions would remain disgruntled, as all of those fights have the same verdict under the Majority/NY Systems. Since our data ends with December 1999 fights, boxing’s most recent “bad decision,” Erik Morales-Marcus Antonio Barrera (February 19, 2000), was not one of our 956 fights examined. We obtained the scorecards from that fight separately, from the Nevada State Athletic Commission. This “bad decision,” too, would be unresolved by consensus scoring as the Majority/NY Systems pick Erik Morales—the controversial winner under the Must system—as the winner (and still champion).

There are several fights in our dataset which were not viewed as “bad decisions” by the public, but likely would have been if the Majority or NY Systems had been used. As Table 8 reveals, there are nine unanimous decisions out of these 956 fights in which the Majority/NY systems do not award the fight to the winner on all three judges’ cards. The highest-profile fight among these nine is the second bout between Pernell Whitaker and Wilfredo Rivera. Whitaker is well known in boxing for his talent and accomplishments (1984 Olympic Gold Medalist, four division titlist in professional boxing), but is equally well known for being on the short end of (at least) two widely-panned “bad decisions:” his first bout with Jose Luis Ramirez and his bout with Julio Cesar Chavez. On the second bout between Whitaker and Rivera, boxing journalist Michael Katz (1996) expressed gratitude for Whitaker’s victory:

> With a 12-round statement that he is still The Man, Whitaker “avenged” his somewhat controversial split decision over Wilfredo Rivera . . . Whitaker (39-1-1) took some incoming but gave the game Puerto Rican a pounding to defend his welterweight title despite the kind of WBC judging that produced the two blemishes on his record. They gave him a 1988 loss against Jose Luis Ramirez and a 1993 draw with Julio Cesar Chavez, and [this] was close to being just as outrageous… The Daily News gave [Rivera] only the second round calling the fifth round even. Incredibly, if [the sixth round] had been even, Rivera would have won. If Whitaker had won by only 10-9, the bout would have ended in a draw and boxing would have had more sores than Rivera’s battered face.

Boston Globe writer Ron Borges (1996) was also impressed with Whitaker’s performance in the rematch with Rivera, saying that “although in the end the scorecards made Whitaker only a close victor, few sitting ringside concurred. . . . Regardless of the questionable judge’s scores . . . [Whitaker] controlled the bout and was never in trouble.”
<table>
<thead>
<tr>
<th>Round</th>
<th>J. Kassees</th>
<th>J. Keane</th>
<th>M. Uchida</th>
<th>Majority</th>
<th>NY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W</td>
<td>R</td>
<td>W</td>
<td>R</td>
<td>W</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>8-1</td>
<td>10</td>
<td>8-1</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>113</td>
<td>115</td>
<td>111</td>
<td>115</td>
<td>113</td>
</tr>
</tbody>
</table>

TABLE 9: Scorecards for Whitaker-Rivera II. Whitaker’s scores are in the columns marked “W”, Rivera’s in the columns marked “R”.

The official scorecards for Whitaker-Rivera II, and the majority scorecards constructed under the Majority and NY Systems, are shown in Table 9. In the fight, the Majority/NY Systems award Rivera victory in six rounds: 2, 4, 5, 9, 10, and 12. What is interesting is that none of these rounds is unanimous; one judge has a different score from his peers in each one of these rounds. In contrast, the three judges agree on the score of three of the five rounds which are given to Whitaker on the consensus scorecards. One plausible explanation is that the rounds Whitaker won on the Majority/NY cards were “clear” Whitaker rounds, while the rounds Rivera won were “close” or uncertain rounds. By forcing every round to have one score on the master scorecard, the Majority/NY Systems make no such distinctions; in this case, Rivera’s success in the “close” rounds would have cost Pernell Whitaker the victory.

More recently, Stevie Johnston was named the WBC Fighter of the Year in 1999, a year which began with him regaining the WBC lightweight title from Cesar Bazan. However, if the Majority or NY System were in effect for the title fight against Bazan, Johnston would have never regained the title. Under the 10-Point Must System, Johnston won a split decision: 115-112, 114-113, and 112-116 on the judges’ scorecards. Under the Majority system, the score of the bout is 113-113, and 114-114 under the NY system. Both the Whitaker-Rivera II and Johnston-Bazan II were generally perceived as definitive performances by Whitaker and Johnston, respectively, following sub-par performances in their initial bouts with these opponents. Under the Majority/NY Systems, both these rematches would have resulted in draws, seemingly depriving Whitaker and Johnston of their hard-fought “revenge” against their opponents.

13In 1999, Johnston four fights and four victories, winning the title from Bazan and then defending it successfully against Aldo Rios, Angel Manfredy, and Billy Schwer.
### Table 10: Scorecards for Carbajal-Pastrana. Carbajal's scores in the columns marked “C”, Pastrana's in the columns marked “P”.

<table>
<thead>
<tr>
<th>Round</th>
<th>M. Glienna C</th>
<th>J. Roth C</th>
<th>D. Ford C</th>
<th>Majority C</th>
<th>NY C</th>
<th>M. Glienna P</th>
<th>J. Roth P</th>
<th>D. Ford P</th>
<th>Majority P</th>
<th>NY P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>116</td>
<td>112</td>
<td>113</td>
<td>115</td>
<td>114</td>
<td>115</td>
<td>113</td>
<td>115</td>
<td>115</td>
<td>113</td>
</tr>
</tbody>
</table>

An example of a split decision being “reversed” under the Majority/NY Systems is the Michael Carbajal-Mauricio Pastrana bout. The scores for this bout are in Table 10. Of the decision, Borges (1997) notes that “about the only person who thought Carbajal had won the bout was a judge, because not even Carbajal was willing to argue for himself after it was over.” Yet Carbajal wins the fight 115-113 on either the Majority or NY System scorecard. Table 10 shows that judges Roth and Ford agree on eight rounds (rounds one through seven, and round ten); since they agree on these rounds, they define the master scorecard entry for these rounds. Of these eight rounds, five are scored for Pastrana and three for Carbajal. On every round which Roth and Ford disagree—rounds 8, 9, 11, and 12—judge Glienna scores the round for Carbajal. Applying the Majority/NY systems, these rounds then become Carbajal rounds, and Carbajal wins the bout under these systems, seven rounds to five (115-113). Carbajal-Pastrana is a good example of a fight where the scoring of a minority judge drives the master scorecards of the Majority and NY Systems and reverses a split decision.

## 5 Conclusions and Recommendations

The May 2000 report of the National Association of Attorneys General Boxing Task Force states, “While the Task Force is not specifically recommending consensus scoring or any other system of scoring, it does recommend that state commissions evaluate and consider such alternative systems.” Two “alternative” systems are specifically mentioned in the report, Dr. Ralph S. Levine’s “consensus” scoring procedure (which this paper refers to as the NY System) and New Jersey’s 10-Point Majority System. This report examines these two systems, as well as the status quo 10-Point Must System. The analyses of this paper have yielded several findings about these scoring systems by both considering the conceptual properties of the systems and applying them to the judges’ scorecards of all world title fights sanctioned by the WBA, WBC, and IBF (1986–1999) and WBO (1990–1999) that went to
judges’ decision. We believe that this paper constitutes a thorough evaluation of these scoring systems, as called for in the Boxing Task Force report, and provides much-needed information in a debate area too often framed around the performance of a proposed system on one fight only.

Theoretically, the Majority and NY Systems are very similar in that they return the same decision in rounds where judges score 10-9 or 10-10, the most commonly occurring round scores (Table 2). In practice, we find that the Majority and NY Systems are virtually identical. In its current form, the 10-Point Majority system inappropriately treats point deductions by the referee (see Section 3.3.1), so that a single-point deduction can effectively count as either a zero- or two-point deduction, depending on the judges’ scores for that round. If the 10-Point Majority System remains in use, these inconsistencies can and should be rectified by applying the Majority System procedure first, then taking the point deduction, rather than deducting the point first. If such modification is made, then the (corrected) Majority System and the NY System return the same decision for every one of the 956 fights in the dataset. Given that they are virtually equal, we believe that—were it desired to implement a consensus scoring method—the NY System would be preferred, simply because the procedure used in the NY System is easier to explain and simpler to understand than the more convoluted process the Majority System uses to arrive at a round score.

We find that many of the claims made in support of the Majority and NY Systems to be unsubstantiated. The claim that the Majority System gives the fight to the fighter with the most total points is simply false. Conceptual counterexamples to this claim are shown in Tables 3 and 4; more telling, on the 956 fights examined in this study, the Must System awards the fight to the boxer with the most points more often than the Majority system. As discussed in Section 3.3.2, we do not agree that using the most total points as the sole criterion for victory is a sound idea, in light of the fact that boxing is a subjectively scored sport.

We also find no support for the claims that the Majority and NY Systems are less vulnerable to the influence of one “bad” judge than the Must System. The Must System’s principle of one judge, one vote makes it clear in advance how much weight a single judge’s vote has on a decision; moreover, since a fighter needs two judges in his favor in order to win, a single “rogue” judge alone cannot determine the outcome of the fight. The same is not true for the Majority and NY Systems; the influence of one judge depends on the level of agreement between the remaining two judges. As shown conceptually in Section 3.3.3 and in the real-life example of the Michael Carbajal-Mauricio Pastrana fight (Table 10), it is possible for Boxer A to win the contest even if two judges score the bout for Boxer B; the phenomenon where the majority judges’ verdict is rejected by the Majority/NY Systems occurs 62 times in the 956 fights in the dataset.

The question of whether consensus scoring systems would reduce the amount of “bad decisions” in boxing has usually been answered by pointing to their performance on one or two specific fights (most recently, Holyfield-Lewis I). However, we find that while the Majority or NY Systems would resolve three important “bad decisions”—Leonard-Hearns II, Whitaker-Chavez, and Holyfield-Lewis I—in a manner consistent with public sentiment, the systems would actually induce more “bad decisions.” They would do so either by converting draws into victories that are not consistent with public opinion of who won the fight (e.g. Azumah Nelson-Jeff Fenech I scored for Nelson and Lonnie Bradley-Otis Grant scored for Bradley) or by converting unanimous decisions—cases where all three judges agree, in their scorecard totals, that one fighter on the bout—into reversals or draws. This latter phenomenon occurs nine times in our 956-fight dataset, including the relatively high-profile Whitaker-Rivera II (Table 9).
One claim about the Majority/NY Systems which is supported, in part, by our analysis is the argument that the Majority System would "minimize draws." Under the Must System, 47 out of 956 fights result in draws, while 43 out of 956 fights end in draws under the Majority/NY Systems—roughly, one less draw every 3.5 years of title fight decisions. However, while there are fewer draws under the Majority/NY systems, the fights which are identified as draws are not limited to cases that were draws under the Must System. Only eleven of the 956 fights are identified as draws under both the Must and the Majority/NY systems; the remaining 32 fights which the Majority/NY system identifies as draws have a winner under the Must system.

Based on the preceding analysis, we conclude that if boxing commissions and sanctioning bodies face a choice between the 10-Point Must System or either of the consensus scoring methods, the choice is clear: the 10-Point Must System is, and remains, the best of the three systems on both conceptual and empirical grounds.

The Must System avoids the trap of assuming that there is one, unequivocally correct score for each round of boxing. The descriptions and justifications of the Majority and NY Systems suggest that the only reason why there would be disagreement among judges would be if one of the judges had their view obstructed, was incompetent, or was dishonest; the consensus systems would then drop this odd judge's score. However, this argument is at odds with a more compelling reason for disagreement between professional boxing judges—namely, that some rounds of boxing are very close and very competitive, making them tough to call. In rounds like these, manufacturing a single "master" score from the three judges' scores oversimplifies the scoring process and can lead to consensus scorecards which are inconsistent with all of the individual judges' opinions. In the presence of close rounds, on which two judges disagree, a single minority judge could swing the fight under the Majority System—in some cases, sufficiently so to reverse a unanimous decision in which every judge's scorecard total favors one fighter (see Tables 4 and 9). Such reversals run counter to the claim that the Majority or NY Systems would "enhance the integrity of boxing;" they would likely tarnish the sport further.

Boxing fans understandably are disappointed with some draws that crop up in professional boxing matches; the lack of a definitive winner is frustrating. However, we question whether "minimizing draws" is actually a desirable objective. Los Angeles Times writer Earl Gustkey (1988) believes that "a dead heat in boxing match is physiologically and mathematically impossible"—that "a draw is not a decision, It's a cop-out, by a judge who can't make a decision" (Gustkey, 1988). But this argument glosses over the actual process of scoring a bout; it seems predicated on the idea that a judge makes one decision at the end of a fight—Boxer A, Boxer B, or draw. This is simply not the case; it is the combination of individual round scores submitted by each of the three judges that factors into a verdict on the fight. One might argue that a judge's propensity to score even rounds reflects indecisiveness, but a draw on a judge's total scorecard must be understood as the end result of multiple round-by-round decisions. Moreover, public and media observers of a fight are not bound to the same scoring criteria as official judges. Observers have the luxury of unofficial punch statistics and expert opinions offered on television broadcasts and the possibility of consulting their peers in forming a decision—none of which is true for official judges at ringside; they also have the option (and, perhaps, propensity) to "score" a bout based on their overall impression of a fight, rather than documenting a specific score for each round. More fundamentally, we have noted that there are rounds (and entire fights) in professional boxing that are very competitive and difficult to call, one way or the other. Duane Ford, boxing judge and former Chairman of the Nevada State Athletic Commission notes that "judges look at each round as a (separate) fight, and most don't keep a running score. If at the end, it comes out even, then that's
how it is. If three judges can’t agree, then a draw is a good decision. What’s the alternative, flipping a coin? Drawing Straws?” (Gustkey, 1989). We believe that there are close rounds and close fights in professional boxing, and that a draw is not necessarily a bad outcome or an indecisive “cop-out.”

Still, are there too many draws in boxing? Our data shows that roughly 5% of all title fights that go to decision under the Must system are ruled draws. Keep in mind, many title fights end in knockout, so the actual percentage of title fights that end in a draw is substantially less than 5%. The question before policy makers and boxing commissions is whether 5% of decisions ending in draws is too high. If not, then we feel that the 10-Point Must System is superior to the consensus scoring methods considered in this paper. If this proportion of fights ending in a draw is deemed too high, though, then one possibility to consider is a mixed approach: use the 10-Point Must System as the scoring system, but use the NY System as a tie-breaker in the event of a split decision draw. The split decision draw is the only type of decision under the Must System where at least two judges do not agree on the outcome (see Table 1). Such a hybrid approach would resolve some unsatisfying draws (notably, Holyfield-Lewis I) but not allow for the detrimental possibility of a unanimous decision being reversed. It is important to consider, though, that the Majority/NY systems do not always resolve draws in a fashion consistent with public opinion, such as the Nelson-Fenech I and Bradley-Grant bouts. Again, if boxing deems that too many fights end in draws (and we do not believe this is the case), then a hybrid approach using the NY System may be a way to get some benefit from consensus scoring without encountering the more egregious problems that would result if such a system were implemented fully.

As boxing undergoes attempts to clean up its often-unsavory public image, the system by which boxing matches are scored and decided will understandably undergo scrutiny. From our analysis, we believe that a rush to “improve” boxing scoring by changing the way the scores are combined to reach a decision is misguided; neither the Majority nor the NY System outperforms the existing Must System. But we do believe that it is vitally important that the sport continue to make strides to ensure that individual judges’ scores are honest and well-informed. To that end, workshops and seminars like the WBC’s recent Ring Officials Congress (Feour, 2000a) are vital; giving official judges more training and experience will help ensure quality decisions. The New York Athletic Commission recently adopted reforms designed to alleviate structural impediments to the ringside judges’ ability to see the fight fully, by raising the judges’ seats an additional six inches and prohibiting still photographers from entering zones three feet to either side of a judge (Associated Press, 2000). Ultimately, the quality of boxing decisions depends on the quality of boxing judges—people who are asked to filter large amounts of information through several criteria and instantaneously produce a round score. Improving the education and training of boxing judges—not post hoc tweakings of the way scores are combined—is the best way to ensure that boxing decisions are honest (if not always popular).
References


