An Analysis of Scoring Methods for the Women’s Heptathlon

Hailey Dammeier

SENIOR HONORS THESIS

Submitted in Partial Fulfillment of Requirements of the
College Scholars Program
North Central College

Date

Approved: _______________________________ Date: __________
Thesis Director Signature
David Schmitz

Approved: _______________________________ Date: __________
Second Reader Signature
Diane Anstine
Abstract

The current scoring method of the women’s heptathlon creates an imbalanced point distribution across the seven events. Each event equation standardizes performances into points based on the ideal of all-around talent; however, studies of point distribution in the NCAA from 2015 to 2018 reveal an undervaluation of the throwing events. Other researchers addressing this issue propose new scoring methods or changes to the events within the heptathlon, but fail to consider their relationship. Because a majority of the events rely on sprint speed, most athletes succeed by capitalizing on those events and sacrificing points in throwing events. To correct the imbalance within the event structure of the heptathlon, we propose applying a multiplier to weight the equations based on the skills of sprint speed and technical strength. This establishes a point premium on underrepresented events to more effectively measure all-around talent and forms pathways for more diverse athletes to succeed.

Section 1

Introduction

The heptathlon is a women’s track and field event consisting of seven events completed over two days. The first day includes the 100 meter hurdles (100H), the high jump (HJ), the shot put (SP), and the 200 meter dash (200m), and the second day includes the long jump (LJ), the javelin throw (JT), and the 800 meter run (800m). At least a 30 minute time period must elapse between the end of one event and the beginning of the next to allow for warmups and rest. Because these events cover a wide variety of disciplines and require many types of athletic talents, the winner is titled the “greatest athlete in the world” (Donovan, 2016). The heptathlon is not the only combined event, an amalgamation of several events into one competition, in track and field. The men’s decathlon adds the pole vault, discus, and the 100m dash; furthermore, it lengthens the 100H to 110 meters and the 800m to 1,500 meters. Because
of their similarities, research on the decathlon often relates to the heptathlon as well, and the training mechanisms and scoring techniques are intimately related.

The seven events within the heptathlon each has its own independent history, skill base, and technique. The following descriptions provide background and skills needed to be successful in each event.

100 Meter Hurdles

The 100 meter hurdles is a sprint event in which the athlete runs 100 meters while also clearing 10 hurdles spaced evenly throughout. As the shortest running distance in the heptathlon, a great amount of sprinting speed is needed to be successful. The athlete begins in blocks, a metal contraption allowing the athlete to begin the race in a crouched position, which aids in the acceleration through the first few steps. Because of the obstacles, the maximum speed reached in the hurdle race is less than in the 100 meter dash. Hurdle technique attempts to minimize this decrease in speed during hurdle clearance. (Tsiokanos, Tsaopoulos, Giavroglou, & Tsarouchas, 2017).

High Jump

The high jump technique has evolved immensely over the past 100 years. At first, high jump was competed in both standing and running forms without a mat to cushion the fall. Because there was no mat, two techniques were commonly used: the Western and Eastern Rolls. The Eastern Roll was the first high jump technique widely used. In this technique, the athlete would take off with the foot farther from the standards and clear with their backs to the bar (Templeton, 1926). High jumpers using the Western Roll would take off of the foot closest to the standards and clear with their stomach facing over the bar (Templeton, 1926). In these techniques, the jumper would land on their feet. After the introduction of the mat, however, these techniques became obsolete as Dick Fosbury invented the current “Fosbury Flop” method of high jump in 1968 (Goldman, 2018). In this technique, the jumper
runs along a curve towards the bar and jumps off of the outside foot similar to the Eastern Roll, but clears the bar with the arms and head first, then raises the hips over the bar, and finally inverting their positioning to allow the legs to clear. They land on the mat with their shoulders first, an impossible position without a cushioned mat. Within 8 years of Fosbury winning gold in the Olympics using this revolutionary technique, no medals were ever won again by anyone using another technique (Goldman, 2018).

There are several different variations to the Fosbury Flop (Pavlovic, 2017). There are speed and power jumpers, right and left footed jumpers, single and double armed jumpers, and various knee angles that all create the greatest conversion from horizontal to vertical motion for different athletes (Pavlovic, 2017). Finding the perfect technique and speed for a specific jumper to transfer horizontal to vertical speed can be a long process and requires extensive coaching and experience (Templeton, 1926). In fact, the amount of speed brought in the approach has only a 0.79 correlation with the height reached by the jumper (Pavlovic, 2017). Even research from the 1920s realized that a fast approach produces a distance effect rather than a height effect, and too much speed cannot be converted or controlled by the takeoff leg (Templeton, 1926). Because of this, high jump, though partially reliant upon speed, is mostly reliant upon strength and technique to launch the jumper over the standard.

**Shot Put**

There are two techniques generally used in the shot put: rotational and glide. Both methods attempt to gather the greatest amount of momentum while in the ring to provide the most amount of power to the throw. The ring the athlete must throw within is only 2.135 meters in diameter, so forward speed or momentum is not extremely beneficial. Instead, both methods rely heavily upon creating torque from the rotation of the body (Błażkiewicz, Łysoń, & Chmielewski, 2016). The energy transfer from this rotational element creates more energy transfer than is possible from forward muscle
movement in the human body, which further emphasizes the reliance on explosive energy and technique rather than speed (Błažkiewicz et al., 2016).

An athlete’s ability to throw the shot put depends at least partially on body type, too. In a study of the Polish National Team, a positive correlation of 94% was found between the distance thrown and increased body mass (Błažkiewicz et al., 2016). A higher body mass allows for greater muscle mass to power the throw; however, the excess weight would be detrimental in other speed-based events within the heptathlon. In fact, a study on body types in Division I track and field athletes revealed an average of a 50 pound difference between the weight of sprinters and throwers (Hirsch, Smith-Ryan, Trexler, & Roelofs, 2016). Shot put relies nearly entirely on an athlete’s strength and requires conflicting body composition with the other events.

200 Meter Dash

The 200 meter dash is the final event on the first day. Athletes run halfway around the track as fast as possible after starting out of blocks. This event is categorized as a short-sprint event and relies nearly entirely on an athlete’s ability to sprint.

Long Jump

Long jump consists of 4 phases: the run-up, the take-off, the aerial phase, and the landing phase (Guiman, 2015). While all 4 phases are important in the final distance of the jump, the first phase creates the horizontal momentum that causes the greatest horizontal distance (Guiman, 2015). The run-up distance is measured as the number of steps needed to reach maximum speed. The take-off phase attempts to convert the horizontal speed of the run-up into vertical speed to combat gravity and carry the athlete far into the sand. The proper angle for takeoff typically falls between 20° and 30° (Guiman, 2015). Third, the aerial phase occurs after the athlete has left the ground and focuses on controlling momentum as aerodynamically as possible before landing (Guiman, 2015). Finally, the landing phase
occurs as the athlete attempts to straighten their legs in front of them until impact and then pull their body forward to where their legs landed (Guiman, 2015).

The first phase depends entirely on sprinting speed and creates most of the distance in the jump (Guiman, 2015). The next three phases are mostly strength based as the athlete attempts to convert horizontal speed into a slight vertical component and to engage the core muscles to control their momentum and implement proper landing technique. The majority of the athlete’s success in this event, however, results from the speed of the approach.

**Javelin Throw**

A javelin’s flight depends primarily on 3 factors: release speed, release angle, and release height (Saratlija, Zagorac, & Babić, 2013). The release speed depends partially on the speed of the approach. The standard javelin approach is a straight running motion converted into a series of lateral cross-overs before throwing. The cross-overs slow the body down and make the final positioning more accessible. The proper technique translates as much of this forward momentum as possible into the javelin when it is thrown. This translation of energy is the most important factor on distance in this event (Saratlija et al., 2013).

The ideal release angle based on projectile motion and air resistance is between $33^\circ$ and $36^\circ$ (Saratlija et al., 2013). This creates the farthest horizontal distance possible. The ideal height of release is 105% of the athlete’s body; both the release angle and height rely solely on the athlete’s technique, and excess speed can hinder the athlete’s control (Saratlija et al., 2013). Based on this study, it would appear the javelin throw heavily relies upon both strength and speed to maintain control and launch the javelin correctly.
800 Meter Run

The 800 meter run is the final and longest event of the women’s heptathlon. It consists of running two laps around the track and is classified as a middle-distance event, a hybrid between a sprint and endurance event. This race is the only individual event that includes an element of endurance in the heptathlon.

While these seven events collectively comprise the current women’s heptathlon and serve as an evaluator of overall athleticism, the method to discern the greatest athlete has been under debate for centuries. The first combined event pentathlon occurred around 700BC in the Greek Olympic Games with a combination of jumps, throws, sprints, and wrestling (International Association of Athletics Federations, 2016). A more modern pentathlon occurred hundreds of years later in the 1851 Olympic Games, where wrestling was replaced with climbing a rope. A version similar to the current decathlon was submitted to the 1912 Stockholm Olympics, which had to be competed over three days to accommodate the number of entries (International Association of Athletics Federations, 2016). Winner Jim Thorpe was declared “the greatest athlete in the world” by Swedish King Gustaf V, and the Olympic champions in both the decathlon and heptathlon continue to unofficially receive this title (Donovan, 2016). The current men’s decathlon began its official recognition in 1914 (International Association of Athletics Federations, 2016).

Women’s athletics progressed much more slowly, with the women’s pentathlon and heptathlon not accepted into the Olympic Games until after World War II (International Association of Athletics Federations, 2016). The societal uncertainty surrounding women’s athletics prevented a unified, international body from organizing the event, allowing for much greater variation in the events contested over time. The following table shows the progression of the event structure in both the women’s pentathlon and heptathlon.
Though the current scoring model for the heptathlon was adopted in 1985, only two years after the adoption of final heptathlon event variation, it is beneficial to examine how the scoring of men’s and women’s events evolved and how the value of all around talent culminated into the current equation.

**Prior to 1884**

Competitions were scored based on the athlete’s respective place in every event with the lowest total sum as the winner (International Association of Athletics Federations, 2016). Using this method, it was nearly impossible to compare athletes in different meets, and it failed to take into consideration the margin of victory and loss in any given event. For example, if one athlete outperformed all other athletes by a large margin in one event, with all else equal, the athlete could lose to another who bested him by an infinitesimal margin in another two events.
The Linear Model

The Linear Model, used from 1884 to 1934, marked the first use of equations to discern the winner of the overall event (International Association of Athletics Federations, 2016). To create the model, the IAAF simply drew a straight line between (0,0) and an allocation of 1,000 points to the world record based on the 1908 Olympic records (International Association of Athletics Federations, 2016). Because this method scores based on performance rather than place, individual scores do not depend on the competition, allowing comparability between meets and point distribution based on the greatness of the performance.

The linearity of this model posed issues, too, for the same increase in performance was awarded the same number of points regardless of how exceptional the performance. Besting the world record by an inch in the shot put is indisputably more difficult and less probable than the same improvement of a mediocre put. Because improvement in an event becomes more improbable as one approaches the limit of the human body, it was argued that athletes should receive points on a progressive rather than a linear scale (International Association of Athletics Federations, 2016).

The earliest tables also required decimal results exceeding two digits (International Association of Athletics Federations, 2016). Coaches and athletes disliked this aspect of the tables, forcing a revision in 1920 to operate only in integers (International Association of Athletics Federations, 2016). The current model incorporates this preference by rounding down all scores to the nearest whole number.

The Progressive Model

The Progressive Model was the first attempt to add progressivity into the equation and debuted in 1934 for the men's combined event (Trkal, 2003). Progressivity awards greater points to greater performances, addressing the issues in the linear model by replacing the straight line with an exponential curve. The equation, completed by J. Ohls, was based on the equation $P = e^M$ where $P$ is
the number of points awarded, e is the natural base of logarithms, and M is the performance or mark completed by the athlete (Trkal, 2003).

The first women’s tables were formed during this time, but they did not receive the same guidance as the men’s equation. Germany’s Deutsche Sportbehörde fur Leichtathletik (DSfl) published the women’s tables, nicknamed the “100-points” table because the maximum number of points was 100 for each event (International Association of Athletics Federations, 2016). This was also a progressive curve, but it was not drawn up by mathematicians; rather than a smooth curve, the curve was a progression of straight lines connected at angles to create the different progressive slopes (International Association of Athletics Federations, 2016). Despite the comparatively crude nature, this table continued to be used until 1954.

Post-War Tables

Following World War II, training techniques and technology greatly improved, leading to imbalances in the scoring equations as athletes improved disproportionately throughout the events. In 1950, Swedish scientists Gosta Holmer and Axel Jörbeck designed a new set of progressive tables with a maximum point total of 1500 points per event (Trkal, 2003; International Association of Athletics Federations, 2016). The researchers used Performance Curves, developed using average speeds required to break world records, to determine where to set the key equivalent performance scoring, the point at which the difficulty is equal across events, and used 1,300 points for the most influential equal point (International Association of Athletics Federations, 2016). This was the first time true scientific measurement entered the scoring tables rather than intuition from experience in the sport. In 1952, these tables were reconsidered and altered due to excess progressivity. The same basic concepts were applied, but the progressivity was reduced to create fairer scoring.
The women’s tables were updated in 1954 after examining a study by Karl Uhlbrich, who determined that the velocity, rather than the speed, of the runner or object thrown should be converted into points (International Association of Athletics Federations, 2016). The resulting tables created by Jörbeck lasted for the next 30 years due to his fine use of statistical data relating to women’s athletics. The men’s tables were recalculated in 1962 using this new technique as well (Trkal, 2003). Despite its relative success, Ulbrich’s method posed certain issues. For running events, where the human body is the measure of velocity, the curves were progressive. For the field events, where an object or body is projected both upwards and outwards, the velocity must be calculated from the distance travelled. According to the Ulbrich principles, the initial velocity of the attempt can be calculated through the square root of the performance (International Association of Athletics Federations, 2016). The result creates regressive tables for field events, meaning the growth in the number of points as an athlete improves is smaller at higher levels (International Association of Athletics Federations, 2016). Finally, only slight adjustments to the 1954/1962 tables were needed in the 1970s to account for the inception of electric timing, with which times were recorded to the hundredth of a second, and for changes in the event structure (Trkal, 2003; International Association of Athletics Federations, 2016).

The Current Table

In the next twenty years, changes in technology and technique created imbalances in the scoring of events. Such changes included new plastic fibers for the pole vault in the men’s decathlon, the introduction of the Fosbury Flop in 1968, and improvements in footwear (Trkal, 2003). Based on the 1962 tables, what was then becoming an average pole vault at 5.10 meters scored approximately the same number of points as a nearly world record breaking 100m dash at 9.99 seconds (Trkal, 2003). Additionally, the regressive nature of the field events was angering athletes as their improvements at high levels rewarded them few additional points. This posed a psychological effect on the athletes, who despaired at the lack of points possible in the field events, and after reaching a certain point in the
tables, found no benefit from trying to improve in those events (Trkal, 2003). These imbalances created pressure on the sports community to create new equations to balance out these changes.

The 1984 tables used kinetic energy emitted by the athlete to create a slightly progressive curve. The creator of the tables, Victor Trkal, purposely weighted the longest distance events, the 1500 and 800 meter runs in the decathlon and heptathlon, slightly higher because of the unique talents required for these compared to the other events (Trkal, 2003). Since 1984, the tables have not been updated or changed significantly.

The current formula adopts the following equations for the field and running events, respectively:

\[ P(x) = A(x - B)^C \]

\[ P(x) = A(B - x)^C \]

where \( P(x) \) is the total points awarded from the event rounded down to the nearest integer, \( x \) is the athlete’s performance, \( A \) is the constant based on kinetic energy, \( B \) is a threshold performance required to earn one point, and \( C \) is the exponent providing progressivity to the curve (World Sports Intelligence, 2012). The following table shows the constants for each event:
An Analysis of Scoring Methods for the Women’s Heptathlon

### Heptathlon Events

<table>
<thead>
<tr>
<th>Event</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Javelin Throw</td>
<td>15.9803</td>
<td>3.8</td>
<td>1.04</td>
</tr>
<tr>
<td>Shot Put</td>
<td>56.0211</td>
<td>1.5</td>
<td>1.05</td>
</tr>
<tr>
<td>High Jump</td>
<td>1.84523</td>
<td>75</td>
<td>1.348</td>
</tr>
<tr>
<td>Long Jump</td>
<td>0.188807</td>
<td>210</td>
<td>1.41</td>
</tr>
<tr>
<td>100 Hurdles</td>
<td>9.23076</td>
<td>26.7</td>
<td>1.835</td>
</tr>
<tr>
<td>200m Dash</td>
<td>4.99087</td>
<td>42.5</td>
<td>1.81</td>
</tr>
<tr>
<td>800m Run</td>
<td>0.11193</td>
<td>254</td>
<td>1.88</td>
</tr>
</tbody>
</table>

Figure 2. Heptathlon equation constants (International Association of Athletics Federations, 2016)

Though this set of equations has been used for over 30 years, imbalances in scoring still occur. Studies of elite heptathletes competing in the World Athletics Championships and the Olympic Games from the years 1987 to 2013 found average scores in the 100H of approximately 1,050 points while mean javelin scores only reached approximately 775, with no competitors reaching more than 1,000 points (Gassman, Fröhlich, & Emrich, 2016). These point discrepancies reveal lingering imbalances within the current scoring model and suggest further revision is necessary.

Additionally, Gassman et al. (2016) found that heptathletes tended to fall into two categories: specialists and generalists. Specialists tended to score more points in the 100H, 200m, and LJ, while generalists tended to score more evenly throughout the events (Gassman et al., 2016). The generalist scoring method better embodies the value of an all-around athlete, but results found that specialists not only tend to outscore generalists by about 170 points, specialists are also 12% more likely to appear in elite competitions (Gassman et al., 2016). These results appear to contradict the value of determining the best all-around athlete, and again suggest the deficiency of the current scoring model.
The following guidelines were proposed and used to create the current model and can be used to determine best practices for revisions:

1. The tables should only be used for combined events.
2. The results in different disciplines that are evaluated with approximately the same point value should be comparable as far as the quality and difficulty of achieving these results are concerned.
3. The tables in all disciplines should be:
   a. a modification of current tables
   b. linear in all disciplines
   c. very slightly progressive in all disciplines
4. The tables must be usable with combined events for beginners and juniors as well as top-class athletes.
5. There will be separate tables for men and women.
6. The tables must be based on decathlon statistics, taking into account the statistics of specialist athletes in the individual disciplines.
7. The tables should be usable now and in the future.
8. The sum of points scored by world-class athletes should remain approximately the same.
9. As far as possible, the tables should eliminate the possibility that an athlete specializing in one discipline is able to acquire sufficient points in that discipline to overcome low scores in weaker disciplines and beat more versatile, all-round athletes (International Association of Athletics Federations, 2016).

These guidelines will serve as a basis for judging the proposed equations.
Methodology

Data from results of 4 years of NCAA collegiate data were obtained from a public database. The data was then transferred, checked for accuracy through recalculation, and ranked. Total scores reported without a breakdown of event performances were removed from the data set and comprised of less than 0.3% of the sample population.

No human subjects were used to collect this data, rather, it was found using the Track and Field Results Reporting System (TFRRS) as sponsored by the United States Track and Field and Cross Country Coaches Association (USTFCCCA). This database organizes the results for collegiate competition in the NCAA and NAIA as well as other smaller associations. The sample population includes the top performance of every athlete who competed in the women’s heptathlon in years 2015, 2016, 2017, and 2018 in NCAA Divisions I, II, and III. This dataset satisfies the 4th guideline above, for the high performing athletes in Division 1 are those who could eventually compete at the elite level, while the freshmen in the study and many of the athletes in Divisions 2 and 3 with no aspirations to compete at the elite level would represent the beginner and junior level competitors.

Data Analysis

An analysis of the collected data provides similar results as research on elite athletes. When comparing the average scores in each event from the sample, one sees that the 100H followed by the 200m score the most points while the shot put and javelin score the least points.
An Analysis of Scoring Methods for the Women’s Heptathlon

In an event stressing the importance of all-around athleticism, it appears unfair that two events with arguably the highest correlation in required skills achieve the greatest number of points. A percentage breakdown of the total score further suggests the athlete’s reliance upon points earned in those two disciplines over the throwing disciplines.

<table>
<thead>
<tr>
<th>Event</th>
<th>Percentage of Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>100H</td>
<td>17.14%</td>
</tr>
<tr>
<td>HJ</td>
<td>15.19%</td>
</tr>
<tr>
<td>SP</td>
<td>12.12%</td>
</tr>
<tr>
<td>200m</td>
<td>17.19%</td>
</tr>
<tr>
<td>LJ</td>
<td>13.42%</td>
</tr>
<tr>
<td>JT</td>
<td>10.31%</td>
</tr>
<tr>
<td>800m</td>
<td>14.63%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 4. Average percentage breakdown of total scores under current method

If all events were weighted equally, each event should contribute \( \frac{1}{7} \), or 14.3% of the score. Based on the above data, the 100H and 200m contribute significantly more than the two throwing events. The total contribution of 100H and 200m is nearly 35% of the total score, while the throws disciplines contribute just 22%. According to researcher Wim Westera (2007) of the Open University of the Netherlands, “when starting from the principle of all roundness, the ideal score distribution should be
uniform over the disciplines” (Westera, 2007). Therefore, the current model has failed to fairly distribute points across the events.

A study of the decathlon in 2000 examined the effect of the scoring model change in 1986 and revealed a shift in the accumulation of points. In fact, the performances in the 1,500 meter run, the men’s counterpart to the 800m, and the high jump declined, while the throws disciplines actually began performing at higher levels (Tidow, 2000). The differences in performances demonstrate how the values embedded in the equation affect the training and success of combined event athletes. Altering the current model may create more balanced scoring through changing the implicit values represented within the heptathlon equations.

Section 2

Alternate Scoring Methods

Several researchers have proposed solutions to the current model with varied success. It is imperative to note that any suggested method will incorporate some amount of bias. The current method was devised in an attempt to value the various disciplines on the principle of the all-roundness of the athlete, yet appears to favor the 100H and the 200 based on the score distribution. An analysis of fairness and unfairness depends on a person’s experiences and perception. In other words, no truly objective or correct answer exists, rather, the following are other biased possibilities attempting to improve the current system.

Special Totals

John Barrow (2013) from the University of Cambridge created an entirely new method of scoring the men’s decathlon that could also be adapted to the women’s heptathlon using a “Special Total”. He
created this method by grouping events based on those measured in time versus those measured in distance (Barrow, 2013). To that end, his converted heptathlon equation would be

\[ ST = \frac{LJ \cdot HJ \cdot JT \cdot SP}{100H \cdot 200m \cdot 800m} \]

where \( ST \) is the special total or total score achieved, and the event variables are entered in either seconds or meters. Though an interesting attempt, entirely new issues arise. For instance, a change of 0.5 seconds in the 100H is a significant improvement, but it would be completely voided by a very slight change in the 800m of 0.6 seconds. Because of the differences in times and distances between events, this equation would weigh the largest measured performances, the JT and the 800m, greater than those with the smallest measured performances, the HJ and the 100H. It also completely alters the current equation, lacks progressivity, and fails to utilize statistical data, all violations of the IAAF guidelines for a new equation.

The Power, Parabolic, and Exponential

The previously mentioned Wim Westera (2007), a researcher from the Open University of the Netherlands, also proposed several revision possibilities. His solution derives from changing the disproportionate exponents to create a universal exponent. The universal exponent would potentially increase the points in the throwing and jumping events, where the exponents are smaller, to create more even scoring. He hypothesizes that the difference in the exponents could be a factor in the imbalance of points (Wester, 2007). He creates two different types of equations:
An Analysis of Scoring Methods for the Women’s Heptathlon

\[
S(P) = A \cdot \left( \frac{P - P_0}{P_1 - P_0} \right)^C
\]

and

\[
S(P) = A \cdot \frac{e^{\lambda P_N} - 1}{e^\lambda - 1}
\]

where \( S(P) \) is the total points for the event, \( A \) is a constant 957.83 to create a similar total score to the current method, \( P \) is the athlete performance, \( P_0 \) is the threshold value (the equivalent to \( B \) in the current equation), \( P_1 \) is an event specific constant to standardize the equation, and \( C \) provides progressivity (Westera, 2007). The first equation is the basis for his first two models, the Power Law and Parabolic Model. The only difference between these two models is the \( C \) value. The second equation applies to the final of his three models, the Exponential Model, where \( S(P) \) and \( A \) mimic the first equation, \( \lambda \) is the exponent, \( P_N \) is the athlete performance, and \( e \) is the base of natural logarithms (Westera, 2007).

In his first model, named the Power Law, he creates a \( C \) value through the mean of the current equation’s \( C \) values, 1.481857 (Westera, 2007). In the Parabolic Model, Westera simply creates a perfectly parabolic curve by setting \( C = 2 \) (Westera, 2007). Finally, his Exponential Model minimizes the total squared differences between the curves, resulting in an exponent, \( \lambda \), of 1.6054 for all events (Westera, 2007). After rescoring the top 99 performances in the heptathlon of all time, he found that all athletes improving more than 20 places over their peers were considered “good throwers” (Westera, 2007). Notably, current World Record holder Jackie Joyner Kersee, who also holds the next 5 highest scores, is displaced by Russian Olympian Larisa Turchinskaya, the 24th best performance under the current model (Westera, 2007). Her strength in the throws and underrated performance again provide evidence of bias against the throws disciplines in the current model (Westera, 2007).
Despite providing evidence against the current model, his models are not entirely practical. Averaging the current exponents into a hybrid and creating a parabolic curve out of the data are not a statistical calculation of the events and serve only as reference curves for the Exponential model.

**Z-Score Model**

Another scoring methodology could be adapted from powerlifting research. Powerlifting consists of 3 events: bench press, squat, and deadlift. Bench press is performed by lying on one’s back and pushing upwards on a barbell. Squats are performed with the barbell across the shoulders and performing a squatting motion. Finally, deadlift is performed by grasping the barbell from the floor, and, while maintaining a straight back with the chest pressed upward, picking the weight up from the floor. Based on the physiology of the human body, the legs have a greater capacity to lift heavy objects than the arms; in powerlifting events, the greatest weight lifted is typically in the deadlift, followed by the squat, and then the bench press (Bishop, Williams, Heldman, & Vanderburgh, 2017). The scoring for powerlifting, however, fails to recognize the limitations of the human body and sums the total weight lifted, which creates an imbalance in the scoring (Bishop et al., 2017). Those who tend to win competitions are particularly good at the deadlift, while those with greater comparative strength in the bench press tend to compete at a disadvantage simply because the total weight lifted in the event is lower (Bishop et al., 2017).

In an attempt to amend this inequity, researchers chose to gather the available public data on weightlifting results and perform statistical analyses (Bishop et al., 2017). Specifically, they calculated z-scores for the data. Z-scores show the number of standard deviations above or below the mean calculated using the following equation:
where \( z \) is the \( z \)-score, \( x \) is the measurement (in this case the weight lifted), \( \mu \) is the average of the population, and \( \sigma \) is the standard deviation. Rather than valuing performances based on the amount of weight lifted, in which one pound represents greater difficulty in certain lifts over others, the performance is valued based on its distance from the average performance. This removes the imbalance caused by weight and creates a more objective scoring method.

To apply this scoring method to the heptathlon, the mean and standard deviations were calculated using the NCAA dataset. The results were then transformed to eliminate negative points and to maintain similar point totals to the current model. The average performance was set at 400 points and any performance below four standard deviations was allocated 0 points. The following chart summarizes the constants used for each event:

<table>
<thead>
<tr>
<th>Event</th>
<th>100H (s)</th>
<th>HJ (cm)</th>
<th>Shot (m)</th>
<th>200m (s)</th>
<th>LJ (cm)</th>
<th>JT (m)</th>
<th>800m (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>16.13567</td>
<td>150.7921</td>
<td>9.435868</td>
<td>27.14476</td>
<td>499.1959</td>
<td>27.68652</td>
<td>157.4952</td>
</tr>
<tr>
<td>Std Dev</td>
<td>1.752637</td>
<td>12.2022</td>
<td>1.493449</td>
<td>1.603351</td>
<td>49.80017</td>
<td>6.876849</td>
<td>14.61542</td>
</tr>
</tbody>
</table>

Figure 5. Constants for Z-Score Equation

After applying the \( z \)-score formulas, the data was analyzed again to determine if the scoring became more balanced. The following graph shows the average points scored in each event:
Figure 6. Score distribution after applying Z-Score Equation

Because the averages and standard deviations are based on this dataset, the average points scored across the events is nearly perfectly aligned at 400 points. This balances out any inequities in scoring tendencies between events; consequently, the percentage of the total score from each event is now leveled.

<table>
<thead>
<tr>
<th>Event</th>
<th>Percentage of Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>100H</td>
<td>14.41%</td>
</tr>
<tr>
<td>HJ</td>
<td>14.16%</td>
</tr>
<tr>
<td>SP</td>
<td>14.49%</td>
</tr>
<tr>
<td>200m</td>
<td>14.30%</td>
</tr>
<tr>
<td>LJ</td>
<td>14.15%</td>
</tr>
<tr>
<td>JT</td>
<td>14.23%</td>
</tr>
<tr>
<td>800m</td>
<td>14.24%</td>
</tr>
</tbody>
</table>

Figure 7. Percentage of total score by event

While at first appearing to be a nearly ideal solution, this equation only suits this dataset. The deep-rooted issue with this method arises in determining the appropriate mean and standard deviation with which to judge performances. The mean and standard deviations of the total population will continuously change as athletes compete. In powerlifting, the researchers recommend recalculating the mean and standard deviations at least every 2 years (Bishop et al., 2017). This makes using such a measure for the heptathlon extremely unlikely, for it would be nearly impossible to track world records.
with continuously changing equations or to gather the perfect dataset to fairly calculate these constants for all performance levels. Because this method only works through continuous updates and completely abandons the current model, it fails to be a reasonable solution.

Further issues are exposed from implementing this equation, too. To begin, the table below lists the constants again, but also lists the ratio of the standard deviation to the mean.

<table>
<thead>
<tr>
<th></th>
<th>100H (s)</th>
<th>HJ (cm)</th>
<th>SP (m)</th>
<th>200m (s)</th>
<th>LJ (cm)</th>
<th>JT (m)</th>
<th>800m (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>16.13567</td>
<td>150.7921</td>
<td>9.435868</td>
<td>27.14476</td>
<td>499.1959</td>
<td>27.68652</td>
<td>157.4952</td>
</tr>
<tr>
<td>Standard Deviation as % of Mean</td>
<td>11%</td>
<td>8%</td>
<td>16%</td>
<td>6%</td>
<td>10%</td>
<td>25%</td>
<td>9%</td>
</tr>
</tbody>
</table>

Figure 8. Standard Deviations as Percentage of Mean

The extremely high standard deviation in the javelin throw prevents the allocation of points beyond -4.00 standard deviations. Allowing standard deviations beyond -4.00 would award points to javelin throws thrown backwards. While in most elite cases, the cutoff measures would be moot, approximately 20 marks in the data set, about 0.6%, were allocated 0 points when they had earned some number of points in the current model. All of these discounted cases came from the running events. Only allowing points to four negative standard deviations in each event actually limits certain events more than others for beginners. The javelin, where a throw greater than 0.2m would still earn at least one point, then is at an advantage in comparison to the 200m, where the cutoff mark decreased from 42.50 seconds in the current model to 33.56 seconds. This could have a detrimental effect on
beginners and younger competitors. After examining these previously researched methods, the pattern of poor throws results coupled with high variability caused concern and demanded further analysis.

Section 3

Analysis of Event Performance

To determine the cause of the discrepancies, a comparison of performances using a platform other than the current equation or suggested solutions was required. The IAAF published an event comparison chart independent of any combined events tables that uses statistical analysis to allow a person to compare event performances between athletes in different disciplines (Spiriev & Spiriev, 2017). For example, a racewalker could be compared to a 4x400 meter relay or a hammer thrower and so on by looking at the standardized point allocations in the table. The points increase from 0 to 1,400 points with increasing performance values. The following chart is the result of applying the point allocations of the comparison table to the average performances in the dataset:

<table>
<thead>
<tr>
<th></th>
<th>100H</th>
<th>HJ</th>
<th>SP</th>
<th>200m</th>
<th>LJ</th>
<th>JT</th>
<th>800m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>16.12</td>
<td>149</td>
<td>9.44</td>
<td>27.14</td>
<td>495</td>
<td>27.30</td>
<td>155.66</td>
</tr>
<tr>
<td>Comparison Table</td>
<td>766</td>
<td>725</td>
<td>552</td>
<td>755</td>
<td>773</td>
<td>476</td>
<td>613</td>
</tr>
</tbody>
</table>

Figure 9. Comparison of Event Performance Using an Independent Source

From this chart, one sees that heptathletes are significantly more skilled in the 100H, HJ, 200m, and LJ than the other three events; therefore, the significantly higher variability in the throwing events, as shown by the high standard deviations, is because of the athletes’ relative weakness in these events. The current scoring model, then, may fairly allocate points when based solely on athlete performance. Rather than an issue with the model, the athletes competing in the event may simply perform at a
higher level in the sprinting-based disciplines than the throwing-based disciplines. Further evidence emerges in a longitudinal analysis from 2015 to 2018 of NCAA athletes competing in the heptathlon throughout those four years of their collegiate career. The percentage improvement shows that athletes tend to improve the most in the throwing events over time and very little in the running or jumping events, again causing the differences in overall variability in the dataset.

<table>
<thead>
<tr>
<th>Event</th>
<th>Average Percentage Improvement from 2015 to 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>100H</td>
<td>6.0%</td>
</tr>
<tr>
<td>HJ</td>
<td>6.0%</td>
</tr>
<tr>
<td>SP</td>
<td>12.3%</td>
</tr>
<tr>
<td>200m</td>
<td>1.8%</td>
</tr>
<tr>
<td>LJ</td>
<td>5.4%</td>
</tr>
<tr>
<td>JT</td>
<td>16.6%</td>
</tr>
<tr>
<td>800m</td>
<td>9.5%</td>
</tr>
</tbody>
</table>

Figure 10. Improvement of athletes in each event over 4-year period

Based on learning curves discovered by Hermann Ebbinghaus, humans tend to improve very rapidly at new tasks, and eventually stagnate as they reach the limits of their bodies or minds. This points to the possibility that the athletes entering the heptathlon are typically strong in the sprint-based disciplines before training for combined events. The exceptionally high javelin improvement rate could be due to the unfamiliarity with the javelin, for in many states it is the only event in the heptathlon not in high school competition.

Cumulatively, this data suggests the current IAAF model weighs the events as stated in its values. It bases the equation on statistical data from performances of each individual event. The athletes
competing in the heptathlon receive more points in the 100H and the 200m not due to imbalances in the equation, but due to their relative strengths in that discipline.

Another way to view this evidence is from a comparison to world record holders in the events. The following table was created using information from TrackStats.com (Trackstats, 2016). It shows to what percentage of execution athletes in the heptathlon perform compared to the world record. The heptathlon best performances were determined from the best event performance within the ten highest scoring heptathlons in history. The comparison between the world record, heptathlon best performances, and dataset performances is not perfect. Combined event athletes, beyond the need to train for seven different events, are also affected by other factors. For example, individual events usually start in the afternoon, meaning athletes have time to wake up, prepare their muscles, and have a nutritious breakfast and lunch to prepare for their event (Tidow, 2000). Combined events, especially in collegiate competitions, often start in the mornings between 8:00 and 11:00 (Tidow, 2000). Additionally, specialists in individual events such as the long jump, shot put, and javelin can qualify for finals in which they receive 3 extra attempts, doubling the number of attempts of heptathletes (Tidow, 2000). Despite these inconsistencies, the comparisons shed light on the performance levels between events and the standards of the current equations.
<table>
<thead>
<tr>
<th></th>
<th>100H (s)</th>
<th>HJ (m)</th>
<th>Shot (m)</th>
<th>200m (s)</th>
<th>LJ (m)</th>
<th>Jav (m)</th>
<th>800m (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>World Record</td>
<td>12.20</td>
<td>2.09</td>
<td>22.63</td>
<td>21.34</td>
<td>7.52</td>
<td>72.28</td>
<td>1:53.28</td>
</tr>
<tr>
<td>(1) Heptathlon best</td>
<td>12.54</td>
<td>1.98</td>
<td>16.44</td>
<td>22.57</td>
<td>7.27</td>
<td>59.28</td>
<td>2:07.57</td>
</tr>
<tr>
<td>(1) Percentage of World Record</td>
<td>97%</td>
<td>95%</td>
<td>73%</td>
<td>96%</td>
<td>95%</td>
<td>82%</td>
<td>89%</td>
</tr>
<tr>
<td>(2) Bests from Data Collected</td>
<td>13.00</td>
<td>1.95</td>
<td>14.99</td>
<td>23.22</td>
<td>6.40</td>
<td>50.65</td>
<td>2:06.34</td>
</tr>
<tr>
<td>(2) Percentage of World Record</td>
<td>94%</td>
<td>93%</td>
<td>66%</td>
<td>92%</td>
<td>85%</td>
<td>70%</td>
<td>90%</td>
</tr>
<tr>
<td>(3) Average from Data Collected</td>
<td>16.14</td>
<td>1.50</td>
<td>9.44</td>
<td>27.14</td>
<td>4.99</td>
<td>27.69</td>
<td>2:37.50</td>
</tr>
<tr>
<td>(3) Percentage of World Record</td>
<td>76%</td>
<td>72%</td>
<td>41%</td>
<td>78%</td>
<td>66%</td>
<td>38%</td>
<td>72%</td>
</tr>
</tbody>
</table>

Figure 11. Comparison of Scores to World Records

The evidence again highlights the heptathletes’ relatively poor performance in the throwing disciplines. In every event but the SP, JT, and 800m, the heptathlon best (1) of the event is greater than 90% of the world record, with the 800m only slightly behind at 89%. Similarly, the best performances within the NCAA dataset (2) remain over 85% in all events except the throws, and those percentages decrease at nearly double the rate. Interestingly, the 800m is actually faster in this dataset, which points to an abandonment of further training in this event. Finally, the average performance of the NCAA dataset (3) caused percentages to dip reasonably further in all events, but all remain over 65% of the world record performance except the SP and JT, which are performed at 41% and 38% of the world
An Analysis of Scoring Methods for the Women’s Heptathlon

record, respectively. Rather than a problem with the IAAF scoring model, the heptathletes simply perform at a much lower level in the throws disciplines compared to individual event athletes.

The scoring model further accentuates these disparate performances because of the progressive nature of the curve. By performing closer to the world record in the event, the athlete gets a greater incremental point increase for every improvement. Because the throwing events are performed at a lower percentage of the world record, the points earned for improvements in the throwing events are much smaller. For example, below is a graph of the 800m scoring curve created using its current equation.

![Figure 12. Scoring curve of the 800m Run](image)

This graph demonstrates how progressivity impacts point accumulation. On the far right of the curve, where the runners take the greatest number of seconds to complete the event, the absolute value of the slope increases at a shallow rate. This means that for every drop in time, the athlete achieves a slight increase in points. When examining the far left side of the graph, where the athlete completes the event in fewer seconds, the absolute value of the slope is much greater. This provides an
An Analysis of Scoring Methods for the Women’s Heptathlon

incentive to focus on improving in those events already performed at a high level, where they receive a large reward for small improvements, rather than improving in the throwing events where their improvements would be minimally rewarded.

Other causes could contribute to the imbalance in throws performance, too. As previously stated in the event descriptions, an athlete’s ideal weight and body composition differs based on the event (Hirsch et al., 2016). Because the optimal body weight for sprints and jumps is around 50 pounds less than those of throwing events, athletes in the heptathlon most likely compete in a significantly different weight class than a throwing specialist (Hirsch et al., 2016). The difference in weight could then create a difference in throwing ability. Additionally, a gender hierarchy study from the Sport, Education and Society Journal evaluates how society’s views of femininity shaped the popularity and participation in women’s track and field (Ashbolt, O’Flynn, & Wright, 2018). All female track and field athletes upset the traditional, patriarchal view of women as weak and fragile through their musculature. Some women within the sport attempt to counteract this through the use of makeup, hairstyle, and dress, which is seen particularly in women with sponsorship or in highly televised events like the sprints and jumps (Ashbolt et al., 2018). The violent nature of the throwing events embodies an even greater masculine perception, and it also requires a bulkier body type to transfer more power into the throw, both straying even further from the ideal, societal image of femininity. This creates a disincentive for a woman concerned with appearance or sensitive towards gender norms from competing in the throwing disciplines. Within the study, researchers interviewed a triple jumper and a shot putter. The triple jumper stated she typically wore crop tops to compete and was “supposed to look good,” while the thrower admitted to not bothering to brush her hair before competing at a televised meet, knowing from experience her event would not be shown (Ashbolt et al., 2018). This difference in popularity and acceptance of women’s bodies may further affect a heptathlete’s desire to train in the throwing disciplines or to sacrifice her body shape for greater strength in the throws.
Through this analysis, we determine that the current scoring method is not entirely at fault for the imbalance in scoring within the heptathlon. Instead, the athletes tend to have less experience in the throwing events when choosing to become heptathletes, and the throwing events require different body types, skills, and values than the other events. To create more balanced scoring, a scoring method should recognize these differences and provide a pathway for athletes with these strengths to succeed.

Section 4

Proposal of Solution: Multiplier Scoring Method

The previous analysis suggests the current equation is not solely at fault for the scoring imbalances in the women’s heptathlon. By further examining correlations between event performances within the NCAA 2015-2018 data, certain events are more positively correlated both to each other and to the total score. This would mean certain common skills benefit multiple events and more greatly influence an athlete’s success in the heptathlon.

<table>
<thead>
<tr>
<th>Event</th>
<th>100H</th>
<th>HJ</th>
<th>SP</th>
<th>200m</th>
<th>LJ</th>
<th>JT</th>
<th>800m</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>100H</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HJ</td>
<td>58%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SP</td>
<td>53%</td>
<td>50%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200m</td>
<td>77%</td>
<td>53%</td>
<td>53%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LJ</td>
<td>70%</td>
<td>60%</td>
<td>54%</td>
<td>74%</td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JT</td>
<td>42%</td>
<td>42%</td>
<td>63%</td>
<td>39%</td>
<td>47%</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>800m</td>
<td>49%</td>
<td>43%</td>
<td>39%</td>
<td>58%</td>
<td>52%</td>
<td>41%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>84%</td>
<td>75%</td>
<td>73%</td>
<td>84%</td>
<td>84%</td>
<td>67%</td>
<td>72%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 12. Correlations from NCAA Dataset
Most notably, the 100H, 200m, and LJ all have greater than 70% correlation to each other, leading to a correlation of 84% to the total score. All three incorporate maximum speed in a straight line, so it appears an athlete’s ability to sprint is influential to success in the heptathlon. The next highest correlations are between the two throwing events (63%) and the two jumping events (60%).

When addressing the common skills between heptathlon events, another group of researchers grouped the events in the classifications of maximum speed, explosive power, and speed endurance (Gassman et al., 2016). In their analysis, maximum speed incorporated the 100H, 200m, and LJ, while explosive power primarily included the HJ, SP, and JT. One issue they discuss is the difficulty of classifying the high jump within these categories because of it requires both speed and explosive force (Gassman et al., 2016). Speed endurance was exclusive to the 800m, for the 800m run falls under the category of middle-distance in track and field, a hybrid of a sprint and distance event.

In the following analysis, the breakdown of skills is limited to sprint speed and technical strength, but allows each event to be classified across both categories based on a percent breakdown to capture the multidisciplinary skills needed to succeed in any event. These categories, along with classifying each event within these categories, is entirely subjective. The category of sprint speed refers to the athlete’s ability to generate running speed, and technical strength refers to the athlete’s ability to control or convert speed to generate the greatest performance. Examples of technical strength include those discussed in the event descriptions: release angle and height of the javelin, the takeoff and landing phases in the long jump, starting from blocks, or the execution of the Fosbury Flop in the high jump. Because technical strength encompasses many different skills and types of control, excellence across events in this area is extremely difficult.

No endurance category will be added for the 800m. As the shortest middle-distance event, it requires the least endurance of any distance event, and its uniqueness within the heptathlon prevents a
fair, mathematical analysis of an endurance category. Additionally, the current IAAF model already recognizes the uniqueness of the 800m and claims to weight it slightly greater than the other events (International Association of Athletics Federations, 2016). Because of this, it will be classified primarily as a speed event.

The following skills breakdown is based on experience within the events, the correlations, and the resources cited in the event descriptions.

<table>
<thead>
<tr>
<th></th>
<th>100H</th>
<th>HJ</th>
<th>SP</th>
<th>200m</th>
<th>LJ</th>
<th>JT</th>
<th>800m</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sprint Speed</td>
<td>75%</td>
<td>40%</td>
<td>0%</td>
<td>95%</td>
<td>85%</td>
<td>25%</td>
<td>100%</td>
<td>60%</td>
</tr>
<tr>
<td>Technical Strength</td>
<td>25%</td>
<td>60%</td>
<td>100%</td>
<td>5%</td>
<td>15%</td>
<td>75%</td>
<td>0%</td>
<td>40%</td>
</tr>
</tbody>
</table>

Figure 13. Classification of events between speed and technical strength

According to this proposed model, a total of 60% of the heptathlon is comprised of sprinting speed and 40% technical strength. If truly attempting to find the best all-around athlete, the athletic skills should be more balanced in weight. Currently, the title of “the greatest athlete in the world” is limited to those who excel in sprint speed skill. An athlete with great skill in the throws cannot compete fairly in the heptathlon under the current model because their technical strengths are outnumbered by their weakness in sprint speed.

One solution to this problem would be to change the events in the heptathlon to promote greater diversity of talents. Suggested substitutions include exchanging the JT with the hammer throw to increase the amount of strength required (Gassman et al., 2016). The hammer throw occurs in a ring with a small diameter rather than a runway, which, like the shot put, forces the athlete to rely almost entirely on rotational technical strength rather than sprint speed. Other recommendations include replacing the 800m run with 3,000m run or 5,000m run to create a true test of endurance within the
heptathlon (Gassman et al., 2016). Including an endurance event would incorporate a currently untested athletic skill and decrease the prevalence of sprint speed, for endurance training would convert the fast-twitch muscles required for sprinting into slow-twitch muscles. Another way to remove some of the sprinting bias is the elimination of the 200m (Gassman et al., 2016). Changing these events could balance the structure to support more diverse athletes and remove the premium from sprinting events.

Changing events and structure, however, is not entirely feasible. Though the events have changed throughout history, the current structure is adopted internationally and has developed a rich history. Performances by World Record Holder Jackie Joyner-Kersee can be updated using a new scoring method, but cannot be converted to a new set of events. With such a tradition of fantastic female athletes, changing the heptathlon events would be radical, and the current structure would be fiercely defended. Additionally, the event would no longer be the heptathlon in many cases, rather a hexathlon or an octathlon.

Another way to fix the imbalance lies in disproportionally allocating points in events to create a proportionate allocation of points to skills. Rather than creating a perfectly equal distribution of points as in the Z-Score Model, a distribution can be created in which the underrepresented skill component of each event is weighted more highly using a coefficient to counteract the excess weight of speed within the current heptathlon structure.

**Multiplier Method**

To create a scoring system that equally values the talents of an athlete, it is the skills rather than the events that need to be weighted equally. Using the above percentage breakdown from Figure 13 converted to decimals, the following table demonstrates the total events allocated to the skills of sprint speed and technical strength. To mathematically balance the structure of events and skills required, the technical strength component must be multiplied by the coefficient 1.50.
### Table 1: Determination of Coefficient

<table>
<thead>
<tr>
<th>Skills</th>
<th>Amount of Events (a)</th>
<th>Coefficient (b/a)</th>
<th>Events (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sprint Speed</td>
<td>4.20</td>
<td>1.00</td>
<td>4.20</td>
</tr>
<tr>
<td>Technical Strength</td>
<td>2.80</td>
<td>1.50</td>
<td>4.20</td>
</tr>
</tbody>
</table>

This coefficient is then allocated using the percentages from Figure 13, resulting in the following coefficients for each event. This will increase the athlete’s earned points by the factor of the coefficient.

### Figure 15. Coefficient for Events

<table>
<thead>
<tr>
<th>Event</th>
<th>100H</th>
<th>HJ</th>
<th>SP</th>
<th>200m</th>
<th>LJ</th>
<th>JT</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.125</td>
<td>1.30</td>
<td>1.50</td>
<td>1.025</td>
<td>1.075</td>
<td>1.375</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The new equation would then be as follows:

\[ P(x) = D \cdot A(x - B)^C \]

\[ P(x) = D \cdot A(B - x)^C \]

where \( D \) would be the coefficient for each event, and the other variables would remain the same as the current model.

When applied to the ranked NCAA data, this change caused the top 10% of SP competitors within the heptathlon to place cumulatively 600 places higher within the 3,167 performances, which shows the increased value on technical strength. A true evaluation of this method’s efficacy, however, is nearly impossible because the dataset is based on bias from the current method. As seen in Figure 11, few outstanding throws performances exist within this dataset, so only if this method were adopted could one see if a greater number of throws specialists would enter the event and if the skill in the
throws disciplines would increase compared to specialist performances. The effects on the skill level of the other events, too, could be hypothesized, but the true effect could not be seen until implemented.

This method follows nearly all of the IAAF guidelines, for it is based entirely on the current model. One drawback is its failure to maintain similar point totals as the current model, with the highest point total within the data increasing about 1,200 points due to the $D$ coefficients. This could be fixed through reducing $D$ by about 15% in all events. The reduction would reduce the total score without interfering with the percentage allocation of sprint speed and technical strength originally addressed as it affects each event equally. The most challenging component in implementing this model, then, would be finding a universally agreed upon breakdown of skills, for both the categories listed and the percentages are based on individual judgment. Despite these drawbacks, the utilization of this method could create a more balanced system in determining the greatest all-around athlete.

Summary

Numerous methods have been proposed to create a heptathlon more suited towards distinguishing the “greatest athlete in the world.” Every attempt relies upon differing biases, and nearly all proposals blame the equation for the disparity in point distribution across events. To correct this, the Z-Score Model created a nearly ideal point distribution across events, but relied on virtually impossible population estimates and yearly updates. It also fails to recognize the unfair allocation of points to the underlying skills within the events. The Multiplier Method consists of inserting coefficients into the current model to equally weight the events based on their sprint speed and technical strength requirements. This ignores the endurance component of the 800m.

By weighting events in this manner, it allows athletes of more diverse backgrounds to find success in the heptathlon. Currently, sprints specialists dominate the competition because their strengths are used most frequently within the events and therefore rewarded most highly. By weighting
the technical strength skill component slightly higher, it allows talented throwers to compete against sprinters on a level playing field. The increased diversity will demonstrate the many ways in which a person could be the greatest athlete. A limitation in diversity continues to exist in regards to endurance athletes because the 800m run represents the only pseudo-endurance event, and there is no way to fully incorporate athletic diversity into the heptathlon without changing the event structure. Despite this drawback, the Multiplier method improves upon the current method in distinguishing the greatest athlete in the world through its recognition of diverse athletic skills.
References


An Analysis of Scoring Methods for the Women’s Heptathlon


