Toward a More Objective Understanding of Professional Wrestling: The Multidimensional Scale for the Analysis of Professional Wrestling (MSAPW)

Tyson L. Platt

Alabama State University

tplatt@alasu.edu

Aaron D. Horton

Alabama State University

ahorton@alasu.edu

The quality of professional wrestling matches has commonly been evaluated using the five-star system popularized by Dave Meltzer of the Wrestling Observer *newsletter. Though useful for describing match quality in casual settings, the scale is insufficient for academic analyses of the perception of match quality. The present article proposes a multidimensional scale for evaluating professional wrestling across the dimensions of spectacle, in-ring technique, in-ring difficulty, storytelling, and spectator engagement. Each dimension is explicitly defined and can be assessed using an 11-point scale, which allows researchers to better understand how professional wrestling viewers interpret match quality. The scale can also be weighted in a manner that allows for comparisons across different categories of professional wrestling viewers. Information on the application and analysis of the scale are discussed.*

Keywords: professional wrestling; pro wrestling; combat sports; fandom

> *"When you break down what a five-star match is, what are you looking for?"*
> *"I have no idea …"*
>
> Dave Meltzer, Wrestling Observer Radio Live Q&A
> from Starrcast II, May 24, 2019

The evaluation of professional wrestling has generally been based on the five-star system popularized by Dave Meltzer. The scale, first utilized by Norm Dooley in his *Weasel's World* professional wrestling newsletter, allows observers to rate match quality using quarter-star increments. Jim Cornette, who worked as a ringside photographer in Memphis and supplied photos to Dooley's newsletter, suggested in 1979 that Dooley adopt a star rating system akin to that used in *TV Guide* for movies in its listings. The scale was originally one to four stars until the pair witnessed a wild no-disqualification match between Jerry "The King" Lawler and Terry Funk in Memphis on March 23, 1981. Dooley, believing the match to be the best he'd ever witnessed, gave the match five stars in the next issue of his newsletter (Greene). This scale-breaking rating began the transition to the five-star system that is currently employed by many wrestling fans and journalists, which allows match quality to be assigned one of twenty-one values (excluding negative star values and DUD ratings). However, the scale is too flawed and subjective to be useful as a tool for understanding professional wrestling viewing as a form of behavior in scientific inquiry.

Any scale that requires viewers to make aesthetic judgements about performance will be inherently subjective due to factors such as preference, previous experiences and personality type contributing to judgements. In his introduction of the scale in 1985, Meltzer briefly defined the meaning of star ratings as such:

> ½* is for a terrible match, but at least there was a high spot or something. * is a bad match, *½ is below average but tolerable; ** average; **½ kind of good; *** Quite good; ***½ almost great; **** excellent; ****½ better than you can ask for.

He also indicated that a five-star rating indicates that a match is "stupendous." To the authors' knowledge, the meaning of star ratings has not been re-published or expanded upon since its introduction in the 1984 Wrestling Observer Yearbook (Meltzer, *Wrestling Observer 1984 Yearbook*). Thus, while initially defined, the meaning of star ratings has largely been lost to time. This may be due to the inconsistency with which Meltzer often applies the scale. This can be observed in Meltzer's acknowledgement that he is uncertain of what contributes to his rating of a five-star match.

This is understandable, as the scale was originally created as a shorthand means of describing one observer's experience of a match. Over time, the scale has been adopted in some form by most wrestling journalists and many wrestling fans. It is useful to the extent that it provides journalists and fans with a simple numerical value that largely describes the extent to which they enjoyed the match. However, because of the lack of stated objective criteria by which each match should be

evaluated, it is difficult to make meaningful comparisons of ratings assigned between viewers. Furthermore, there are no anchor points from which deviation can be understood. Presumably, in a five-point scale a value of 2 ½ stars should represent a match that was neither perceived to be good nor bad. However, Meltzer originally indicated that a 2 ½ star rating reflects a match being "kind of good" (*Wrestling Observer 1984 Yearbook*).

Another issue related to the five-star scale is its occasional deviation from a 5-point system. Since the introduction of the five-star system, Meltzer has assigned negative values to some matches (e.g., minus-3 stars). While introducing the scale in 1985, he indicated that negative star ratings would be assigned periodically to indicate that a match was "obnoxiously bad." The presence of negative star ratings illuminates the importance of anchor points; if the scale may take on values ranging from -5 to 5, with 0 representing "neither good nor bad", any positive value should suggest that a match is perceived as being above average. However, a review of ratings assigned in the 1-2 star range consistently demonstrates that these matches are considered below average in quality. Furthermore, Meltzer will occasionally assign a rating of DUD to a match in order to indicate that a match has "no redeeming social value" (*Wrestling Observer 1984 Yearbook*). Presumably, a negative star rating ("not only was the match worthless, but obnoxiously bad") indicates a worse evaluation than a DUD rating ("without any redeeming social value"). However, it is entirely unclear under what circumstances a particularly unappealing match will receive a low positive star rating, a negative star rating or a DUD rating.

The application of the five-star scale may less frequently result in ratings greater than 5 stars. As mentioned previously, the original Dooley four-star scale was converted to a five-star scale due to an exceptional match. Thus, even in its nascent stages, the highly subjective scale was subject to deviation from its stated parameters. Following the conversion to a five-star system, Dooley again deviated from his scale by assigning a 6-star rating to a match featuring Bill Dundee and Dream Machine vs. Wayne Ferris and Kevin Sullivan in May 1981. Meltzer first assigned a 6-star rating to a 1989 house show match between Ric Flair and Ricky Steamboat in the midst of their memorable series of televised matches that year. He again assigned a 6-star rating to a match between Mitsuharu Misawa and Toshiaki Kawada in June 1994 (Dixon). Since that time, he has assigned ratings greater than 5 stars to nearly 20 matches, including one (Kazuchika Okada vs. Kenny Omega; June 9, 2018) that earned an unprecedented 7-star rating ("Dave Meltzer 5 Star Matches List"). By deviating from his own scale, Meltzer has made it more difficult to interpret the meaning of star ratings. Meltzer has acknowledged this shortcoming and has indicated that the scale was never intended to objectively quantify the

performance of professional wrestling. Rather, it serves as a convenient means of describing his experience of the match and for communicating with others about match quality.

As a tool for describing wrestling fans' often-hyperbolic interpretation of match quality, the five-star system has been an invaluable asset for communicating issues of match quality. It has provided viewers with a means to quantify their viewing experiences for purposes of comparison with other viewers. However, the broad variability in ratings of individual matches speaks to the insufficiency of the scale as a means of understanding how individuals consume and understand the performance of professional wrestling. For instance, Meltzer assigned a rating of minus-4 stars to the main event of WWF's WrestleMania III, a long-anticipated showdown between champion Hulk Hogan and challenger Andre the Giant ("Andre The Giant"). Despite Meltzer's rating, the match is venerated by the WWF/WWE and its most loyal fans as one of the best and most historically relevant matches in history (Felstead). Though the match was undoubtedly relevant from a historical perspective, the differences in ratings of quality between Meltzer and many WWF/WWE fans must be attributable to differences in the aspects of the match to which different viewers assigned value.

Professional wrestling matches may be conceptualized as a *gestalt* event (Koffka). That is, they consist of interwoven parts that contribute to the entire experience of viewing the match. While non-fans may only consider bell-to-bell action as the substance of professional wrestling, seasoned fans recognize that the bell-to-bell action is only a part of a larger narrative. The bell-to-bell action is generally the primary criterion by which wrestling journalists and many fans assign star ratings. However, hedged into an interpretation of the action are assumptions about the characters performing the match, including their motivations to fight and their moral authority to do so. Because many of these assumptions may occur at a preconscious level, viewers may not be aware of how issues such as character and storyline development influence their match ratings.

The purpose of the present article is to propose a more objective model for evaluating professional wrestling matches. Given the shortcomings of the five-star system that is ubiquitously used, we propose a multidimensional model that allows for a more nuanced understanding of how viewers interpret pro wrestling matches. The model is intended to serve as a bridge between the present system and a more objective system that can be used for scholarly research into pro wrestling. As discussed below, some compromises were made in the development of the scale in order to balance usability with meaningful interpretation. The proposed model is theoretical in nature and has not yet been used to generate data on pro wrestling

match viewing. As such, it represents a model in its nascent stage that is intended to be modified and revised as a result of empirical observation.

Toward a More Objective Scale

The Likert method was first proposed in 1932 as a solution to the problem of quantification of phenomena that had traditionally been investigated using qualitative techniques. By representing qualitative data as quantitative data, the data can be subjected to a variety of statistical manipulations that allow for more precise interpretation and greater generalization of findings. The original Likert scale was a 5-point system ranging from "strongly disapprove" to "strongly approve." A key feature of the original Likert scale was the presence of an anchor point in the form of the response "neutral." This item represented a midpoint from which deviation can be understood (Likert). Likert-based scales have since been modified for purposes of understanding a variety of phenomena ranging from attitudes toward sex (Spector et al.) to treatment of illnesses (Gutknecht et al.) to student evaluations of classroom content (Ivanov et al.).

Because of the wide variety of phenomena studied, the labels used for numerical values on the scale vary widely (Vagias), although most formulations include a neutral anchor point (e.g., neither agree nor disagree) from which deviation can be understood. Some modifications have adopted a "forced choice" approach in which the neutral anchor point is excluded in order to require some degree of bias in responses. The present scale includes neutral anchor points for each dimension because the authors believe that (a) a neutral response is both valid and reflective of the experience of the viewer and (b) a neutral point may assist viewers in more precisely considering his/her experience of a professional wrestling match.

The present scale differs from traditional Likert scales in two meaningful ways. First, the scale requires respondents to evaluate broader characteristics than are normally rated using Likert techniques. While Likert scales generally require respondents to rate the extent to which he/she agrees with a particular statement, the proposed scale requires respondents to rate predefined dimensions of professional wrestling. Though the scale could be represented as an expansive series of statements, we believe constructing the scale in such a manner would be overly cumbersome and would dissuade use of the scale. By folding such statements into dimensions, we believe the scale is both easily accessible and useful for the interpretation of professional wrestling viewing. Second, only the values of the two extreme endpoints (1 and 10) and the neutral point (5) are explicitly defined. Though this introduces some ambiguity regarding the specific meaning of intermediate

values, we believe that it is more consistent with the manner in which viewers currently interpret professional wrestling. As the five-star system does not explicitly define the meaning of any values, professional wrestling fans are accustomed to assigning values based on gut feelings rather than objective criteria. We believe that allowing a level of subjective interpretation makes the scale more accessible to casual users and allows us to better capture the subjective aesthetic experience. As we will discuss later, the absence of defined intermediate values may also permit statistical analyses that would not be permitted with traditional Likert scales.

In order to better understand the consumption and interpretation of professional wrestling, a scale that is more objective and expansive than the five-star system will be required in order to distinguish between elements of professional wrestling matches. Such a scale should (a) be multidimensional, (b) clearly define the nature of each rated dimension, and (c) provide anchor points for determining how each dimension should be understood and assessed. We propose a Likert-derived scale based on five dimensions: spectacle, in-ring technique, in-ring difficulty, storytelling, and spectator engagement. Each dimension can be rated using an eleven-point scale in which a value of 5 represents a neutral anchor point. Any deviation away from a value of five can be interpreted as indicative of increasing or decreasing quality.

Spectacle

The dimension of spectacle is related primarily to out-of-ring aspects of professional wrestling used to create interest in the in-ring performance. This includes pre-match promos, interviews, match introductions, post-match events, match stipulations, outcomes of previous bouts between the performers and aspects of the venue where the match is held. The dimension of spectacle has been a key element of professional wrestling since its infancy. As professional wrestling is promoted primarily for the purpose of generating revenue, spectacle is the key element in match promotion. For instance, in promoting a 1911 rematch between Frank Gotch and George Hackenschmidt, Hackenschmidt accused Gotch of relying on "dirty tactics" in their 1908 match and claimed to have suffered an illness that left him unable to compete at a high level (Trowbridge). The combination of the accusation of cheating and claims of illness were used to promote the need for a rematch. The rematch at Comiskey Park in Chicago drew 30,000 spectators, an increase of nearly 24,000 compared to their 1908 match ("George Hackenschmidt Bio").

The value of spectacle in the interpretation of professional wrestling matches is apparent in reviews. For instance, a match between commentators Michael Cole

and Jerry Lawler at WrestleMania 27, built entirely around spectacle and interference, was the culmination of a lengthy storyline in which Cole had repeatedly insulted Lawler, including making light of the latter's mother's recent passing. Most fans expected that the match would be short and feature Lawler quickly defeating Cole as comeuppance for the latter's consistent harassment over the months leading to the match. Instead, the entire spectacle lasted nearly fifteen minutes, with interference from guest referee Steve Austin. Scott Keith emphasized the match's excessive length, giving it a minus-3 star rating ("Wrestlemania Countdown 2018") while in his one-star rating Meltzer seemed to find parts of it entertaining, especially a good-looking dropkick from the 61-year-old Lawler, though he also noted that the match lasted too long given its performers' limitations (*Wrestling Observer Newsletter*, April 11, 2011). This discrepancy highlights the inconsistency of the scale, as reviewers occasionally employ negative star ratings to criticize a match they find particularly awful; though Keith saw nothing of value in the match, Meltzer found parts of it entertaining enough not to resort to a negative star rating.

Only promos and interviews that are directly related to the match should be considered when evaluating this dimension. Though promos and interviews may serve as an important element of storytelling, they should not be considered part of the storytelling dimension as that dimension applies only to bell-to-bell elements of the match. The scale ranges from Extremely Poor Spectacle (0) to Exceptional Spectacle (10) with a value of 5 indicating Neutral Spectacle (neither good nor bad). Generally, matches that have been built around clearly defined characters and conflicts and in which the consequences of the outcome of the match are clear to viewers will be scored high on the spectacle dimension. Matches in which there is no discernable reason for the match and/or the consequences of the outcomes of the match are unclear should be scored low on the Spectacle dimension.

0	5	10
Extremely Poor Spectacle	Neutral Spectacle (neither good nor bad)	Exceptional Spectacle
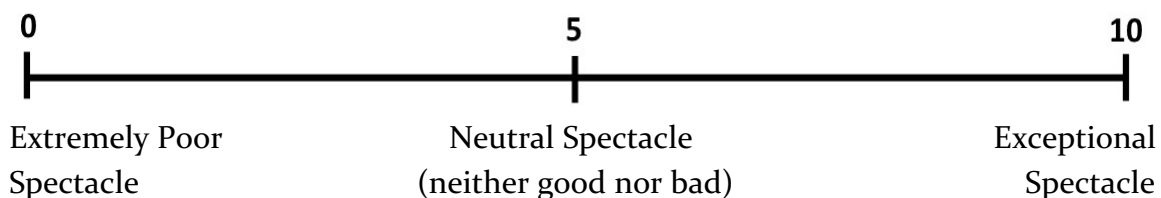
In-Ring Technique

The dimension of in-ring technique is related to the execution of in-ring moves independent of the difficulty of the moves being performed. Instead, it is related to how well moves, holds and strikes are executed. Because professional wrestling matches are generally designed to suggest that a legitimate fight is

occurring, in-ring technique should be evaluated based on crispness of execution and realism. In-ring technique is a key dimension of many current reviews. For instance, in his review of the match between Chris Jericho and Kenny Omega at AEW Double or Nothing (May 25, 2019), T.J. Hawke emphasized the bout's botched reversal spot, in which Jericho tried to escape Omega's One-Winged Angel finisher and ended up dropping the latter on his head, and gave the match 2 ½ stars. In contrast, both Meltzer (*Wrestling Observer Newsletter*, June 3, 2019) and Keith ("The SmarK Rant for AEW Double or Nothing") praised the match not only for its in-ring action but for its storytelling, with each awarding the match 4 ¼ stars, a very high score. Hawke clearly valued the perceived lack of in-ring execution in the match while Meltzer and Keith focused on its overall storytelling value.

As the performance of moves often requires the coordinated execution of sequences by both protagonist and antagonist, this dimension should be used to evaluate quality of execution as a function of all parties involved in the match. The scale ranges from Extremely Poor Technique (0) to Exceptional Technique (10) with a value of 5 indicating Neutral Technique (neither good nor bad). Matches that contain interactions between performers that suggest realism and that contain consistently crisp execution of moves, strikes and holds will be scored higher on the in-ring technique dimension. Matches that contain disjointed interactions between performers, lack realism and/or contain inconsistent execution of moves, strikes and holds should be scored lower on the in-ring technique dimension.

0	5	10
Extremely Poor In-Ring Technique	Neutral In-Ring Technique (neither good nor bad)	Exceptional In-Ring Technique
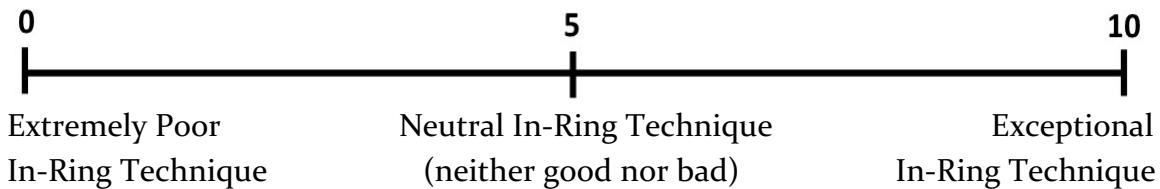
In-Ring Difficulty

The dimension of in-ring difficulty is related to the complexity and difficulty level of the moves performed in a match irrespective of the quality of execution of the moves. Due to the variability in the moves employed by wrestlers over time, in-ring difficulty must be assessed relative to the moves that are common at the time the match was held. Thus, in-ring difficulty may be considered a rating of a performer's move set relative to the moves most commonly observed during that performer's era. As the moves performed are the fundamental ingredients of any professional wrestling match, in-ring difficulty is often addressed in match reviews. For example, in his review of the Kazuchika Okada-Tetsuya Naito IWGP Heavyweight Championship match at WrestleKingdom 12 on January 4, 2018,

Meltzer praised the match's technical qualities, with both wrestlers using a series of difficult and impressive moves and sequences throughout, giving the match a near-perfect 4 ½ star rating (*Wrestling Observer Newsletter*, January 15, 2018). Kevin Pantoja of 411Mania was, however, less impressed, and seemingly prized storytelling over in-ring difficulty, as Meltzer had. Pantoja was disappointed that Naito lost the match, and felt the match was "missing something," giving the match 3 ¾ stars ("Kevin's NJPW Wrestle Kingdom 12 Review"). In this instance, Meltzer clearly valued the high level of difficulty and technical execution in the match and did not seem to factor the match result into his overall evaluation of it, while the fact that Naito lost clearly affected Pantoja's overall score, indicating that storytelling was a more influential factor.

The dimension of in-ring difficulty is primarily related to the degree of athletic skill and timing required to execute a move set within a match. As the performance of moves often requires the coordinated execution of sequences by both protagonist and antagonist, this dimension should be used to evaluate in-ring difficulty as a function of all parties involved in the match. The entire repertoire of a wrestler should not be considered when evaluating this dimension. Instead, only the moves performed in the match under evaluation should be considered. Across time, this dimension should be considered the most variable of all dimensions. As professional wrestling has developed over time, the level of difficulty in the execution of moves has increased greatly. For instance, in the 1980s the moonsault (top-rope back flip splash) was considered a very difficult highspot. Today, such a move is performed routinely by a wide variety of performers. For this reason, in-ring difficulty is difficult to compare over time. The scale ranges from Extremely Low Difficulty (0) to Extremely High Difficulty (10) with a value of 5 indicating Moderate Difficulty (neither low nor high). Creative matches that involve a high level of athletic skill and timing will be scored above a value of 5, whereas matches that rely on less athletic skill and timing will be scored below a value of 5.

0	5	10
Extremely Low Difficulty	Moderate Difficulty (neither low nor high)	Extremely High Difficulty
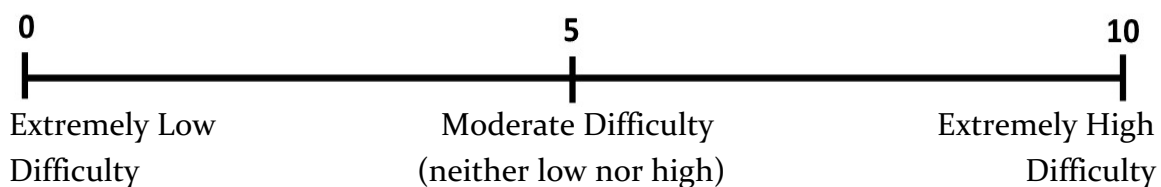
Storytelling

A key aspect of professional wrestling matches is the story that is conveyed in the ring. Attention to the narrative of professional wrestling matches is often noted in reviews. For instance, Keith panned the Miz-Shane McMahon falls count

anywhere match at WrestleMania 35 ("The Smark Rant for WWE Wrestlemania 35"), claiming it was "just two guys walking around the arena" and that it was generally terrible, awarding it a 1-star rating. Similarly, Meltzer emphasized the poor execution of moves, particularly on the part of McMahon. Though he did acknowledge that the highspots "got over" with the live crowd, Meltzer gave the match 1 ¾ stars (*Wrestling Observer Newsletter*, April 15, 2019). However, Pantoja of 411Mania gave the match 4 ¼ stars, calling it an "incredible match" for its highspots and storytelling elements, including the involvement of Miz's father ("Kevin's WWE WrestleMania Review"). In this instance, Keith clearly valued the match's workrate and in-ring action over storytelling, while Pantoja seemingly prized the latter element above all else.

The dimension of storytelling is related to how well a match conveys its intended narrative. The dimension is associated with the order and manner in which moves and breaks between moves are executed in order to tell a story. It also involves non-combative aspects of matches such as facial expressions, expressions of physical and emotional turmoil and taunting. In some cases, storytelling may be aided by non-wrestlers such as managers and valets. Additionally, storytelling may be aided or diminished by play-by-play/color commentary during the match. In many cases, match stipulations may contribute to storytelling, although in some cases match stipulations may be used solely as contributors to the spectacle dimension of the scale. The viewer should only consider match stipulations as an element of storytelling if they are exploited to convey the match narrative. The narrative of matches can be wide-ranging and may include conflicts between babyfaces (virtuous) and heels (villainous), competition to determine which wrestler is more skilled, and matches intended to establish the skills and personal characteristics of wrestlers (e.g., "squash matches"). The scale ranges from Extremely Poor Storytelling (0) to Exceptional Storytelling (10) with a value of 5 reflecting Average Storytelling (neither good Nor bad). Matches in which a clear and cohesive narrative structure can be understood by viewers should be scored a value greater than 5, whereas matches lacking in a clear narrative structure or inconsistent in their conveyance of a narrative should be scored a value lower than 5.

0	5	10
Extremely Poor Storytelling	Neutral Storytelling (neither good nor bad)	Exceptional Storytelling
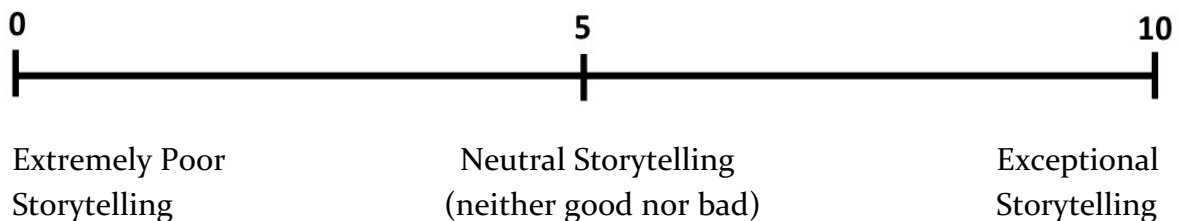
Spectator Engagement

The dimension of spectator engagement is related to the extent to which the live viewing audience appears to be actively engaged in the viewing experience. It should be viewed as a measure of the excitement that is produced by the live viewing experience. The degree of spectator engagement is frequently noted in match reviews. For example, in his review of the Triple H-Batista match at WrestleMania 35 ("The Smark Rant for WWE Wrestlemania"), Keith not only criticized the match's slow pace but also the crowd's seeming lack of interest. He gave the match half a star, a very low score. In contrast, Steve Cook of 411Mania awarded the match 3 ½ stars, calling it an "impressive showing" despite perhaps running too long. Cook indicated his enjoyment of the match and did not make any mention of the disengaged audience. Pantoja of 411Mania went even higher, calling it an "absolute spectacle of a match," and rated it 4 stars by virtue of its highspots, making no mention of the fans' interest or lack thereof ("Kevin's WWE WrestleMania Review").

The variability in ratings may be attributed to the fact that this dimension is difficult to assess outside of live settings. That is, because cameras generally focus on the wrestling ring, it may be difficult for viewers of recorded matches to assess this dimension. When this is the case, the dimension should be excluded from analysis.

It is relevant to note that there are cultural differences in the expression of excitement. Thus, though the loudness of crowd responses may be considered when evaluating this dimension, it should not be the sole factor considered when evaluating spectator engagement. Other factors such as proportion of the audience who appears to be actively viewing the match and the number of viewers who are involved in non-viewing activities (e.g., talking to neighbors, using phone or other device, playing with beach balls) should also be considered. Similarly, though crowd size may be considered when evaluating this dimension, it should not be the primary factor considered. The level of active engagement of the spectators present should be considered a more important criterion than the total number of spectators viewing the event. The scale ranges from Extremely Low Engagement (0) to Extremely High Engagement (10) with a value of 5 reflecting Moderate Engagement (neither low nor high).

0	5	10
Extremely Low Engagement	Moderate Engagement (neither low nor high)	Extremely High Engagement
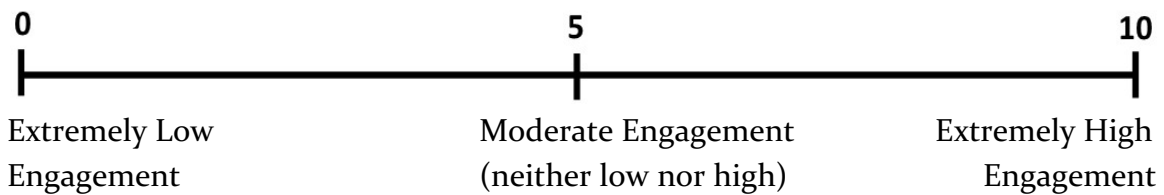
Calculation of Match Ratings

The proposed scale is highly flexible due to its ability to account for viewers' expectations in assigning match ratings. In its most basic form, the scale assumes that viewers equally value each of the five dimensions of the scale. Thus, the standard application of the scale can be stated as:

(Spectacle + In-Ring Technique + In-Ring Difficulty + Storytelling + Spectator Engagement) / 5 = Rating

For instance, a match may be rated by a viewer as follows:

Spectacle – 6
In-Ring Technique – 9
In-Ring Difficulty – 7
Storytelling – 8
Spectator Engagement – 4

The rating for the match would be calculated as:

(6 + 9 + 7 + 8 + 4) / 5 = **6.8**

In this application, each dimension contributes 20% to the overall match rating. While the scale allows for assignment of a single value that describes the overall perceived quality of a match, it preserves the elements that contributed to the rating. This allows for a more accurate understanding of how matches are interpreted by viewers. For example, two viewers may assign an overall rating of a match a score of 5. Using a unidimensional scale, observers may conclude that the two viewers interpreted the match identically. However, using a multidimensional scale, observers can better understand differences in how the two viewers interpreted the match. Consider the following ratings for two viewers:

Dimension	Viewer #1	Viewer #2
Spectacle	6	6
In-Ring Technique	3	3
In-Ring Difficulty	1	8
Storytelling	8	1
Spectator Engagement	7	7
MATCH RATING	**5**	**5**

In this scenario, identical match ratings can be observed to be a result of different interpretations of the match. Both viewers interpreted the spectacle, in-ring technique, and in-ring difficulty dimensions identically. However, they profoundly differed in their interpretation of in-ring difficulty and storytelling dimensions. By relying on a match rating that is derived as an average of the five dimensions, observers can reach a general understanding of an observer's overall impression of a match. But by using a multidimensional scale, observers may disambiguate identical ratings, thus better understanding how professional wrestling is interpreted by different viewers.

Furthermore, the multidimensional scale allows for viewers' previous assumptions about the nature of professional wrestling to be accounted for prior to match evaluation. This can be accomplished by differentially weighting scale dimensions. For example, consider the presentation of professional wrestling by WWE. WWE identifies itself as "sports entertainment" rather than "professional wrestling." In doing so, it has influenced its fanbase's interpretation of professional wrestling by emphasizing the importance of the spectacle and storytelling dimensions of professional wrestling at the expense of the In-Ring Difficulty and In-ring technique dimensions. Thus, professional wrestling fans who view WWE exclusively may evaluate matches with a bias toward the spectacle and storytelling dimensions. In order to better understand how WWE fans interpret match quality relative to their expectations, a modification of the standard scale may be required. This can be accomplished by differentially weighting each dimension to reflect viewer expectations. In a restated form, the standard scale can be represented as:

(Spectacle x 0.2) + (In-Ring Technique x 0.2) + (In-Ring Difficulty x 0.2) + (Storytelling x 0.2) + (Spectator Engagement x 0.2) = Rating

Because WWE fans may value spectacle and storytelling, the scale might be reweighted to better reflect the relationship between the viewers' expectations and their experiences. Such a reweighting might be stated as:

(Spectacle x 0.275) + (In-Ring Technique x 0.15) + (In-Ring Difficulty x 0.15) + (Storytelling x 0.275) + (Spectator Engagement x 0.15) = Rating

In this case, in-ring difficulty, in-ring technique and spectator engagement are assumed to be 45.5% less valuable in the interpretation of a match than the spectacle and storytelling dimensions. If we apply this modified model to the ratings assigned by Viewer #1 in the previous example, we would obtain the following rating:

(6 x 0.275) + (3 x 0.15) + (1 x 0.15) + (8 x 0.275) + (7 x 0.15) = 5.5

Whereas the rating for Viewer #2 would be expressed as:

$$(6 \times 0.275) + (3 \times 0.15) + (8 \times 0.15) + (1 \times 0.275) + (7 \times 0.15) = 4.625$$

Thus, by reweighting the scale in a manner that reflects the assumptions of the viewers, the identical ratings obtained using the standard scale are no longer identical. Again, such weighting allows for a better understanding of the relationship between viewers' expectations and their experiences.

The valid weighting of individual dimensions can only be determined empirically. For instance, surveying WWE fans on the relative importance of each dimension in their expectation of matches would allow for the determination of the appropriate weights for each dimension. Those weights could then be applied when interpreting how self-identified WWE fans experience both WWE and non-WWE matches. Differential weighting is acceptable under the following conditions:

1) The reviewer or researcher indicates explicitly how dimensions are weighted. This can be accomplished by placing the weights in italicized parentheses: *(.05, .3, .3, .3, .05)*
2) The sum of the weights is equal to 1 (i.e., .05 + .3 + .3 + .3 + .05 = 1)

Statistical Analysis

There is considerable debate regarding the nature of Likert-derived data, with the more conservative argument being that Likert-derived data represents ordinal rather than interval data. This is an important consideration as it dictates the statistical analyses that are permissible during data analysis (Jamieson). The issue is ultimately related to whether or not the distances between values are equivalent. Because traditional Likert scales qualitatively define the meaning of each value, a conservative interpretation suggests that we should not assume equivalent distances between values. For instance, we may not assume that the distance between "neutral" and "agree" is perceived as being equivalent to the distance between "agree" and "strongly agree" by respondents. If we adopt this assumption, interval-based analyses such as mean calculation cannot be employed. Instead, median values are considered an acceptable measurement of central tendency. As a result, tests such as Kruskal-Wallis test and the Mann-Whitney U test are sometimes employed to compare medians.

Because the proposed scale does not explicitly define the meaning of intermediate values, we believe that the assumption of equivalence between values is more justifiable. For instance, users should interpret the distance between a value of 1 and 2 as equivalent to the distance between 5 and 6 on the Storytelling dimension by users. For this reason, the scale can be considered to produce interval

data. This assumption allows for the calculation of a mean as an appropriate measure of central tendency and permits the use of a t-test or ANOVA for the comparison of means across individuals.

Conclusions

Professional wrestling has often been overlooked or simplistically reduced by academics due partially to its neither-fish-nor-fowl status: it is often disparaged as being less-than-sport by sports enthusiasts and dismissed as lowbrow by entertainment experts. However, professional wrestling represents a unique and enduring worldwide phenomenon that deserves more careful consideration and analysis than it has generally received. One missing element in the careful analysis of professional wrestling is the proper tool for understanding the consumption of professional wrestling.

The issue of quantifying the aesthetic experience of viewing and interpreting professional wrestling is difficult but necessary for a better, more nuanced understanding of the phenomenon. For nearly 40 years, the five-star system has served as the gold standard for quantifying the experience of professional wrestling viewership. The scale has been incredibly useful for purposes of describing the extent to which viewers enjoy professional wrestling matches, but it fails to disambiguate the reasons for the enjoyment. The present article represents the beginning of the construction of a more useful model for evaluating professional wrestling. As such, it will be modified and refined as a function of empirical observation.

The proposed Multidimensional Scale for the Analysis of Professional Wrestling (MSAPW) represents an attempt to introduce a better tool for the quantitative analysis of professional wrestling viewership. By requiring viewers to rate professional wrestling matches across five dimensions (spectacle, in-ring technique, in-ring difficulty, storytelling, and spectator engagement), we believe that researchers will be better equipped to understand the pro wrestling viewing experience. Though more expansive than the commonly employed five-star system, we believe the proposed scale is sufficiently simple to allow for application in both casual and academic settings. Professional wrestling fans are often enthusiastic and dedicated, and we believe that they will welcome a more sophisticated means of describing their experience of pro wrestling. Similarly, academics may use the scale to more thoroughly investigate the phenomenon of professional wrestling. To the extent that a lack of useful quantitative tools prohibits insightful quantitative research into pro wrestling, the proposed scale may also help to stimulate interest and meaningful research in the field.

The five dimensions of the scale were derived from extensive reviews of match evaluations by a variety of sources. Aspects of spectacle, in-ring technique, in-ring difficulty, storytelling, and spectator engagement were both directly and indirectly addressed in reviews from a variety of reputable reviewers. Because of their frequent appearance in reviews, the dimensions were selected as key elements of professional wrestling match analysis that could be disambiguated from each other. However, the authors recognize that data collected on use of the scale may result in an expansion, reduction or reinterpretation of each of the dimensions.

One goal in developing the proposed model was to find a balance between usability and objective interpretation of matches. We were sensitive to the fact that an overly complex scale may be perceived as overly cumbersome and, as a result, would not be commonly used, and an unused scale is of little value in understanding match interpretation. The purpose of the proposed scale is to serve as a bridge between the tool that is currently used (the five-star system) and a more formal and objective scale (MSAPW). As such, some compromises had to occur to produce a scale that is both usable and more objective. One such compromise was the decision to allow users to rate whole dimensions rather than creating sub-items for each dimension. Though we considered representing each dimension as a set of up to five ratable items, we believe that the requirement to rate up to twenty-five individual items would be perceived as burdensome by reviewers and would ultimately dissuade use of the scale. However, future empirical observations may necessitate the development of sub-scale items. This may prove especially true if researchers attempt to investigate the assessment of individual dimensions of the scale. The authors intend to apply and refine the scale based on data acquired from a wide variety of reviewers who use the MSAPW, and we strongly encourage other researchers to use and refine the scale in order to create a more meaningful empirical knowledge base about professional wrestling viewership. The authors claim no ownership of the scale, and it may be used freely for academic or personal use.

Works Cited

"Andre The Giant." *Internet Wrestling Database (IWD)*, www.profightdb.com/wrestler-star-ratings/andre-the-giant-15.html

Cook, Steve. "Cook's WrestleMania 35 Review." *411 Mania*, 8 Apr. 2019, 411mania.com/wrestling/cook-wrestlemania-35-review/

"Dave Meltzer 5 Star Matches List (1983 to 2020)." *IWNerd.com*, https://www.iwnerd.com/dave-meltzer-5-star-matches-list/

Dixon, James. "First Six-Star Match Revealed And It Is Not What You Would Think!" *WrestleTalk*, 28 Oct. 2018, https://wrestletalk.com/news/first-six-star-match-revealed-and-it-is-not-what-you-would-think/

Felstead, Scott. "The 25 Greatest Matches in WWE History." *Muscle & Fitness*, Weider Publications, 2020, https://www.muscleandfitness.com/athletes-celebrities/news/25-greatest-matches-wwe-history

"George Hackenschmidt Bio." *Professional Wrestling Historical Society*, https://www.prowrestlinghistoricalsociety.com/w-hackenschmidt.html

Greene, Dan. "The History of Wrestling Match Star Ratings." *Sports Illustrated*, 15 Aug. 2018, https://www.si.com/wrestling/2018/08/15/dave-meltzer-observer-star-ratings-wwe-njpw

Gutknecht, Mandy, et al. "Assessing the Importance of Treatment Goals in Patients with Psoriasis: Analytic Hierarchy Process vs. Likert Scales." *The Patient-Patient-Centered Outcomes Research*, vol. 11, no. 4, 2018, pp. 425-37.

Hawke, T.J. "Hawke's AEW Double or Nothing Review." *411Mania*, 26 May 2019, https://411mania.com/wrestling/hawkes-aew-double-or-nothing-review/

Ivanov, O. A., V. V. Ivanova, and A. A. Saltan. "Likert-Scale Questionnaires as an Educational Tool in Teaching Discrete Mathematics." *International Journal of Mathematical Education in Science and Technology*, vol. 49, no. 7, 2018, pp. 1110-18.

Jamieson, Susan. "Likert Scales: How to (Ab)use Them." *Medical Education*, vol. 38, no. 12, 2004, pp. 1217-18.

Keith, Scott. "The SmarK Rant for AEW Double or Nothing–05.25.19." *Scotts Blog of Doom!*, 28 May 2019, http://blogofdoom.com/index.php/2019/05/27/the-smark-rant-for-aew-double-or-nothing-05-25-19/

---. "The SmarK Rant for WWE Wrestlemania 35–04.07.19." *Scotts Blog of Doom!*, 9 Apr. 2019, http://blogofdoom.com/index.php/2019/04/07/the-smark-rant-for-wwe-wrestlemania-35-04-07-19/

---. "Wrestlemania Countdown 2018: The SmarK Rant for WWE Wrestlemania 27." *Scotts Blog of Doom!*, 9 Mar. 2018, http://blogofdoom.com/index.php/2018/03/09/wrestlemania-countdown-2018-the-smark-rant-for-wwe-wrestlemania-27/

Koffka, K. *Principles of Gestalt Psychology*. Brace, 1935.

Likert, Rensis. "A Technique for the Measurement of Attitudes." *Archives of Psychology*, vol. 140, pp. 1-55.

Meltzer, Dave. *Wrestling Observer 1984 Yearbook,* January 1985.

---. *Wrestling Observer Newsletter*, April 11, 2011.

---. *Wrestling Observer Newsletter*, April 15, 2019.

---. *Wrestling Observer Newsletter*, January 15, 2018.

---. *Wrestling Observer Newsletter*, June 3, 2019.

Pantoja, Kevin. "Kevin's NJPW Wrestle Kingdom 12 Review." *411Mania*, 5 Jan. 2018, https://411mania.com/wrestling/kevins-njpw-wrestle-kingdom-12-review/

---. "Kevin's WWE WrestleMania Review." *411Mania*, 8 Apr. 2019, https://411mania.com/wrestling/kevin-wwe-wrestlemania-review/

Spector, Ilana P., Michael P. Carey, and Lynne Steinberg. "The Sexual Desire Inventory: Development, Factor Structure, and Evidence of Reliability." *Journal of Sex & Marital Therapy*, vol. 22, no. 3, 1996, pp. 175-90.

Trowbridge, Tony. "When Gotch Beat the 'Russian Lion' - Again." *Chicago Tribune*, 8 Sept. 2018, https://www.chicagotribune.com/news/ct-xpm-2013-09-08-ct-per-flash-wrestling-0908-20130908-story.html

Vagias, Wade. M. "Likert-Type Scale Response Anchors." *Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management*. Clemson University, 2006.