2009 MFTHBA Celebration Summary of Judge Performance

# 2009 MFTHBA Celebration Summary of Judge Performance

The performance of the 2009 Celebration judges was typical of what has been seen over the last dozen years. There were a few classes where they agreed on placings, several with moderate disagreement and a few where there was almost no agreement. Before reviewing these results it would be very helpful for you to be familiar with the basics of what is being measured and how the statistics reported herein should be interpreted. If you have not already done so, I recommend you read the information in the Judge Performance Tutorial. There are more examples given in the MFT2003 and MFT2004 summaries.

The preferred classes to use in evaluating judge performance are the open classes because it is within them we have the best horses and best exhibitors in general and that translates to the entries being quite consistent in their performance as they circle the arena which gives all the judges a very similar view of each exhibitor and horse. Only classes with at least 10 entries are eligible for judge evaluation because placing smaller classes is not nearly as challenging as placing classes with 10 or more entries and we want to see how judges perform when challenged. There was a problem finding enough open classes that qualified for use in 2009 judge evaluation because there were very few open classes with 10 or more entries in the 2009 Celebration.  As a result there are only 5 open classes included in the 2009 judge evaluation summary. Five other classes were included which were open amateur and amateur WGC classes in order to get enough data to do a reasonable evaluation of the 2009 judges' performances. The use of amateur classes in all likelihood caused more disagreement among the judges' placings because of less consistency in the exhibitors' performances. The extra variation in exhibitor performance doesn't affect the comparisons of the 2009 judges to one another as we are still able to evaluate which were more consistent with their peer judges even though they all were evaluating exhibitors who may not have been as consistent in their performances as we would have liked in some classes. However, the fact that there were five amateur classes in the 2009 results should be kept in mind when comparing the 2009 summary results to other years that didn't include any amateur classes in their summary data. We can assume the 2009 judges would have been somewhat more consistent than they were over the 10 classes chosen here had I been able to base their evaluation on 10 open classes.

The judging tutorial explains how the distance measure we use to quantify how different a judge is from his peers when placing a class. It shows that if a judge agrees completely with the consensus of his peers, his distance measure will be zero. That is the goal we seek but seldom occurs. In the discussion below we discuss the distance the judges are from the consensus of their peers for the first 3 places because we feel they are the most important places as far as exhibitors are concerned. If a judge switches an adjacent pair in the top 3 positions such as places horses 2, 1, 3 rather than 1, 2, 3, his distance for the class would be 1.41. If he picks the top 3 horses that were picked by his peer judges but places them in reverse order, i.e. places them 3, 2, 1 rather than 1, 2, 3, his distance will be 2.83. Ideally we would like to see judges have distances that are in the range of zero to 3 because that range reflects that there is class agreement among a judge and the consensus placings of his peers. When distance values get into the 4 and up range it indicates a judge didn't agree very well with his peers in placing a class. We also use *theta* as a generalized measure of agreement. It may be easier to interpret when comparing the amount of disagreement between judges because its range is 0 to 1. 0 is perfect agreement and 1 is no agreement at all. *Theta* less than 0.3 correlates to a short 1-3 distance and indicates good agreement among judges in the first 3 positions. Values near 0.5 and larger indicate the judges disagreed considerably on how the horses should have been placed.

Although I concentrate exclusively on the first 3 places which is referred to as the 1-3 distance, note that the 1-5 distance and 1-10 distance are also reported for each class. These are measures of the distance between a judge and his peer consensus ranking for the first 5 and all 10 places, respectively. They may be of interest at times even though I do not mention them in my discussion of results.

# Class #178 - Open 3 yr old WGC - 2009

```
 5 judges
10 horses ranked

    Rankings are:                Orderings based on all 5 judges:

  b    m    t    r    s          b    m    t    r    s
 115  115   97  115  115         1    1    3    1    1
  42   42   42   97   98         2    2    2    3    8
 347  392  236   42   42         5    6    4    2    2
  97   98  744  347  744         3    8    7    5    7
 392   97  115  236  236         6    3    1    4    4
 236  236  347   98   97         4    4    5    8    3
 744  744  392  392  347         7    7    6    6    5
 177  347  177  744  392         9    5    9    7    6
  98  177   98  177  177         8    9    8    9    9
 974  974  974  974  974        10   10   10   10   10

            Orderings from deleting one judge at a time:

                  b    m    t    r    s
                  1    1    3    1    1
                  2    2    2    3    8
                  8    8    6    2    2
                  3    7    8    7    7
                  7    3    1    4    4
                  4    4    5    8    3
                  6    6    7    6    5
                  9    5    9    6    6
                  5    9    4    9    9
                 10   10   10   10   10
                                               Averages
  1-3 Distance  5.000  5.000  3.606  1.414  6.083      4.221
  1-5 Distance  5.477  6.164  6.708  3.464  6.856      5.734
 1-10 Distance  7.211  7.211  8.485  5.099  8.000      7.201

Theta hat for class: 0.425
```

There was considerable disagreement in this class. Both theta and the average 1-3 distance are larger than desirable and judges b, m, and s all disagreed strongly with their peers in how the class should  be placed. A WGC class should be placed with considerably more agreement than seen here.

## Class #183 - Open 5 yr old WGC - 2009

```
 5 judges
10 horses ranked

     Rankings are:                Orderings based on all 5 judges:

  m    t    r    s    g           m    t    r    s    g
 152  745  745  739  152          4    1    1    2    4
 739  682  739  745  745          2    3    2    1    1
 470  739  682  682  674          7    2    3    3    6
 682  100  674  470  100          3    5    6    7    5
 745  470  746  674  739          1    7    8    6    2
 439  152  152  152  682         11    4    4    4    3
 100  674  100  746  746          5    6    5    8    8
 746  746  734  100   76          8    8   10    5   12
 762  391  391  734  391         13    9    9   10    9
 277   76  470  391  734         14   12    7    9   10


            Orderings from deleting one judge at a time:

                  m      t      r      s      g
                  4      2      1      2      4
                  2      4      2      1      2
                  8      1      4      4      7
                  3      6      7      8      6
                  1      8      8      6      1
                 14      3      3      3      3
                  6      5      6      7      8
                  7      7     12      5     14
                 13     11     10     11     10
                 12     12      5      9      9
                                                  Averages
1-3 Distance    5.831  3.000  1.000  1.732  5.000    3.313
1-5 Distance    7.141  4.690  4.359  4.472  6.708    5.474
1-10 Distance  11.705  6.633  8.426  6.557  9.644    8.593

Theta hat for class: 0.365
```

Note the 1-3 distance and theta both show the disagreement is not great but that there was some which came primarily from judges m and g not being in agreement with the other 3 judges. This class verifies the importance of using Olympic style judging which eliminates the extreme score at either end of the rankings for each horse. Note judge r was in close agreement with the consensus of his peers except for their 3<sup>rd</sup> place choice.

This is a very poor performance as far as being the class where the breed champion was selected. These judges could not have all been going by the same breed standard in choosing their top placings.

Note that in the 2003 TWH WGC class the average 1-3 distance was 1.4 and theta was 0.16. Comparing these values to those above pretty well summarizes the judge situation in the MFTHBA.

## *Class #132 - Open 5 Yr & Older Mares - 2009*

```
 5 judges
10 horses ranked

    Rankings are:                 Orderings based on all 5 judges:

  r    g    s    m    b          r    g    s    m    b
  12  739  739  739  682         3    1    1    1    2
 739  391  470  470   12         1    5    4    4    3
 391  682  682  698  739         5    2    2   10    1
 682  417   12  682  417         2    7    3    2    7
 417   51   51   51  470         7    6    6    6    4
  51   12  153   12  391         6    3    8    3    5
 470  153  391  153   51         4    8    5    8    6
 756  756  417  391  756         9    9    7    5    9
 698  470  756  756  153        10    4    9    9    8
 153  698  698  417  698         8   10   10    7   10

        Orderings from deleting one judge at a time:

                 r      g      s      m      b
                 3      1      1      1      2
                 1      6      7      7      3
                 6      2      2     10      1
                 2      7      3      2      7
                 7      5      6      6      5
                 5      3     10      3      4
                 4      9      4      8      6
                10     10      5      4     10
                 9      4      9      9      8
                 8      8      8      5      9
                                                    Averages
 1-3  Distance  3.742  4.123  5.099  8.602  2.449      4.803
 1-5  Distance  4.690  5.099  5.292  8.888  3.873      5.568
 1-10 Distance  6.325  8.485  8.124 11.402  5.099      7.887

Theta hat for class: 0.499
```

Note the magnitude of the average 1-3 distance and theta indicating a lot of disagreement in this class. Judge m picked his peer's 10th place horse as his 3rd place horse and their 7th place horse as his 2nd place. Judge b did really well relative to his peers.

## Class #134 - Open 5 Yr & Older S&G - 2009

```
 5 judges
10 horses ranked
```

```
     Rankings are:                  Orderings based on all 5 judges:

  b     t     r     g     s            b     t     r     g     s
 745   674   745   745   745           1     2     1     1     1
 152   745   100   674   152           4     1     3     2     4
 100   100   674   152   674           3     3     2     4     2
 674    76   746   100   695           2     6     8     3     5
 695   152   695   734   100           5     4     5     7     3
 734   695    76   695   277           7     5     6     5     9
 746   734   152   746    76           8     7     4     8     6
 277   277   734    76   746           9     9     7     6     8
  76   746   277   277   734           6     8     9     9     7
 764   764   764   764   764          10    10    10    10    10
```

```
           Orderings from deleting one judge at a time:

                      b     t     r     g     s
                      1     2     1     1     1
                      4     1     4     2     4
                      3     4     2     4     2
                      2     8     8     3     5
                      5     3     5     8     3
                      8     5     7     5     9
                      7     7     3     7     8
                      9     9     6     6     7
                      6     6     9     9     6
                     10    10    10    10    10
```

```
                                                        Averages
 1-3 Distance   2.000  1.732  2.236  1.000  2.236        1.841
 1-5 Distance   2.828  4.796  4.583  3.317  3.162        3.737
 1-10 Distance  4.690  5.831  6.481  4.000  5.477        5.296
```

Theta hat for class: 0.268

This class was well judged with only minor disagreements in the first 3 placings. Note that all the 1-3 distances are small and so is theta.

## Class #155 - Am 4 Yr Old WGC - 2009

```
 5 judges
10 horses ranked

    Rankings are:                Orderings based on all 5 judges:

  t    g    m    s    b           t    g    m    s    b
 333  333  382  333  411          1    1    3    1    6
 401  379  333  401  382          2    4    1    2    3
   8  401  454  382  379          7    2    9    3    4
 379  382  401   34  415          4    3    2    8    5
 411    8    8  415  333          6    7    7    5    1
 415  415  415    8   34          5    5    5    7    8
 378  396  379  411  378         10   11    4    6   10
 454  411  411  379   17          9    6    6    4   13
  34  378  234  454  401          8   10   12    9    2
 382   34  378  234  396          3    8   10   12   11

              Orderings from deleting one judge at a time:

                   t         g         m         s         b
                   1         1         4         1         7
                   3         6         1         4         3
                   7         3        11         3         5
                   4         2         2        10         6
                   6         7         7         6         1
                   5         5         6         7         9
                  10        13         3         5        10
                   9         4         5         2        13
                   8        10        13         9         2
                   2         8         9        13        11
                                                          Averages
 1-3  Distance   4.123   4.000   8.602   2.000   6.403      5.026
 1-5  Distance   4.243   4.899   9.055   6.403   7.810      6.482
 1-10 Distance   9.695   9.055  11.136   9.539  12.410     10.367

Theta hat for class: 0.480
```

Another blown class where judges m and b were clearly lost.

## Class #73 - Open 3 Yr Old Mares - 2009

```
 5 judges
10 horses ranked

     Rankings are:                 Orderings based on all 5 judges:

   g    b    m    r    t          g    b    m    r    t
  115  115  115  115  115         1    1    1    1    1
  392   97   42   97   42         5    2    4    2    4
   97  744  744  744   97         2    3    3    3    2
  744  392  392   42  744         3    5    5    4    3
   42   42  347  347  347         4    4    6    6    6
   44  347   97  392  392         8    6    2    5    5
  177   44  177  177  177         7    8    7    7    7
  347  177  639  639   44         6    7    9    9    8
  639  639   44   44  241         9    9    8    8   10
  241  241  241  457  639        10   10   10   11    9

             Orderings from deleting one judge at a time:

                     g       b       m       r       t
                     1       1       1       1       1
                     5       3       4       2       5
                     2       4       3       3       2
                     3       5       5       4       3
                     4       2       6       6       6
                     8       6       2       5       4
                     7       8       7       7       7
                     6       7       9       9       8
                     9       9       8       8      10
                    10      10      10      11       9
                                                          Averages
   1-3  Distance   3.162   1.414   2.000   0.000   3.162   1.948
   1-5  Distance   3.464   3.464   2.449   1.000   3.464   2.768
   1-10 Distance   4.472   3.742   4.899   2.236   4.243   3.918

Theta hat for class: 0.251
```

A good job of judging on this class! Note judge r was in near complete agreement all the way to 10th place.

### Class #44 - Open Am 4 Yr Old Mares - 2009

```
  5 judges
 10 horses ranked

     Rankings are:                Orderings based on all 5 judges:

   g    m    b    s    r          g    m    b    s    r
   8  382  382  410  124          6    2    2    5    3
 454  454  355    8  410          1    1    4    6    5
 378  415   17  454  355          8    7    9    1    4
 355  124  124  382   17          4    3    3    2    9
 415  410  454  124  454          7    5    1    3    1
 382  378  415  355    8          2    8    7    4    6
 124  267  378   34   88          3   12    8   10   11
 410  355  384   28  382          5    4   14   13    2
  34    8    8   88  415         10    6    6   11    7
  17   88  410   17   34          9   11    5    9   10


          Orderings from deleting one judge at a time:

                g        m        b        s        r
                6        5        5        7        4
                3        1        6        6        6
                9        8        9        1        3
                5        3        3        2        9
                8        6        1        3        1
                1        9        7        4        5
                2       14        8       13       14
                4        2       14       14        2
               11        4        4       10        7
                7       11        2        9       10
                                                          Averages
  1-3  Distance   7.874    6.481    8.246    7.483    5.000      7.017
  1-5  Distance   8.485    6.633    9.220    8.000    8.124      8.092
  1-10 Distance  12.288   12.806   14.560   11.916   12.490     12.812

Theta hat for class: 0.625
```

Holy cow! The data say it all. Chaos!

## Class #133 - Open Am 3 Yr Old Mares  - 2009

```
  5 judges
 10 horses ranked

    Rankings are:                Orderings based on all 5 judges:

  t    r    g    s    m           t    r    g    s    m
 434   27  772  434  434          1    3    2    1    1
  27  772  434  772  772          3    2    1    2    2
 772  434   27   27  594          2    1    3    3    5
 677  594  594  677   27          4    5    5    4    3
 656  677  677  594  677          7    4    4    5    4
  84  656  656   84   84          6    7    7    6    6
  26   84   84  656   47          9    6    6    7    8
 594  299  299   47   26          5   10   10    8    9
  47   47   47   26  656          8    8    8    9    7
 299   26   26  299  299         10    9    9   10   10

         Orderings from deleting one judge at a time:

                t       r       g       s       m
                2       3       2       1       1
                3       2       1       2       2
                1       1       3       3       5
                5       5       5       4       3
                7       4       4       5       4
                6       7       7       7       7
               10       6       6       6       8
                4      10      10       8       9
                8       8       8       9       6
                9       9       9      10      10
                                                    Averages
 1-3 Distance   2.449   2.828   1.414   0.000   2.000   1.738
 1-5 Distance   3.317   3.162   2.000   0.000   2.449   2.186
 1-10 Distance  6.164   4.243   3.464   1.414   4.243   3.906

Theta hat for class: 0.258
```

A well judged class. All judges kept the top 3 at or near the top.

### Class #47  Open Am 4 Yr & Older S&G - 2009

```
  5 judges
 10 horses ranked

     Rankings are:                Orderings based on all 5 judges:

   s    r    t    g    m           s    r    t    g    m
  379  379  401  396  333          2    2    1    7    3
  401  205  379  401  222          1    4    2    1    6
  333  333  207  333  411          3    3    8    3    5
  205  411  205  205  401          4    5    4    4    1
  290  401  411  379  205         10    1    5    2    4
  222  396  396  222  379          6    7    7    6    2
  411  234  222  411  234          5    9    6    5    9
  234  207  333  207  396          9    8    3    8    7
  207  290  234  290  207          8   10    9   10    8
  396  222  290  234  290          7    6   10    9   10

          Orderings from deleting one judge at a time:

                 s       r       t       g       m
                 2       2       3       8       4
                 1       4       2       2       8
                 4       3      10       3       5
                 3       6       4       4       2
                10       1       5       1       3
                 7       7       7       6       1
                 5      10       6       5      10
                 9       8       1       7       6
                 8       9       8      10       7
                 6       5       9       9       9
                                                      Averages
  1-3  Distance  1.732   2.236   7.280   7.000   7.000     5.050
  1-5  Distance  5.385   5.000   7.280   8.062   7.550     6.655
  1-10 Distance  7.211   7.746  10.296   8.485  10.000     8.748

Theta hat for class:  0.532
```

Another class with serious disagreements.

## Class #67  Open Am 5 Yr & Older Mares - 2009

```
  5 judges
 10 horses ranked

     Rankings are:              Orderings based on all 5 judges:

   g    b    m    r    t           g    b    m    r    t
   41   41   41   41   41          1    1    1    1    1
  321  250  250  321  250          2    3    3    2    3
  394  321  321  250  110          9    2    2    3    7
  424  424  296  424  321          4    4   10    4    2
  350  416  424  276  350          5    8    4    6    5
  250  350  110  110  424          3    5    7    7    4
  276  276  350  350  276          6    6    5    5    6
  416  394  276  416  416          8    9    6    8    8
   56  110  416   56  394         11    7    8   11    9
  110   56  394  394   56          7   11    9    9   11


          Orderings from deleting one judge at a time:

                  g      b      m      r      t
                  1      1      1      1      1
                  3      3      3      2      3
                 10      2      2      3      9
                  4      4     11      4      2
                  6      9      4      7      5
                  2      5      7      6      4
                  7      7      5      5      6
                  8      8      6      9      7
                 11      6      8     11      8
                  5     11      9      8     11
                                                   Averages
 1-3  Distance  7.071  1.414  1.414  0.000  6.083    3.196
 1-5  Distance  7.141  4.243  7.211  2.000  6.403    5.400
 1-10 Distance  9.798  5.385  7.937  4.123  7.000    6.849
```

Theta hat for class: 0.383

This class was marginal at best. Three judges agreed with their peers on the first three places but even then m placed entry 296 4th which was way too high according to his peers. Judges t and g apparently made bad choices for their 3rd place picks. Judge r did well with all 10 places.

## *Grand averages for 10 classes - 2009*

```
       Placed an 8,9,10   Placed a 1,2,3    Was the only one to place
       as 1, 2, or 3      as 8, 9, or 10    a particular horse          Average 1st two
  Judge
    b          0.29               0.29               0.29                      0.29
    g          0.33               0.00               0.00                      0.17
    m          0.67               0.11               0.56                      0.39
    r          0.00               0.11               0.11                      0.06
    s          0.12               0.12               0.12                      0.12
    t          0.25               0.25               0.00                      0.25
```

Average theta: 0.409

```
              1-10     1-5     1-3    Variance     # classes(10)

   Judge-b:   7.59     5.27    3.85    7.403             7
   Judge-g:   7.74     5.46    4.52    6.056             9
   Judge-m:   9.04     6.39    5.21    7.910             9
   Judge-r:   6.35     4.04    2.05    2.769             9
   Judge-s:   7.28     4.95    3.30    6.730             8
   Judge-t:   7.29     5.11    3.93    3.497             8
   Averages   7.55     5.21    3.81
```

The summary shows that as a group these judges did not agree was much as we would have liked
as is reflected by the average value of theta being 0.41 and the average 1-3 distance being 3.81.
Remember that values of theta 0.3 or less and 1-3 average distance of 3 or less are what would
Indicate a set of judges had generally agreed. Judge r did an outstanding job with an average 1-3
distance of 2.05 and a very small variance of 2.8 which indicates that his 1-3 distance was generally
small because his average was small and there was little variation from his average. Judge s performed
well having an average distance of 3.3 over all 10 classes. Judges b & t had average distances larger
than desired but judge t showed consistency by having a relatively small variance. Judges g & m
had large average distance values and large variances indicating that they often but not always judged
very differently than their peers.  All the judges except r and s picked horses differently than their peers
as is indicated by the frequency with which they picked a horse in the lower third of their peers' composite
rankings as one of their top 3 choices and judges b and t also made the reverse picks by placing a top 3
consensus horse as either 8, 9, or 10. Judges b and m also were the only ones to pick a particular horse
more often than the other judges. Judge m frequently picked horses very differently than the others as did
judge b but judge b's differences were less extreme. Clearly judges r & s performed better than the others.
These two are judges whom we would recommend be hired for future shows.