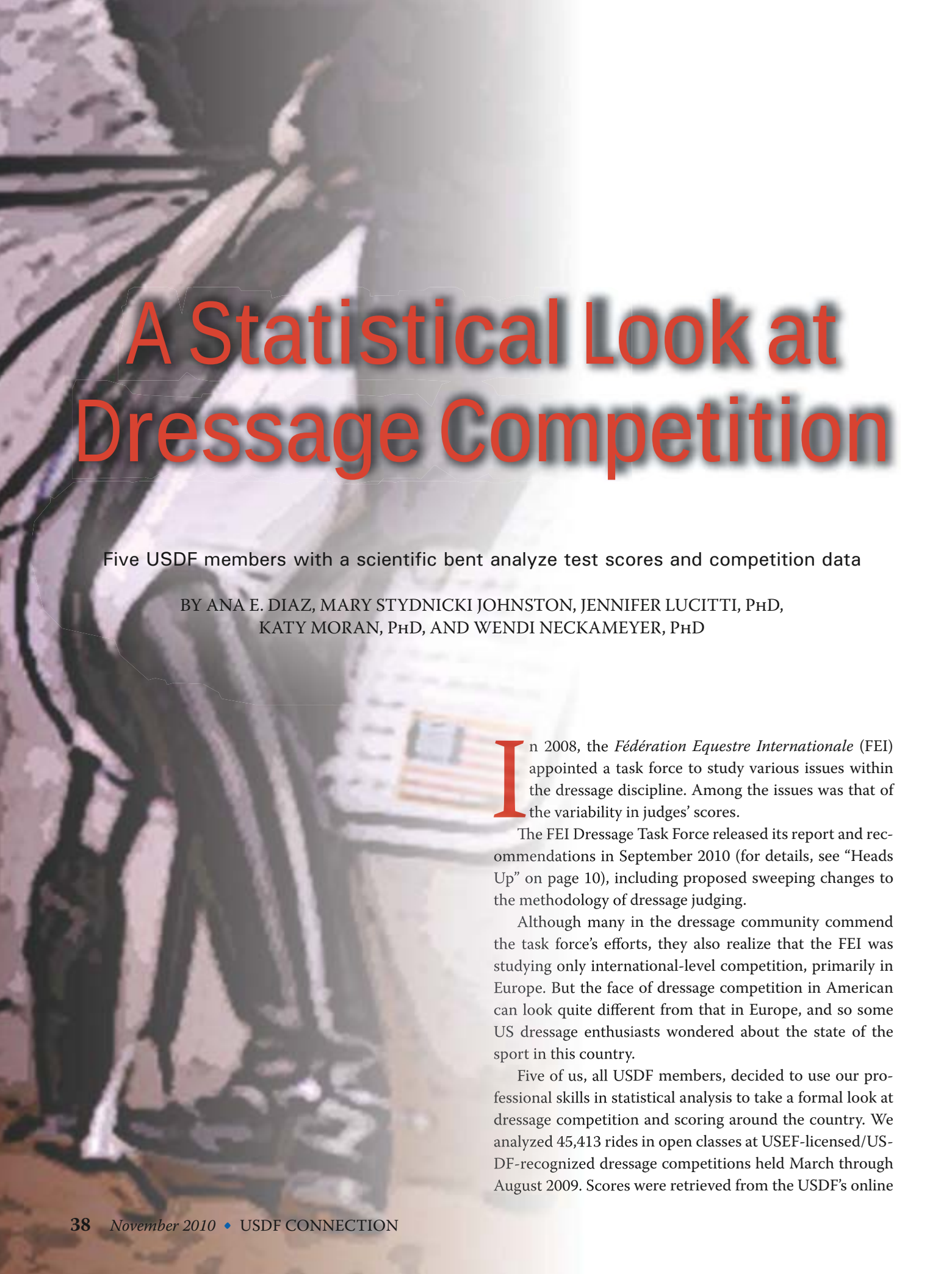


A Statistical Look at Dressage Competition

Five USDF members with a scientific bent analyze test scores and competition data

BY ANA E. DIAZ, MARY STYDNICKI JOHNSTON, JENNIFER LUCITTI, PhD, KATY MORAN, PhD, AND WENDI NECKAMEYER, PhD

In 2008, the *Fédération Equestre Internationale* (FEI) appointed a task force to study various issues within the dressage discipline. Among the issues was that of the variability in judges' scores.

The FEI Dressage Task Force released its report and recommendations in September 2010 (for details, see "Heads Up" on page 10), including proposed sweeping changes to the methodology of dressage judging.

Although many in the dressage community commend the task force's efforts, they also realize that the FEI was studying only international-level competition, primarily in Europe. But the face of dressage competition in American can look quite different from that in Europe, and so some US dressage enthusiasts wondered about the state of the sport in this country.

Five of us, all USDF members, decided to use our professional skills in statistical analysis to take a formal look at dressage competition and scoring around the country. We analyzed 45,413 rides in open classes at USEF-licensed/USDF-recognized dressage competitions held March through August 2009. Scores were retrieved from the USDF's online

listings of show results. The resulting peer-reviewed paper, "Scoring Variables and Judge Bias in United States Dressage Competitions," was published in the *Journal of Quantitative Analysis in Sports* (JQAS), Vol. 6 (2010), Issue 3. (Read the paper online at bepress.com/jqas/vol6/iss3/13.)

Our study identified three dominant variables that affect overall score patterns in US dressage competition:

- Location
- Breed of horse
- Variability among judges.

In this article, we'll share our findings and their possible implications for the sport.

Scoring Trends

Of the 45,413 rides that we studied, only 641 (1.41 percent) earned scores below 50 percent, which could be considered unsatisfactory. (According to the zero-to-10 scale of marks, a score of 5 is "marginal.")

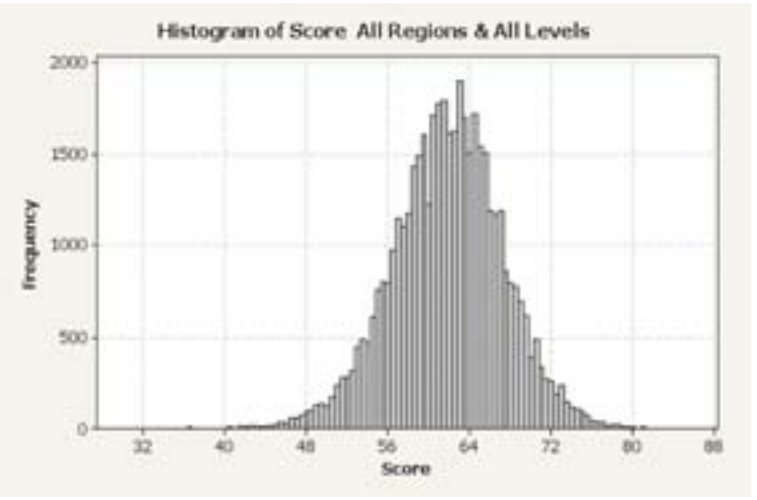


FIGURE 1. *Histogram of scores across all levels.*

Scores in the 50-to-59-percent range accounted for 28 percent of the total rides, with 62 percent of all rides earning scores of between 60 and 69 percent. Figure 1 shows the curve for all scores.

Average scores earned in the 2009 competition year were 0.5 point higher than the averages from 2008. When analyzed by region and level, there were no general statistical differences between 2008 and 2009 scores; but there were some notable differences in certain specifics. For instance, the overall Region 9 scores for 2009 were significantly higher than those for 2008, mainly as a result of higher scores earned by Training and Second Level riders. Looking at scores across all levels, Regions 2, 3, and 7 had the highest averages. Region 4 reported the lowest average scores, as it did in 2008.

COURTESY OF KIMBERLY TAYLOR

Breed	Number of Horses	Mean Score
Quarter Horse	1,555	59.966
Appaloosa/Paint	1,034	60.076
Thoroughbred/TB cross	3,277	60.109
Arabian/Arabian cross	2,782	60.180
Morgan	964	60.758
Iberian	1,785	61.245
Connemara/Welsh	844	61.369
Percheron	268	61.449
Warmblood	27,973	62.482

TABLE 1. *Average scores by breed.*

Breeds and Scoring

We wanted to put numbers to the often-heard claims that non-warmbloods score lower in dressage than warmbloods. After grouping dressage scores by nine breed types (Table 1), we found that, on average, warmbloods score just 2.5 percentage points higher than other breeds.

These findings indicate that many breeds can in fact compete successfully in dressage. Considering that many non-warmbloods and their riders come to the sport from other disciplines, these findings are encouraging indeed.

Judges and Scoring

Dressage judges do two things when they preside over a class: They compare horse/rider combinations to one another (to determine class rank), and they compare each competitor to the defined standards for each movement (to determine the score).

Judging variability means that two or more judges award different marks for the same movement. The vantage point affects what a judge can see and assess—which is why classes with multiple judges seat them at different places around the arena. Although some variation in individual marks is therefore expected and tolerated, it would be considered highly irregular for, say, one judge to award a ride a score of 50 percent and another judge to give the same a ride a score of 70 percent.

To study the issue of judging variability, we looked for similar classes in similar regions with similar horse breeds. We decided to examine scores from open Training Level classes (the classes with the largest numbers of rides), ridden on Dutch Warmbloods, Hanoverians, and Oldenburgs (the predominant breeds in these classes) in Regions 2 and 7 (the regions with the largest numbers of rides). We chose these criteria because they yielded the greatest number of

comparable rides and the most "robust" (strongly significant) data set possible. The variations in average scores, as grouped by the highest-scoring 10 percent of judges and the lowest-scoring 10 percent of judges, are listed in Table 2.

The highest-scoring judges in our study awarded scores of 10 to 13 percentage points higher than other judges. This finding suggests that the judge can have up to four times as much influence on the final score as the horse breed, which is related to a variation of only 2.5 percent.

Of course, issues regarding judging variability are not unique to the sport of dressage. Perhaps the USDF "L" Education Program faculty, the USEF Dressage Committee, and other bodies can use this information to evaluate their training and continuing-education programs. Meanwhile,

Meet the Authors

Ana E. Diaz is a registered professional engineer with degrees in mechanical and metallurgical engineering. She has a 28-year career at a large chemical company working

on quality and manufacturing-process improvements along with product, technology, and operations cost reduction. She owns two Lusitano horses and is a member of the Potomac Valley Dressage Association (PVDA).

Mary Stydnicki Johnston is a longtime PVDA member and a former PVDA treasurer. She is retired from federal service,

where her last assignment was as the leader of a multimillion-dollar Operations Research, Modeling, and Simulation office, which provided decision-analysis and process-modeling tools to improve government operations. She owns three horses and has competed through First Level.

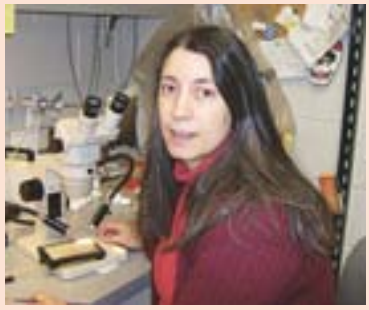
Dr. Jennifer Lucitti has trained her own horses and shown successfully through

Fourth Level, and has earned her USDF bronze medal. She holds an MS and a PhD in physiological sciences and is currently a research associate in the Department of Cell and Molecular Physiology at the School of Medicine, University of North Carolina at Chapel Hill. She is experienced in using statistics to design experiments and to evaluate data, and she has reviewed peers' material for scientific journals and the American Heart Association.

Dr. Katy Moran holds undergraduate and graduate degrees in anthropology and works as a freelance photographer, writer, and graphic designer. She lives in the rural Pacific Northwest,

where she has recently converted from hunters to dressage with her young horse.

Dr. Wendi Neckameyer is a professor in the Department of Pharmacological and Physiological Science at Saint Louis (MO) University School of Medicine. She has reviewed papers for more than 30 scientific journals. She has shown successfully through First Level and



is a past board member of the St. Louis Area Dressage Society. She is also co-president of the charitable organization Special Horses Inc.

The authors thank Lita Dove, S. M. L. Gray, and Amanda M. Jay for their invaluable contributions to the final paper.

PHOTOS COURTESY OF THE AUTHORS

Region	Average Score, Top 10% of High-Scoring Judges	Average Score, Lowest 10% of Low-Scoring Judges	Spread
Region 2	70.918%	57.810%	13.108%
Region 7	69.552%	59.763%	9.789%

TABLE 2. *Variations in average scores for highest- and lowest-scoring groups of judges.*

dressage competitors may be wise to look not only at their final scores but also at their placings. Especially if you receive a score that is unexpectedly high or low, ask yourself: Did the judge rank your performance correctly among the others in your class?

Where Do We Go from Here?

Our goal in undertaking this nine-month study was to highlight and quantify some of the issues facing US dressage—to let the data highlight trends and suggest directions for further study. The JQAS article was the first statistical paper on dressage to be published in a peer-reviewed journal. The analysis of the main data set points to some possible follow-up questions that could be investigated. ▲

Advertise in the *USDF Connection Yearbook*

The largest and most collected issue of the year.

Place a Spotlight Ad in the USDF Yearbook!

Spotlight ads are a great way of letting everyone know about the success of your riding, breeding, and training program – tell the whole dressage community about your farm successes and happy clients.

Mom, dad, grandparents and friends congratulate your family and friends for a winning year!

Insertion order due 12/5/10
Materials due 12/15/10

Call or e-mail
Shelly Stanley at 859-971-7828
or sstanley@usdf.org.