

Executive Summary: A Statistical Analysis of Dressage Scores

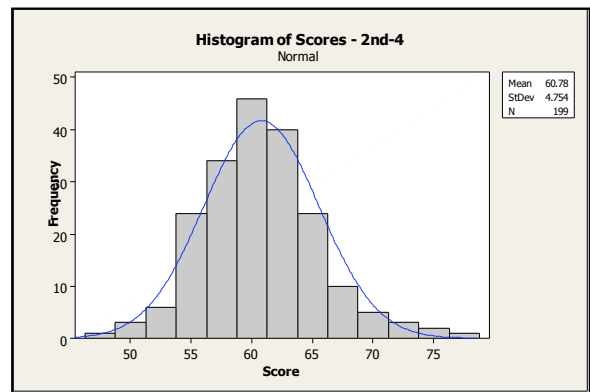
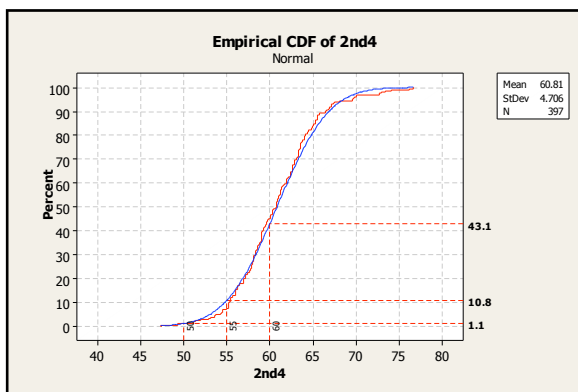
USEF Proposed Rule DR119.3 (tracking number 275-07) had the stated intent of introducing a “timeline for dressage performance standards system.” After considerable concern was voiced by members of the dressage community as a whole, this rule change was tabled by the USEF Board of Directors until 2009. The USEF Dressage Committee is developing more details on the proposed performance standards, which will be released shortly to GMO presidents and participating member delegates for distribution to members for comment.

The proposed rule change created much concern and confusion among dressage riders across the country. Three PVDA members (Yount, Diaz, and Johnston) undertook an analysis of a statistically representative number of competition scores to determine the extent of “poor riding.” This article provides a brief summary of those studies.

This analysis was done on 6,065 rides from “open” dressage performance shows from all USDF Regions (except Region 6, which did not have any competitions in this time frame) between 1-October thru 30-November 2007. The data was “normal” (e.g., bell curve) for all levels, which makes it well suited for statistical analysis. The authors are presently attempting to obtain data for the entire 2007 competition year so that we may extend the analysis.

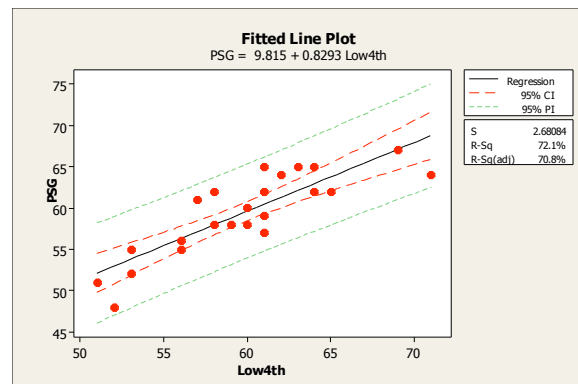
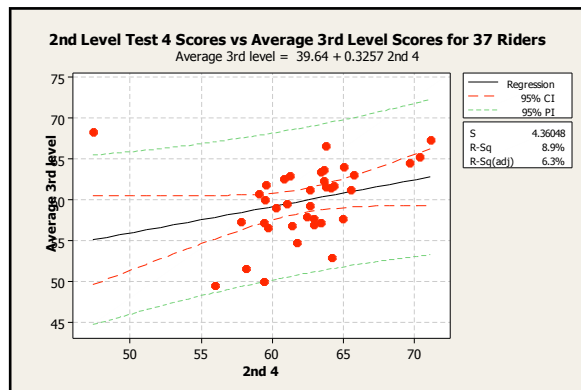
Findings:

- | Level | Average Score | |
|------------------|---------------|---|
| • Training Level | 62.316 | <ul style="list-style-type: none"> • Only 0.02 % of all scores were in the 30-39% range (all were at Training and First Level) • Only 1.78 % of all scores were in the 40-49% range • Only 1.1% of 2nd4 riders scored below 50% • 10.8% of 2nd4 riders scored below 55% • 43.1% of 2nd4 riders scored below 60% |
| • First Level | 61.739 | |
| • Second Level | 60.591 | |
| • Third Level | 60.198 | |
| • Fourth Level | 61.313 | |
| • PSG | 60.339 | |
| • Intermed | 60.122 | |
| • GP | 59.116 | |



Since the Dressage Committee proposed using a system of scores at preceding levels (including the highest tests at 2nd level and 4th level) to qualify riders to ride at 3rd level

and then at PSG, we also looked at this issue. To determine whether lower-level test scores could reasonably be used to predict upper level performance, a regression analysis was run on the 2nd 4 versus 3rd Level average scores and 4th 3 versus PSG scores. This analysis showed that no strong linkage exists between scores earned at 2nd 4 versus average scores at 3rd level. A stronger connection was found between scores at 4th 3 vs PSG, though it was between the LOWEST scores earned at 4th 3 and the LOWEST scores earned at PSG.



Conclusions:

The high average scores earned at each level, the very small percentage of scores at all levels below 50%, and the distribution of scores suggest that if judges feel they are seeing “poor” or “abusive” riding, they are not documenting this on score sheets. Also, 2nd Level Test 4 has little predictive value of success at 3rd Level. The 4th Level Test 3 scores have a higher predictive value for future scores at PSG.

Opinions:

The authors note the following items extracted from the USDF 2007 Strategic Plan:

- USDF needs to have strategic focus and there is more work to be done for members than there are resources to accomplish it.
- Financial issues facing members were identified as a major challenge for people involved in horses.
- There is a declining trend in accessible training.
- Finally, and most important to the qualifying rule question, the Strategic Plan recommends that the USDF should create a more “inclusive” culture.

A system of recording and reporting qualifying scores at one level before being allowed to compete at the next imposes a burden of cost and time on show managers, secretaries, and USDF staff—and these are costs which will be passed on to already burdened riders. Devoting resources to implement and manage a “qualification rule” will dilute and detract both human and financial resources that would be better directed towards furthering the USDF’s Strategic Plan.

Questions, answers & some other ideas:

Why 3 people undertook a statistical study & other questions people are too nice or polite to ask

Why did you 3 people undertake this analysis?

- ✓ Yount and Johnston started the analysis when the statement “seeing 30% rides” attracted their attention. They wanted to determine the extent of this problem.
- ✓ Diaz became involved after seeing the internet discussions. It became evident to her that implementing a qualification rule was analogous to the concept of a “Lower Specification Limit” (LSL) that she sees in her day-job doing quality engineering and contacted Yount with an offer to help with the analysis.

What is your "agenda" in disseminating your analysis before there is a rule change formally on the table?

- ✓ The authors’ “agenda” is to share information. We embarked on the first analysis in response to the original qualification rule proposal. We believe that the analysis created awareness and raised concerns that had not been considered before. The data analysis will be valid for any future qualification proposals.
- ✓ The authors’ agenda is to create awareness of the need for “data-based” decision making.
- ✓ Another item that could be called an “agenda” is that we wanted a dialog and transparency in the discussion with the entire dressage community.
- ✓ But first comes a need to clearly define the problem. If the problem of poor riding is truly widespread, then a potential solution “might be” a qualification rule.
- ✓ If not based on data, statements of a "problem" can actually be only a visible symptom of the real (underlying) problem. Much like a headache can be a symptom of either a hangover or a brain tumor, in either case the symptom may be improved with aspirin, but the root cause of the problem will not be addressed. We questioned whether the USEF Dressage Committee had considered the fundamental “root causes”, such that other potential solutions (such as expanded educational opportunities) could be explored.
- ✓ We are also concerned that a qualification criteria designed absent a clear definition of the problem will have effects on dressage that will not be recognized for several years.
- ✓ We hope that the dialog engendered by this discussion will lead to a proposed course of action that truly improves dressage in the United States.
- ✓ We are interested in:
 - Clear data-based statement of the problem
 - Transparency, inclusiveness and the ability to accommodate multiple perspectives
 - A data-based decision process to define potential solutions to the stated problem

What were your individual roles in this analysis?

- ✓ Yount did the initial data gathering
- ✓ Diaz’ role was of statistical geek. She got involved when it became obvious that having access to software capable of doing statistical analysis would be useful. The Minitab Statistical Software tool was used to perform the regression analyses.
- ✓ Johnston framed the discussion of the results and edited the text of the final document

Additional points considered by the authors:

When you recommend to "simplify" instead of adding complexity, what do you mean and what ideas to you have?

- ✓ Judges already have a scale of 0-10 by which to separate poor, marginal, satisfactory, and exemplary riding. There is no need to layer additional complexity with a qualification rule. The need is to protect the judges when they score according to their conscience under the existing rules.
- ✓ In the case where the riding could be considered abusive of the horse, existing systems can be reconsidered or improved in order to permit Technical Delegates to address abuse without personal repercussions.

If a qualification rule is determined to be the appropriate course of action, what are some alternatives to the original proposal?

- ✓ In order to judge the rider and not the horse, use rider scores or scores from a dressage equitation tests
- ✓ Design a rider qualification test based on 3rd-1 skills for advancement to 3rd and 4th levels and another based on PSG skills for advancement to FEI levels.
- ✓ Utilize the USEF High Performance System to track 3rd, 4th and PSG riders. It currently only tracks Intermediare and GP.
- ✓ It is an "opt in" system, where a rider must declare the intent to participate and maintain the required scores
- ✓ By extending the system to the 3rd, 4th, and PSG riders, it would help to build a base of riders who are both technically qualified and interested in participating, without baring from competition those riders who have no interest in an international competition track.

Who are you guys anyway?

- ✓ **Dr. Rebecca Yount** (dryount@yahoo.com) is a long-time PVDA member, currently on the Board of Directors, Chair of the Education committee and PVDA Rule-Change Liaison. She is a long time USEF and USDF member, who has her Bronze Medal and 4th level scores toward her Silver. In her professional life, she provides psychology services to clients with developmental disabilities, including the administration and interpretation of standardized tests and the analysis of behavioral data.
- ✓ **Ana E. Diaz** (dekko@rcn.com) is a registered Professional Engineer with a strong background in technology and operations. She has a 25-year career at E. I. DuPont de Nemours improving quality thru statistical process control and TQM (total quality management). She finds the question of qualifying scores analogous to the quality problems she faces in her "day-job." She is interested in applying her professional skills to help improve dressage in the US. Ms Diaz is a 4-year member of PVDA.
 - The question being discussed (poor riding) is exactly analogous to poor quality in a manufactured product. The question in both cases is how to improve quality in a cost-effective way.
 - It seemed to Diaz that concepts from quality engineering may help the USEF Dressage Committee. The USEF need not re-invent the wheel, but can utilize the existing body of knowledge already developed for industrial quality management to improve judging and address questions raised by the proposed rule.
 - From the perspective of "Quality Management Systems," putting up a quality standard (Lower Specification) to reduce a quality defect, does not eliminate the defect. The quality standard just eliminates the "defectives" from the population being measured, eg., in this case, competitors presenting themselves to show.
 - Poor riding in the show ring is a symptom of a problem, not the problem itself. There is an underlying question, which is how to improve the quality of riding itself.
- ✓ **Mary Stydnicki Johnston** (shoten.star@verizon.net) is a long-time USDF, USEF, PVDA member and previously served on the PVDA Board in the position of treasurer. She is retired from federal service, where her last assignment was as the leader of a multi-million dollar Operations Research, Modeling and Simulation office which provided decision analysis and process modeling tools used to improve government operations.

The authors have no personal stake in the outcome of the Performance Rule: Yount has her Bronze Medal and will most likely be grandfathered into any system which may eventually be approved. Diaz shows infrequently, though she was a 2nd level demo-rider for the DVCTA L-Judges March session. Johnston is re-habbing an EPM horse and not likely to show any time soon.