

Is the freestyle dressage competition a reliable test of the horse's performance?

*A. Stachurska^{*1}, M. Pięta¹, W. Markowski² and K. Czyrska¹, ¹Agricultural University, Lublin, Poland, ²Medical University of Lublin, Poland*

Corresponding author: A. Stachurska - nowina@ursus.ar.lublin.pl

Abstract

The objective of the study has been to examine if the freestyle dressage competitions are judged reliably and the final results can be used in the horse breeding. Marks given by a judge for single movements of the dressage program were considered. The data contained 9000 marks of ten best ranked horses at seven Grand Prix classes and two Intermediate I classes from six CDI*** and one European Championships. The marks concerned 84 horses in total and were given by 25 judges. The reliability of judging was measured with the index of disagreement (ID). The index measures the disagreement of ranking by a particular judge relatively to the general ranking based on the marks of five judges.

Only four judges had the mean ID lower than 10%, eight judges had it greater than 20% and the highest mean index amounted to 28.9%. The judging was less reliable in Intermediate than in Grand Prix. The average index ranged from 9.5% to 23.6% at particular classes. In both Intermediate and three Grand Prix classes the average ID of at least one judge exceeded 20%. The study shows that the results of the freestyle dressage competition should be used in the breeding with caution. The present system of judging should be permanently checked by statistic methods.

Introduction

The dressage performance may be considered as one of the indicators of the horse's breeding value provided it is judged reliably. The dressage qualities are particularly difficult to judge since they cannot be measured. The judges are taught how to judge during courses and seminars but there is no handbook which would precise the code point system, i.e. which mark to give for a movement executed in a certain way (Niggli, 2003). Many factors decide on the way of judging, first of all the personality of a judge: his knowledge, sport experience, training in judging, ability to concentrate for a long time, consequence, fair play, quickness of reaction etc. The position of the judge at the arena is also important. The judging should be as objective as possible, with the influence of the factors the least differentiated.

The freestyle dressage competitions are run according to Fédération Equestre Internationale (FEI) rules. The program consists of 15 movements which should be executed (can be repeated) in a free order and arrangement to the music of one's choice. The time is allowed from 4'30" to 6'00" according to the class. Each of five judges situated at different places around the arena (E,H,C,M,B) gives a technical mark for each movement with the use of the scale from 0 to 10. Moreover, the judge gives so-called artistic marks in the same scale for the general impression: 1) rhythm, energy and elasticity, 2) harmony between rider and horse, 3) choreography, use of arena, inventiveness, 4) degree of difficulty, well calculated risks, 5) choice of music and interpretation of the music. Some of the marks are multiplied by coefficients. The final ranking of horses in a class is made according to totals: all the marks given by the five judges. The final results are presented in per cents of the total sum which can be scored.

To state if the judging in a class is proper, particular marks for single movements must be taken into account. Two aspects should be considered:

- 1) if the marks from five judges agree among themselves with regard to their level
- 2) if the judges rank the horses in the same way.

The objective of the study has been to examine if the freestyle dressage competitions are judged in agreement with regard to the ranking of the horses at particular movements and so if the final results are reliable enough to be used in the horse breeding.

Material and Methods

The data contained 9000 marks of ten best ranked horses at seven Grand Prix classes (GP) and two Intermediate I (INT) classes from CDI*** (Concour de Dressage Internationale), World Cup and European Championships (Tab. 1). Grand Prix Free Style Test is the most difficult class in dressage free style competitions. Intermediate I is easier by two levels. The marks concerned 90 classes made by 84 horses in total and were given by 25 judges. The coefficients and final classification at the competitions were not considered.

Table 1. The freestyle classes examined

Competition	Year	Intermediate I	Grand Prix
CDI*** Józefin, Poland	1997		x
Volvo World Cup Göteborg, Sweden	1998		x
CDI*** Józefin, Poland	2000	x	x
W/CDI*** Lipica, Slovenia	2001		x
CDI*** Moscow, Russia	2002		x
FEI World Cup Kaposvar, Hungary	2003	x	x
European Championships Hicksted, Great Britain	2003		x

The reliability of judging was measured with the index of disagreement (ID). The index measures the disagreement of ranking by a particular judge relatively to the general ranking based on the marks of five judges (Niewczas, Hulewicz-Stachurska, 1991, Stachurska, 1991, Stachurska et al., 2005). The latter is assumed to be proper. The fact that only ten best ranked horses were considered in each class allowed the ID measure to be identically precise in these classes because the ID level is connected with the number of horses in a class.

HSD Tukey's test was used to determine the significance of differences among ID means for the judges at different positions at the arena, as well as in INT and GP classes in total.

Results

Table 2 shows average ID for particular judges marked with successive alphabet letters, calculated on the bases of one to six classes. Only four judges had the average ID lower than 10%, eight judges had it greater than 20% and the highest mean amounted to 28.9%.

Table 2. Average ID (%) for particular judges

Judges	Number of classes judged		ID mean
	Intermediate I	Grand Prix	
a	1		28.88
b	1	1	23.51
c	1	1	22.87
d	1	2	22.86
e	1		20.99
f		1	20.44
g		1	20.32
h		1	20.09
i		1	19.77
j		1	18.18
k		1	16.35
l	1	1	15.77
m		1	15.69
n	1	5	15.66
o		2	15.52
p		1	15.33
q	1	3	15.07
r	1	3	14.66
s		1	11.88
t		2	11.69
u		1	10.05
v		1	9.03
w		1	8.52
x		2	9.13
y		1	5.50

The ID varied according to the position of the judge at the arena (Table 3). The differences were more pronounced in case of INT classes than GP classes. The main judges at C position had the lowest ID mean. A slight tendency can be noticed: judges at H and M had it greater, whereas at E and B average ID was the greatest. In INT classes ID mean was greater than in GP classes ($P \leq 0.01$). The little number of significant differences results from the fact the standard deviation (SD) was great. It means the IDs varied to great extent in particular movements.

Table 3. Average ID (%) at various positions at the arena

Classes	Judge's position at the arena										Total	
	E		H		C		M		B			
	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>
Intermediatel	24.88	15.23	18.58	16.12	16.48	14.29	17.17	14.17	19.98	17.01	19.42 B	15.54
Grand Prix	16.59	17.23	16.97	16.54	12.68	14.75	15.09	15.62	15.52	17.30	15.37 B	16.34
Total	18.43 a	17.12	17.33	16.42	13.53 a	14.69	15.55	15.30	16.51	17.29	16.27	16.24

The difference between means marked with the same letters is significant at: small letters - $P \leq 0.05$, capitals - $P \leq 0.01$.

The average ID ranged from 9.5% to 23.6% at particular classes (Tab. 4). In both Intermediate and three Grand Prix classes the average ID of at least one judge exceeded 20%.

Table 4. Average ID (%) in the classes at particular competitions

Competition	Intermediate I						Grand Prix					
	E	H	C	M	B	Mean	E	H	C	M	B	Mean
Józefin 1997							14.9	11.5	13.4	20.1	16.4	15.3
Göteborg 1998							6.4	12.5	8.9	10.0	10.1	9.5
Józefin 2000	28.8	24.7	19.0	16.7	28.9	23.6	18.3	29.5	15.0	20.4	26.7	22.0
Lipica 2001							22.3	20.3	20.1	15.6	19.8	19.6
Moscow 2002							14.4	13.1	9.4	5.5	11.9	10.8
Kaposvar 2003	21.0	12.4	13.9	17.7	11.1	15.2	21.8	17.6	13.0	18.6	15.4	17.2
Hicksted 2003							18.2	14.3	9.0	15.3	8.5	13.0

Discussion

The fact that the ID of 25 judges is so differentiated indicates there are some experienced, fair judges and there are also judges who do not judge quite correctly. The latter should perhaps train more, participate at additional courses and pass examinations. The ID may be used as an objective and quantitative scale at such examinations.

The differences among the judges more pronounced and the ID greater in the lower classes may indicate the judging was less reliable in them than in Grand Prix. However, the two INT classes were judged differently. More classes should be examined to prove if the judging in lower free style classes is less qualified. Usually Grand Prix classes are judged by more experienced judges but it is equally important the younger or less trained horses in Intermediate are judged fairly, as well.

The positions of the judges at the arena suggest that marks by H, C and M judges should vary the least among one another, like between E and B judges. Instead, the marks by H, C, M versus E and B should differ the most. The results show such tendency. It is important for a judge if he watches a movement from in front of him or from a side. Not quite the same elements of the execution are judged then. Generally, the C judge sees the greater part of the programme symmetrically and a lot of movements are presented in front of him. The other thing is the distance from which a judge watches the movements. Once it is e.g. 5 m, once 60 m. Both the judges and the riders ought to realize this fact that should be also taken into account in examining the judges.

Average ID in the classes indicates the quality of judging differs at particular competitions. On the bases of the examined competitions it cannot be stated that the judging actually improves.

The study shows that the results of the freestyle dressage competition should be used in the breeding with caution. The present system of judging should be permanently checked by statistic methods which may help to improve its quality.

Bibliography

- Niewczas J., Hulewicz-Stachurska A., 1991: Metoda oceny wyników sędziowania konkursów ujeżdżenia dla potrzeb hodowlanych. *Annales Universitatis Mariae Curie-Sklodowska, Sectio EE, IX, 22*: 177-185.
- Niggli W.M., 2003: Dressage. A guideline for Riders and Judges. J.A. Allen, London.
- Stachurska A., 1991: Analiza i ocena konkursów ujeżdżenia jako hodowlanych prób użytkowości koni wierzchowych. *Annales Universitatis Mariae Curie-Sklodowska, Sectio EE, IX, 21*: 163-175.
- Stachurska A., Niewczas J., Markowski W., 2005: An estimation of reliability of judging the horse dressage competitions. EAAP Meeting, Uppsala, 2005.