HOW APPARATUS DIFFICULTY SCORES AFFECT ALL AROUND RESULTS IN MEN'S ARTISTIC GYMNASTICS

Ivan Čuk[1] and Warwick Forbes[2]

[1] University of Ljubljana, Faculty of Sport, Slovenia
[2]Australian Institute of Sport, Canberra, Australia

Original research article

Abstract

On a sample of 49all-around male gymnasts at the 2009 European Championships the implications of the difficulty scores were tested in relation to the success in all-around competition. After the regression, cluster and ANOVA analysis, three groups of quality all-around gymnasts were determined, while only one group had a chance to win an all-around medal; difficulty scores between all six apparatus were not equal; the highest prediction of the all-around score was the parallel bars difficulty score.

Keywords**:** *artistic gymnastics, man, strategy, all around.*

INTRODUCTION

After the Olympic Games (OG) in 2004 the International Gymnastics Federation (FIG) made changes to the Code of Points. One of these changes was the implementation of a new philosophy of an open scoring system, prepared by Fink and Fetzer (1993), which had previously been introduced at the FIG symposium in Lugano in 1993. Prior to 2006 all disciplines in men's artistic gymnastics (FIG, 2000) were limited to a maximum final score of 10 points. In the past, different maximum scores were allowed, before World War II (WWII) the maximum score was sometimes between 11 and 16 points. After WWII the maximum score was limited to 10 points (Štukelj, 1989). Despite changes to what 10 points actually represented, it was decided that this represented exercise content and exercise presentation. The proportion of exercise content and exercise presentation had also changed; in the end it became equal to approximately 50:50 (Bučar 1998).

Exercise content was mostly characterized by difficulty and special requirements. In the Code of Points 2006 (FIG, 2006) the whole philosophy of evaluating gymnastics exercises changed. No longer was one maximum score (10 points) for evaluating exercises used. New rules (FIG, 2009) defined D and E score, where D score evaluates exercise content (difficulty, special requirements, and bonus points) and the E score evaluates exercise presentation. 'D scores' start at zero points and increase according to the difficulty the gymnast demonstrates, how the exercise is constructed (the exercise must include elements from all five element groups, and no more than 4 from one group), and how difficult elements are connected (bonus points).

The system works well for apparatus specialists; the more you show the greater the score, however in all around (AA) gymnasts a problem can exist. The problem is with the apparent equality between apparatus i.e. the vault has special rules compared to floor exercise, pommel horse,

rings, parallel bars, and horizontal bar. Gymnasts in AA competition only perform one vault, and compared to the other apparatus the vault is similar to only one element from the other exercises. Therefore, on the vault the D score is known in advance (FIG, 2009). According to the results of men's AA qualifications at OG 2008 Čuk and Atiković (2009) found that the vault is considered to be the most valuable apparatus, and the pommel horse was undervalued among AA gymnasts. Using the Code of Points, it is very hard to obtain a high D score on the pommel horse, whereas it is easier to obtain a high D score on the vault. Pairwise *t*-tests showed that D scores between the vault and other apparatus, and between the pommel horse and other apparatus were significantly different.

Table 1. *Average D score (multiplied by 1000) and standard error of sample (N=44) at OG 2008 for MAG all-around gymnasts Čuk and Atiković (2009)*

	Mean	Std. Error
FXA Score	6015.91	50.572
PHA Score	5677.27	69.189
RIA Score	5943.18	95.257
VTA Score	6445.45	65.306
PBA Score	6090.91	84.834
HBA Score	5897.73	80.530

However, the new Code of Points presented in 2009 (FIG, 2009) has a number of changes that impact the D score. In the past (Hadjijev, 1989), it was expected that the least amount of training time was spent on the vault, and the most amount of time was spent on the pommel horse. Training times on other apparatus were similar (the gymnasts preferences, abilities, and individual characteristics are also important in determining training time spent on each apparatus).

Using the 2009 Code of Points, one of the most experienced Slovenian international judges Enis Hodžić calculated maximum difficulty scores for each apparatus. Results were Floor exercise=7.9; Pommel Horse=7.6; Rings=7.6: Vault=7.4; Parallel Bars=8.1; High Bar=8.5. It is clear that the maximum difficulty scores are different for each apparatus.

The 2009 European Championships (EC) in Milan was the first major competition in the world to use the 2009 Code of points. It is therefore interesting to see how the AA gymnasts coped with the new rules, as their performances might be a guideline for the Olympic cycle up to the OG 2012 in London. The number of AA gymnasts has diminished over the last two decades (at OG in 1992 all the gymnasts were competing in AA in order to get into finals, while at OG 2008 and at WC 2007 only half of them competed in the AA competition). It is interesting to see how all-around gymnasts are coping with the new Code of Points and what kind of strategies they are using to improve their results.

METHODS

Our sample was composed of 49 AA gymnasts who competed at the EC in Milan 2009 qualification event. From official results we made 6 variables of D scores: Floor Exercise (FX) , Pommel Horse (PH) , Rings (RI), Vault (VT), Parallel Bars (PB) and Horizontal Bar (HB) . To evaluate the AA we used the AA final score (AAFS). To assist the statistical presentation, D and E scores were multiplied by 1000; so a score of 6 points had a value of 6000. SPSS 15.0 was used to calculate Kolmogorov-Smirnov to test the normality of the variables distributions, Pearson correlations, pair-wise *t*-tests between D scores of all apparatus, and a linear regression analysis between AAFS and D scores (method enter). We also prepared the classification of gymnasts with the method of Euclidian square distances using D scores. Clusters were then compared with one way ANOVA and Tamahne 2 post hoc test. All statistics used an alpha level of $p < 0.05$.

RESULTS AND DISCUSSION

Table 2. *Descriptive statistics*

	N	Minimum	Maximum	Mean	Std. Error	Std. Deviation	K-S test
FXDscore	49	3900	6400	5381.63	77.956	545.693	Normal
PHDscore	49	3000	6800	4997.96	115.709	809.961	Normal
RIDscore	49	2500	6500	5273.47	103.624	725.366	Normal
VTDscore	49	4600	7000	6012.24	76.578	536.048	Not Normal
PBDscore	49	3400	6500	5202.04	99.808	698.656	Normal
HBDscore	49	3100	6800	5269.39	114.070	798.489	Normal
AAFS	49	64325	89150	81395.41	693.419	4853.936	Normal

Table 3. *Pairwise t-test (N=48)*

Pair	t	Sig. (2-tailed)
FXDscore - PHDscore	3.532	.001
FXDscore - RIDscore	1.170	.248
FXDscore - VTDscore	-10.537	.000
FXDscore - PBDscore	2.151	.037
FXDscore - HBDscore	1.149	.256
PHDscore - RIDscore	-2.713	.009
PHDscore - VTDscore	-8.881	.000
PHDscore - PBDscore	-2.148	.037
PHDscore - HBDscore	-2.390	.021
RIDscore - VTDscore	-7.475	.000
RIDscore - PBDscore	.825	.413
RIDscore - HBDscore	.042	.967
VTDscore - PBDscore	9.105	.000
VTDscore - HBDscore	7.087	.000
PBDscore - HBDscore	-.781	.439

Table 4. *Pearson correlation matrix*

	HBDscore	PBDscore	VTDscore	RIDscore	PHDscore	FXDscore
AAFS	.720*	.830*	.606*	.743*	.697*	.710*
HBDscore	1.000	.682*	.452*	.605*	.511*	.537*
PBDscore		1.000	.517*	.639*	.620*	.583*
VTDscore			1.000	.431*	.350*	.700*
RIDscore				1.000	.576*	.511*
PHDscore					1.000	.425*
FXDscore						1.000

*all correlations are significant p<0.01

Table 5. *Regression analysis (method Enter), predicted AAFS variable*

R	R Square	df1	df2	Sig.
.920(a)	.847	6	42	.000

	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta	B	Std. Error
(Constant)	38244.832	3455.332		11.068	.000
HBDscore	.753	.536	.124	1.405	.167
PBDscore	2.324	.687	.335	3.384	.002
VTDscore	.677	.780	.075	.868	.390
RIDscore	1.357	.578	.203	2.349	.024
PHDscore	1.163	.484	.194	2.401	.021
FXDscore	1.868	.826	.210	2.262	.029

Cluster analysis with the method of Euclidian distances gave the best results with 3 clusters, where 21, 6 and 22 gymnasts were grouped. Those with three clusters were used in further analyses via a one way ANOVA.

Table 6. *ANOVA results with Tamahne 2 post hoc test*

		Sum of Squares	df	Mean Square	F	Sig.
FXDscore	Between Groups	7179162.028	2	3589581.014	23.210	.000
	Within Groups	7114307.359	46	154658.856		
	Total	14293469.388	48			
PHDscore	Between Groups	15670726.654	2	7835363.327	22.784	.000
	Within Groups	15819069.264	46	343892.810		
	Total	31489795.918	48			
RIDscore	Between Groups	12291159.555	2	6145579.777	21.806	.000
	Within Groups	12964350.649	46	281833.710		
	Total	25255510.204	48			
VTDscore	Between Groups	6085986.395	2	3042993.197	18.163	.000
	Within Groups	7706666.667	46	167536.232		
	Total	13792653.061	48			
PBDscore	Between Groups	17029687.693	2	8514843.847	61.199	.000
	Within Groups	6400108.225	46	139132.788		
	Total	23429795.918	48			
HBDscore	Between Groups	18955315.399	2	9477657.699	37.426	.000
	Within Groups	11648766.234	46	253234.049		
	Total	30604081.633	48			

						95% Confidence Interval for Mean	
	Group	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound
FXD score	1	21	5057.14	344.342	75.142	4900.40	5213.89
	2	6	4966.67	674.290	275.278	4259.04	5674.29
	3	22	5804.55	342.925	73.112	5652.50	5956.59

	Total	49	5381.63	545.693	77.956	5224.89	5538.37
PHD score	1	21	4904.76	618.447	134.956	4623.25	5186.28
	2	6	3650.00	476.445	194.508	3150.00	4150.00
	3	22	5454.55	578.773	123.395	5197.93	5711.16
	Total	49	4997.96	809.961	115.709	4765.31	5230.61
RID score	1	21	5119.05	520.211	113.519	4882.25	5355.84
	2	6	4166.67	882.421	360.247	3240.62	5092.71
	3	22	5722.73	417.398	88.989	5537.66	5907.79
	Total	49	5273.47	725.366	103.624	5065.12	5481.82
VTD score	1	21	5666.67	425.833	92.924	5472.83	5860.50
	2	6	5800.00	438.178	178.885	5340.16	6259.84
	3	22	6400.00	385.450	82.178	6229.10	6570.90
	Total	49	6012.24	536.048	76.578	5858.27	6166.22
PBD score	1	21	4952.38	400.773	87.456	4769.95	5134.81
	2	6	4000.00	428.952	175.119	3549.84	4450.16
	3	22	5768.18	328.614	70.061	5622.48	5913.88
	Total	49	5202.04	698.656	99.808	5001.36	5402.72
HBD score	1	21	5061.90	529.600	115.568	4820.83	5302.98
	2	6	3916.67	636.920	260.021	3248.26	4585.07
	3	22	5836.36	437.031	93.175	5642.60	6030.13
	Total	49	5269.39	798.489	114.070	5040.03	5498.74

The descriptive statistics and Kolmogorov-Smirnov test (Table 2) showed that only the vault data D scores were not normally distributed. The score distribution was leptokurtic and skewed to the left, meaning that higher values are more common. Despite trying to normalize variables with logarithmic functions (ln and $\log_{10}$), the abnormality persisted, so we decided to continue analyses with raw data. Comparing the average of all-around D scores on the apparatus from OG2008 and D scores from EC2009 it can be noted that there is a huge lowering of D scores when the2009 Code of Points were used. On average, D scores were lower by 0.6 points; the greatest lowering was on parallel bars, and least on the vault. The 2009 Code of Points did not affect AA gymnasts on the vault, but mostly on the parallel bars.

Table 7. *Differences between AA scores from OG2008 and D scores from EC2009*

	OG2008	EC2009	Diference
FX	6015.91	5381.63	634.28
PH	5677.27	4997.96	679.31
RI	5943.18	5273.47	669.71
VT	6445.45	6012.24	433.21
PB	6090.91	5202.04	888.87
HB	5897.73	5269.39	628.34

At the beginning of the Olympic cycle with the adoption of the 2009 Code of Points lower start values (as the value of some elements were lowered, less bonus points on apparatus) were expected, however the drop in scores was more severe than expected (from 0.43 to 0.88 point). If we compare what AA gymnasts could achieve according to maximum scores using the2009 Code of Points, it is noted that they were already achieving 81.2% of maximum possible score on the vault, while on all other apparatus they are below 70% of the maximum score. If we take into consideration the best gymnast by D score on each apparatus, the percentage of

maximum D scores were higher, but vault was still the apparatus where the best AA gymnast was already reaching 94.6%, while the best gymnast on other apparatus was below 90%.

Table 8. *Theoretical maximum D scores by Code 2009 and achieved ones at EC2009*

Max Dscore Code 2009	Average AA gymnasts % of max Dscore	The best AA gymnast Dscore Code 2009 from AA gymnast	The best AA gymnast % max Code 2009
7900	68.1	6400	81.0
7600	65.8	6800	89.5
7600	69.4	6500	85.5
7400	81.2	7000	94.6
8100	64.2	6500	80.2
8500	62.0	6800	80.0

The pair-wise *t*-test (Table 3) showed 10 significant different pairs out of 15 pairs; all pairs with pommel horse and vault were significant different, and floor exercise with parallel bars. The average D scores on the vault were the highest and were lowest on the pommel horse. Similar results were obtained at OG2008 (Čuk, Atiković, 2009).

Pearson's correlations (Table 4) between apparatus D scores were all statistically significant, medium high. Correlations between all AAFS and each apparatus D scores were slightly higher, the highest was with parallel bars D score (0.83 – 68.9% of common variance). Surprisingly the lowest correlation was with the vault D scores (0.61 – 36.3% of common variance). The descriptive statistics and *t*-tests showed that the vault had important differences to other apparatus, but correlations revealed that for AA gymnasts the vault score had the lowest impact on AA score. Coefficient of multiple correlations (Table 5) between dependent variables of the AA final score and independent variables of apparatus D scores were statistically significant and very high (0.92). D scores explained over 84% of the final AA score, in general more difficult exercises attained better results in the AA. Significant predictors of AA success are parallel bars, rings, pommel horse and floor exercise D scores. It was interesting to observe that the vault and high bar D scores were not significant predictors of AAFS. On the vault there was not enough discrimination among gymnast's D scores..

Cluster analyses identified 3 groups of gymnasts. ANOVA (Table 6) showed they differed significantly on D scores. The third group (22 gymnasts) was very good on all events and had significantly higher D scores on all apparatus compared to the other groups. The first (21 gymnasts) and the second group (6 gymnasts) were equal on floor exercise and vault (the second group exceeded the first); while on the other apparatus the first group had higher D scores. Only the third group had the quality (level of D scores) of winning medals, so the questions to be asked are: why do gymnasts from the first and the second group compete in AA at all? Are they just trying to enter AA finals or are they just young gymnasts with a better potential future?

CONCLUSIONS

Based on the results presented it can be concluded that:

- with the 2009 Code of Points, for all-around results the six apparatus are not equal to obtain D scores;
- with the2009 Code of Points, for all-around gymnasts, the vault and the pommel horse D scores significantly differ from other apparatus;
- with the 2009 Code of Points, the vault D scores do not discriminate between all-around gymnasts;
- all-around gymnasts have the lowest D scores on pommel horse;

- with D scores only we can predict 84% of all-around final score;
- after the Code of Points changed in 2009, the all-around gymnast who attained the highest D score on parallel bars has the best chance of good all-around results;
- D scores for the vault and high bar did not significantly predict all-around final scores; vault D scores did not discriminate sufficiently (to many gymnasts with same D score), while on the high bar the lack of discrimination could be due to an increased number of falls. It seems it is more important to perform a slightly less difficult exercise well than a difficult exercise with a fall;
- three groups of all-around gymnasts were classified (with 21, 6, and 22 gymnasts), and only the third group had potential of winning an all-around medal, as their D scores on all apparatus are much higher.

REFERENCES

Bučar Pajek, M. (1998). *Primerjalna analiza tekmovalnih pravil v moški in ženski športni gimnastiki [Comparative analyse of Men's and Women's Code of points]. Diplomsko delo [Unpublished diploma thesis]*. Faculty of Sport. Ljubljana.

Čuk, I. Atiković A. (2009). *Are Disciplines in All-around Men's Artistic Gymnastics Equal?*. Sport Scientific & Practical Aspects. International Journal of Kinesiology. 6, 1&2:8-13.

Code of Points Men Artistic Gymnastics. (2000). Moutier: FIG.

Code of Points Men Artistic Gymnastics. (2006). Moutier: FIG.

Code of Points Men Artistic Gymnastics. (2009). Lausanne: FIG.

Fink, H., Fetzer, J. (1993). Proposals of the new Code of Points. International symposium on Men's Artistic Gymnastics. FIG. Lugano.

Hadjijev, N. (1989). *Some problems of training young male gymnasts*: FIG Symposium during Artistic Gymnastics World Championship. Stuttgart.

Štukelj, L. (1989). Mojih *sedem svetovnih tekmovanj [My seven world competitions]*. Dolenjska založba. Novo Mesto.