ANALYSIS RESULTS JUDGING WORLD CHAMPIONSHIPS IN MEN'S ARTISTIC GYMNASTICS IN THE LONDON 2009 YEAR

Almir Atiković[1], **Sunčica Delaš Kalinski**[2], **Snežana Bijelić**[3], and **Nađija Avdibašić Vukadinović**[1]

[1]Faculty of Physical Education and Sport, University of Tuzla, Bosnia and Herzegovina
[2]Faculty of Kinesiology, University of Split, Croatia
[3]Faculty of Sport and Physical Education, University of Banja Luka, Bosnia and Herzegovina

SUMMARY

On a sample of 176 male gymnasts, we analyzed the score of judges from the World Championship in men's artistic gymnastics, held 2009 in London. The subject of the analysis were the final scores competitors got for the exercises shown in the qualifying competition (C I). Analysis problem was determination of the differences on individual apparatus between judges E1 to E6 and apparatus. The main objective of this study was to determine the reliability of evaluation of judges and whether the current Code of Points (Federation Internationale de Gymnastique, 2009b) should be revised in terms of equalizing score on apparatus. Equality was tested for the achieved D, E and all-around scores on the disciplines of floor exercise, pommel horse, rings, vault, parallel bars and horizontal bar. Vault has the highest D and E scores, while pommel horse the lowest D and E scores. *T*-tests showed that those two disciplines significantly differ from other disciplines. Reliability were calculated (intraclass correlation coefficient ICR, Cronbach's alpha, differences in mean E1 to E6 between judges were tested using factor analyses with method first major component. All data were analyzed using SPSS Statistics 17.0. Results show very high reliability (e.g. Cronbach's alpha range from .94 up to .98).

Key Words: Men's Artistic Gymnastics, Judging, Code of Points (FIG).

INTRODUCTION

In artistic gymnastics, the emphasis is on the aesthetic component which has to be performed in accordance with the specific conventional structure of a movement. Although the methods of assessment in conventional sports differ either by number of judges, set criteria or method of calculating the final results, it is characteristic for the sports industry that judges represent the measuring instrument and that their assessments are instruments of qualities of contestant's results. Each contestant brings his own competitive exercises which are being evaluated from two perspectives: content and performance of an exercise. Today, for the assessment of artistic gymnastic, the international competitive Code of Points for assessment of men's and women's artistic gymnastics is in effect, which are being improved and published after the Olympic Games finish. Male competition Code of Points for the evaluation of the technical commission is composed by Men's Techni-

cal Committee, The *Fédération Internationale de Gymnastique* (FIG).

The first unique instructions FIG for evaluation of gymnastic exercises were created in 1949. year known as "*Code of Points*" for the assessment of the artistic gymnastics includes seven levels of degree of difficulty. Initial degree of severity represents the level A, and the next levels are B, C, D, E, F, and G (Federation Internationale de Gymnastique, 2009b). The latest one is the greatest degree of severity. The main purpose and goal of the Code of Points for evaluating is provision of more objective evaluation of exercises. Independent members of the Refereeing Commission (D & E commission) are on all apparatus: D commission evaluates (weight, special requirements, and bonus points) and the assessment starts from 0.00 points to more and E commission the performance of an exercise (performance techniques, body posture, and balance), and provides deductions for the performance from 10.00 points to lower. D commission determines the initial assessment of an

exercise, and the E commission registers performance errors due to technical performance, body posture and balance of exercise performance so that those two grades would at the end sum up in the final one.

Several authors have tried to evaluate the quality of judging at different competitions. Ansorge, Scheer, Laub, and Howard (1978) found bias in scores induced by the position in which female gymnasts appear in their within-team order. Ansorge and Scheer (1988) found biased judging towards judges' own national team and against immediate competitors' teams. Hraski (1988) analyzed judging at the World Cup in 1982 in all male disciplines.

In rhythmic gymnastics, for the purpose of this study Popović (2000) designated a bias pattern with the international judges at the competition in rhythmic gymnastics at the Olympic games held in Sydney in 2000. The results of analysis conducted on the basis of a test of proportions (relative to the number of major, minor, or identical assessments) in the qualifications for individual all-around competition indicate on biased evaluation of competitors from their own countries. Woman judges evaluated gymnasts from their federations with higher assessment than the other scoring woman judges.

Sands (2010) in his research "Judging in real time" mentioned the biggest problem of evaluation and that is: reliability and validity. In his paper, the author mentions that the judges could use modern technology and with that, immediately after the performance, give their deductions so that a smaller number of judges would stay at rank. Other authors have dealt with this issue, too, such as Čuk and Forbes (2006) who have made the program B Jury Judging Real Time System (RTJS) at the Australian Institute for Sport. The program has improved the objectivity of the evaluation by Jury B Execution Deductions entered during the performance and it cannot be changed, judges must deduct quickly and precisely each time they see an error. This program is approved by the Technical Commission of the *European Union of Gymnastics* (UEG) which was first used at the European Championships in Berlin held in 2011.

Leskošek, Čuk, Karácsony, Pajek, and Bučar (2010) in results show very high reliability and satisfactory validity of judging at the University Games. It should be emphasized that judging quality differs between apparatus, sessions and judges. In different sessions and apparatus all reliability measures (Cronbach's alpha range from .92 up to .99, ICC, Armor's theta) are higher than .90. Those indices tend to be a little lower in the all round finals than in qualification and apparatus finals. There appears to be no systematic differences in reliability between apparatus. Vault scores tend to have lower reliability than other apparatus in qualification and all around. Armor's theta ranged from .92 (on the floor) to .98 (rings and high bar), whereas in Belgrade Armor's theta ranged from .93 (rings and vault all round finals) to .99 (high bar qualifications and apparatus finals). Finals, but not in apparatus finals, high bar scores have the highest reliability in qualification session and apparatus finals, but only average in all around finals.

METHODS

Sample Entity

Our sample was composed of 176 gymnasts who competed at the WCh in London 2009 qualification event C1. On some apparatus, it was a smaller number of gymnast because it comes to qualifying competition where they compete only by specialists on particular apparatus, so the number of gymnasts on individual apparatus is considerably smaller.

Variables

From official Book of results (Federation Internationale de Gymnastique, 2009a) we made six variables of judges E scores, one D variable and one All variable (final score D+E) from 6 apparatus: floor exercise (FX), pommel horse (PH), rings (RI), vault (VA), parallel bars (PB) and horizontal bar (HB).

Data processing methods

To evaluate all judges scores we used SPSS 17.0 to calculate Descriptive Statistics, Interquartile range, Kolmogorov Smirnov test normality of the variables distributions, Pearson correlations, pair-wise tests between scores of all apparatus. The following reliability and statistics were then calculated: Cronbach's alpha. At the end we did also factor analysis, to define important factors. Five percente level of significance ($p < .05$) was considered for all statistic parameters except Pearson correlation was ($p < .01$).

RESULTS AND DISCUSSION

Mean E and D scores (Table 1) vary between apparatus, and for some judges the data is not normally distributed (e.g. valut and high bar during qualification). There is also a large difference in the variability of scores. In general, the smallest score at competition sessions is observed on pommel horse, and the highest in vault.

TABLE 1
Descriptive statistic and Kolmogorov - Smirnov Test

Scores	Apparatus	n	Range	MIN	MAX	M	SE	SD	Skew.	SE	Kurt.	SE	ICR	KS test	p
E	FX	133	3.175	6.100	9.275	8.217	56.087	646.722	-.780	.210	.544	.417	.938	.796	.551
	PH	131	5.725	3.950	9.675	7.739*	83.478	955.464	-1.054	.212	1.961	.420	1.175	1.164	.133
	RI	126	3.900	5.200	9.100	7.935	59.096	663.350	-1.205	.216	2.128	.428	.800	1.114	.167
	VA	176	2.225	7.375	9.600	8.722**	45.451	602.976	-.648	.183	-.798	.364	1.000	1.766	**.004**
	PB	127	4.400	4.800	9.200	8.074	69.012	777.725	-1.114	.215	1.735	.427	1.000	1.294	.070
	HB	127	5.300	3.550	8.850	7.798	75.099	846.323	-1.570	.215	4.247	.427	1.100	1.627	**.010**
D	FX	133	3.000	3.700	6.700	5.549	54.200	625.060	-.065	.210	-.593	.417	.925	.935	.346
	PH	131	5.900	.800	6.700	5.137*	78.681	903.980	-1.081	.211	3.178	.419	1.000	.957	.319
	RI	126	4.400	2.400	6.800	5.427	80.983	909.034	-.495	.216	.210	.428	1.325	.736	.651
	VA	176	3.400	3.800	7.200	6.132**	53.878	714.778	-1.151	.183	1.192	.364	.400	3.360	**.000**
	PB	127	4.700	2.200	6.900	5.312	77.801	876.772	-.858	.215	1.018	.427	1.200	1.024	.245
	HB	127	5.500	1.700	7.200	5.311	88.561	998.032	-.579	.215	.911	.427	1.125	.767	.598
All	FX	133	5.700	10.075	15.775	13.684	91.241	1.052	-.288	.210	.253	.417	1.313	.457	.985
	PH	131	15.475	.800	16.275	12.819*	161.806	1.859	-2.359	.211	12.977	.419	1.850	1.304	.067
	RI	126	6.850	9.050	15.900	13.363	122.623	1.376	-.622	.216	.427	.428	1.744	.762	.606
	VA	176	5.600	11.200	16.800	14.779**	75.941	1.007	-.646	.183	.734	.364	1.138	.724	.672
	PB	127	7.850	8.100	15.950	13.367	127.460	1.436	-1.068	.215	1.805	.427	1.688	1.118	.164
	HB	127	10.350	5.250	15.600	13.109	131.985	1.487	-1.464	.215	5.462	.427	1.681	1.095	.181

Legend: **E** – Judges score; **D** – Judges; **All –** D & E judges score; **FX** – Floor; **PH** – Pommel horse; **RI** – Rings; **VA** – Vault; **PB** – Parallel bars; **HB** – High bar; *n* – noumber of performances; *M* – Mean; **MIN** – Lowest value; **MAX** – Highest value; **SD** – Standard deviation; **Skew.** – coefficients of skewness; **Kurt.** – coefficients of kurtosis; **ICR** – Interguartile range; **KS test** – Kolmogorov Smirnov test normality of the distribution; *p* – Probability at the level of $p < .05$; * – Minimum mean; ** – Maximum mean.

Kolmogorov Smirnov test showed that only vault and high bar were not normally distributed. Analyzing the judicial commission E and the medium scores on all apparatus, the difference is 0.938 points, with D Commission is 0.965 points and the overall result is different for 1.96 points. The pair-wise test (Table 2) showed significant difference with 12 out of 15 pairs; pairs with PH and RI, PH and HB and RI with HB were significant different. The average D scores on the vault were the highest and the lowest on the pommel horse. Similar results were obtained at OG2008 (Čuk & Atiković, 2009). Pearson's correlations between judges (Table 3) are, in the main, very high. One very low correlation is on the floor apparatus between average score and judge E5 from EGY (*r*: .646; *p* < .01) and judge E4 from (*r*: .817; *p* < .01). The reason for such a huge discrepancy of this the two judges in the assessments should be sought in the fact that they come from countries where the sport artistic gymnastics is not very developed, so we think that their experience in the trial of major competitions such as the World Cup, World Championships and Olympic games is not like the other judges.

Despite all results, indices of reliability are generally quite high. In different sessions and apparatus all reliability measures (Cronbach's alpha - α), are higher than .94. There appears to be no systematic differences in reliability between apparatus. Floor scores tend to have lower reliability than other apparatus in qualification. High bar scores have the highest reliability in qualification session .98. This results are simillar like (Leskošek et al., 2010). Although these results are not directly comparable with results from the 1982 World Cup in Zagreb (Hraski, 1988) it seems that reliability is improving over time, and through the introduction of new rules, especially splitting

TABLE 2
Paired Samples Test

Pair	Apparatus	M	SD	SE	95% Confidence Interval of the Difference		t	df	p
					Lower	Upper			
1	FX - PH	500.954	624.156	54.533	393.068	608.841	9.186	130	**.000**
2	FX - RI	356.151	573.548	51.096	255.026	457.275	6.970	125	**.000**
3	FX - VA	-650.752	669.470	58.050	-765.581	-535.922	-11.210	132	**.000**
4	FX - PB	210.039	622.011	55.195	100.811	319.268	3.805	126	**.000**
5	FX - HB	486.220	650.974	57.765	371.906	600.535	8.417	126	**.000**
6	PH - RI	-103.175	611.948	54.517	-211.070	4.721	-1.893	125	.061
7	PH - VA	-1133.397	901.071	78.727	-1289.149	-977.645	-14.397	130	**.000**
8	PH - PB	-253.937	624.932	55.454	-363.678	-144.196	-4.579	126	**.000**
9	PH - HB	22.244	773.919	68.674	-113.660	158.148	.324	126	.747
10	RI - VA	-949.802	721.564	64.282	-1077.024	-822.579	-14.776	125	**.000**
11	RI - PB	-165.476	574.777	51.205	-266.818	-64.235	-3.232	125	**.002**
12	RI - HB	102.976	652.661	58.144	-12.097	218.050	1.771	125	.079
13	VA - PB	811.220	766.749	68.038	676.575	945.866	11.923	126	**.000**
14	VA - HB	1087.402	891.710	79.127	930.813	1243.991	13.743	126	**.000**
15	PB - HB	276.181	688.294	61.076	155.313	397.049	4.522	126	**.000**

Legend: **FX** – Floor; **PH** – Pommel horse; **RI** – Rings; **VA** – Vault; **PB** – Parallel bars; **HB** – High bar; *M* – Mean; *SE* – Standard error; *SD* – Standard deviation; *t* – Student's *t* distribution; *df* – Degrees of freedom; *p* – Probability.

judges' panel into judges for exercise presentation and exercise content. In Zagreb, only 20 gymnasts competed, all in one session; they were evaluated by 5 judges (head judge and four score judges), which were judging exercise difficulty and exercise presentation together.

With the matrix analysis of the first isolated principal component (Table 5), where the matrix coefficients of the mentioned component (vector correlations of variables with the isolated first principal component) are shown, one can see that all 6 judges have the same and very high correlations with the first principal component. The first principal component is saturated very low on the two apparatus; to the FX with the judge from EGY whose value is the lowest on all apparatus (.710), and the judge from QAT on apparatus RI, who had unlike all other judges, the lowest value (.833). We can conclude that these judges do not partially fit into an average assessment of judges on that apparatus. All principal components which explain different total variabilities (cumulative %) with (FX; 83.8%) of common variance of the entire system, (PH; 92.0%), (RI; 82.4%) which represents the lowest value, (VA; 92.0%), (PB; 90.5%) and the highest value of common variance

of the entire systems being on (HB; 93.8%) were isolated on all apparatus (Table 6). The principal isolated components have values (Total): FX: 5.0, PH: 5.5, RI: 4.9 which represents the lowest value VA: 5.5, PB: 5.4 and the highest value HB: 5.6.

The tendency, which appears in the evaluation exercise, is the enlargement of the result of objective factor and the reduction of the result of subjective (human) factor. In addition to various attempts of qualitative improvements, the subjective tone remains necessarily present, sometimes only as a result of different perspectives on a performance, and not so rarely as a result of the subjective interests and prestige. Objectification of the contestant's performance evaluation is increasingly a problem, not only because the effort which increases their results is much bigger than in the past, but also because the differences among the contestants are minimal and thus the ability for a mistake to be made is much higher and the consequences worse.

CONCLUSION

In the analyzed results presented in this paper, vault in comparison with other apparatus have the lowest

deductions of the judging commission E, which evaluates the performance of the jump itself. The average gain on all other apparatus differs from the average for 0.797 points. Average initial assessment given by the judging commission D on the vault, tells us again about the need to revise the current Code of Points (Federation Internationale de Gymnastique, 2009b) because vault has much higher initial assessment than all other apparatus, for 0.784 points, while the total sum of all assessments and value is even higher, amounting to 1.510 points the assessments value on the vault should soon be as equal in the results of all-round which is currently not so. With the 2009 Code of Points, all results for the six ap-

TABLE 3

Pearson correlation coefficients between judges' E scores with average score of 6 judegs

Apparatus	E1	E2	E3	E4	E5	E6
FX E_N133	JPN	VEN	LUX	ROU	EGY	ITA
Average score	.969*	.931*	.974*	.944*	.646*	.952*
PH E_N131	SLO	RUS	PUR	BRA	PRK	DEN
Average score	.963*	.967*	.961*	.973*	.939*	.947*
RI E_N126	BUL	FRA	GRE	QAT	JOR	RSA
Average score	.935*	.954*	.931*	.817*	.897*	.890*
VA E_N176	MEX	NZL	BLR	GER	CAN	ISL
Average score	.944*	.973*	.922*	.978*	.971*	.960*
PB E_N127	NED	KOR	LTU	ARG	CZE	POL
Average score	.956*	.933*	.963*	.942*	.943*	.963*
HB E_N127	ALG	POR	AUT	UKR	HUN	GBR
Average score	.969*	.965*	.970*	.966*	.966*	.971*

Legend: **E** – Judges score; **FX** – Floor; **PH** – Pommel horse; **RI** – Rings; **VA** – Vault; **PB** – Parallel bars; **HB** – High bar; *N* – Number of performances; **JPN** – Japan; **VEN** – Bolivarian Republic of Venezuela; **LUX** – Grand Duchy of Luxembourg; **ROU** – Romania; **EGY** – Arab Republic of Egypt; **ITA** – Italian Republic; **SLO** – Republic of Slovenia; **RUS** – Russian Federation; **PUR** – Puorto Rico; **BRA** – Federative Republic of Brazil; **PRK** – Democratic Pepole's Republic of Korea; **DEN** – Kingdom of Denmark; **BGR** – Republic of Bulgaria; **FRA** – French Republic; **GRC** – Hellenic Republic; **QAT** – State of Qatar; **JOR** – Hashemite Kingdom of Jordan; **RSA** – Republic of South Africa; **MEX** – United Mexican States; **NZL** – New Zealand; **BLR** – Republic of Belarus; **GER** – Federal Republic og Geramny; **CAN** – Canada; **ISL** – Republic of Iceland; **NED** – Kingdom of the Netherlands; **KOR** – Republic of Korea; **LTU** – Republc of Lithuania; **ARG** – Argentine Republic; **CZE** – Czech Republic; **POL** – Republic of Poland; **ALG** – People's Demotraic Republic of Algeria; **POR** – Portugalese Republic; **AUT** – Republic of Austra; **UKR** – Ukraine; **HUN** – Republic of Hungray; **GBR** – United Kingdom of Great Britan and Northern Ireland; ***** – Correlation is significant at the (*p* < .01) level.

TABLE 4

Reliability of judge's E scores

Case Processing Summary	FX		PH		RI		VA		PB		HB	
Reliability Statistics	ȧ	*n*	ȧ	*n*	ȧ	*n*	ȧ	*n*	ȧ	*n*	ȧ	*n*
	.947	6	.982	6	.956	6	.982	6	.978	6	.987	6

Legend: **FX** – Floor; **PH** – Pommel horse; **RI** – Rings; **VA** – Vault; **PB** – Parallel bars; **HB** – High bar; *n* – Number of items; ȧ – Cronbach's index of internal consistency.

TABLE 5

Component matrix E judges

Judges	FX	Component 1	PH	Component 1	RI	Component 1	VA	Component 1	PB	Component 1	HB	Component 1
E1	JPN	.963	SLO	.962	BUL	.935	MEX	.948	NED	.957	ALG	.970
E2	VEN	.930	RUS	.966	FRA	.953	NZL	.971	KOR	.939	POR	.960
E3	LUX	.970	PUR	.961	GRE	.933	BLR	.930	LTU	.961	AUT	.971
E4	ROU	.943	BRA	.971	QAT	.833	GER	.977	ARG	.947	UKR	.965
E5	EGY	.710	PRK	.944	JOR	.902	CAN	.970	CZE	.946	HUN	.968
E6	ITA	.950	DEN	.951	RSA	.888	ISL	.960	POL	.960	GBR	.972
Average E1 - E6		.911		.959		.907		.959		.952		.969

Legend: **E** – Judges score; **FX** – Floor; **PH** – Pommel horse; **RI** – Rings; **VA** – Vault;
PB – Parallel bars; **HB** – High bar; **JPN** – Japan; **VEN** – Bolivarian Republic of Ve-
nezuela; **LUX** – Grand Duchy of Luxembourg; **ROU** – Romania; **EGY** – Arab Repu-
blic of Egypt; **ITA** – Italian Republic; **SLO** – Republic of Slovenia; **RUS** – Russian
Federation; **PUR** – Puerto Rico; **BRA** – Federative Republic of Brazil; **PRK** – De-
mocratic Pepole's Republic of Korea; **DEN** – Kingdom of Denmark; **BGR** – Repu-
blic of Bulgaria; **FRA** – French Republic; **GRC** – Hellenic Republic; **QAT** – State of
Qatar; **JOR** – Hashemite Kingdom of Jordan; **RSA** – Republic of South Africa;
MEX – United Mexican States; **NZL** – New Zealand; **BLR** – Republic of Belarus;
GER – Federal Republic og Geramny; **CAN** – Canada; **ISL** – Republic of Iceland;
NED – Kingdom of the Netherlands; **KOR** – Republic of Korea; **LTU** – Republc
of Lithuania; **ARG** – Argentine Republic; **CZE** – Czech Republic; **POL** – Republic
of Poland; **ALG** – People's Demotraic Republic of Algeria; **POR** – Portugalese Repu-
blic; **AUT** – Republic of Austra; **UKR** – Ukraine; **HUN** – Republic of Hungray;
GBR – United Kingdom of Great Britan and Northern Ireland.

TABLE 6

Matrix of characteristic roots and total variance explained

Apparatus	Component	Total Variance Explained					
		Initial Eigenvalues			Extraction Sums of Squared Loadings		
		Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
FX	1	5.030	83.836	83.836	5.030	83.836	83.836
PH	1	5.522	92.036	92.036	5.522	92.036	92.036
RI	1	4.950	82.497	82.497	4.950	82.497	82.497
VA	1	5.524	92.060	92.060	5.524	92.060	92.060
PB	1	5.433	90.546	90.546	5.433	90.546	90.546
HB	1	5.629	93.822	93.822	5.629	93.822	93.822

Legend: **FX** – Floor; **PH** – Pommel horse; **RI** – Rings; **VA** – Vault; **PB** – Parallel bars;
HB – High bar.

paratus are not equal to obtain D, E and final score (D + E). Reliability is generally quite higher on this WCh 2009 and ranged from .947 to .987. With the help of factor analysis and the matrix analysis of the first isolated principal component, it has been established that the two judges partially fit into an average assessment of all judges on the ground and circles. Computerized system, suggested by authors (Bučar Pajek, Forbes, Pajek, Leskošek, & Čuk, 2011), on next competitions would be good to overcome significant differences in E judge's scores

Coaches can use results from this research for planing of preparation tactics of gymnasts for all round, team and apparatus competition...

REFERENCES

Ansorge, C. J., Scheer, J. K., Laub, J., & Howard, J. (1978). Bias in Judging Women's Gymnastics Induced by Expectations of Within-Team Order. *Research Quarterly for Exercise & Sport*, *49*(4),
399–405.

Ansorge, J. Ch., & Scheer, K. J. (1988). International bias detected in judging gymnastic competition at the 1984 Olympic games. *Research Quarterly for Exercise & Sport*, *59*(2), 103–107.

Bučar Pajek, M., Forbes, W., Pajek, J., Leskošek, B., & Čuk, I. (2011). Reliability of Real Time Judging System (RTJS). *Science of Gymnastics Journal*, *3*(2), 47–54. Retrieved from http://www.fsp.uni-lj.si/mma_bin.php?id=20110602094248.

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. Retrived

from http://www.sportspa.com.ba/download2009.html

Čuk, I., & Forbes, W. (2006). Kam greš, sojenje? [Where do the trial?]. In E. Kolar and S. Piletič (Eds.), Gymnastics for coaches and teachers 2 (pp. 76–86). Ljubljana, Slovenia: Gymnastics Federation of Slovenia.

Čuk, I., & Forbes, W. (2010). How apparatus difficulty scores affect all around results in men's artistic gymnastics. *Science of Gymnastics Journal*, *2*(3), 57–63. Retrieved from http://www.fsp.uni-lj.si/mma_bin.php?id=2010093010022523

Federation Internationale de Gymnastique (2006). *Code of Points for Men Artistic Gymnastics Competitions*. London, GB: FIG.

Federation Internationale de Gymnastique (2009a). *Artistics gymnastics book of results*. London, GB: FIG.

Federation Internationale de Gymnastique (2009b). *Code of Points for Men Artistic Gymnastics Competitions*. London, GB: FIG.

Hraski, Ž. (1988). Valorizacija suđenja u muškoj sportskoj gimnastici [Valuation of the judging in men's gymnastics]. *Kineziologija*, *20*(2), 143–152.

Leskošek, B., Čuk, I., Karácsony, I., Pajek, J., & Bučar, M. (2010). Reliability and validity of judging in mens artistic gymnastics at the 2009 University games. *Science of Gymnastics Journal*, *2*(1), 25–34. Retrieved from http:// www.fsp.uni-lj.si/mma_bin.php?id=2010020920531429

Popović, R. (2000). International Bias Detected in Judging Rhythmic Gymnastics Competition at Sydney 2000 Olympic Games. *Facta Universitatis 1*(7), 1–13.

Sands, B. (2011). Judging in "real time". Retrieved from http://gymnasticscoaching.com/new/2010/11/bills-sands-judging-in-real-time/.

Correspodence to:
Almir Atiković, PhD
Faculty of Physical Education and Sport
2. Oktobra 1
75000 Tuzla
Bosnia and Herzegovina
E-mail: almir.atikovic@untz.ba
Phone: 00387 61 830 730
Fax: 00387 35 278 536