WHY THE VAULT BECAME SUPERIOR TO OTHER EVENTS IN WOMEN'S ARTISTIC GYMNASTICS AT THE OLYMPICS?

Jun He[1], Jeffrey Montez de Oca[2] and Lei Zhang[2]

[1] School of Sports Science and Physical Education, Nanjing Normal University, China
[2] Department of Sociology, University of Colorado Colorado Springs, USA

Abstract

Critics argue that disciplines in women's artistic gymnastics are not equal and the vault is generally scored much higher than the uneven bars, balance beam and floor exercise. The aim of this study is to understand why the vault became superior to other women's events. The data are the official results for the 586 women gymnasts in Qualification at Olympics from 2000 to 2020. The One-Way ANOVA was used to analyze the variance of D-scores, E-scores and F-scores for women gymnasts obtained on each apparatus. Our research shows that disciplines in women's artistic gymnastics have not been equal for gymnasts when trying to obtain high F-scores in the past 6 Olympics. Among the four women's events, the vault came to be the one on which gymnasts are more likely to obtain high F-scores after the 2000 Olympics. We indicate that the strength of vault resulted from the introduction of the new vaulting table in 2001 and the new open-ended rules in 2006. Although the two big changes implemented by the International Federation of Gymnastics in the beginning of the new millennium were aimed at improving safety of the vault and fairness of judging, the interplay of the two big changes unintentionally promoted the vault to become the most powerful event in women's artistic gymnastics. Such unanticipated consequence of purposeful action may constitute the most important element (i.e., imbalance of disciplines) in the sport. Accordingly, this research has the potential to shed new light on not only this important topic of equality between disciplines, but also broader trends in modern artistic gymnastics.

Keywords: artistic gymnastics, vault, discipline, equality, Olympics.

INTRODUCTION

Women's artistic gymnastics (WAG) made its first appearance at the 1928 Olympic Games in Amsterdam. In 1933, the women's technical committee (WTC) was founded and governs the development of WAG (Pajek, 2018). The Code of Points (CoP) is utilized by the WTC to encourage particular movements or styles, including acrobatics and artistry. During the 1950s and 1960s, the CoP was centered around artistry and enlivened by ballet. Consequently, in that period, achievements in the sport favored those who brought elegance to exercises (Atiković, Kalinski, & Čuk, 2017; Kerr, 2006). Since then, however, women's artistic gymnastics evolved from this balletic tradition into a more acrobatic sport. As a result, in the 1970s, this sport saw an abundance of young athletes performing high level acrobatics at the expense of elegance and

artistry (Barker-Ruchti, 2009; Cervin, 2015; Kerr, Barker-Ruchti, Nunomura, Cervin, & Schubring, 2018). Since 1996, gymnasts have no longer performed compulsory routines on each apparatus and instead only competed in optional routines (Pajek, 2018). In 2006, the International Federation of Gymnastics (FIG) removed the "perfect 10" as a maximum score, and replaced it with the open-ended scoring system, in which difficulty scores no longer have a ceiling. In the open-ended format, the total score consists of two scores: the requirements and the difficulty of a routine (D-score) plus execution (E-score). Since 2006, updates to the scoring system in each Olympic cycle have pushed the sport in the acrobatic direction (Cervin, 2015). This trend, however, has been controlled to some degree as the FIG gave more weight to the performance quality of a routine after Beijing 2008. The key aspects are both greater deductions for errors and the scoring of artistry introduced into the women's CoP (on both the balance beam and the floor exercise) which means that the performance quality significantly affects the results in the finals (He, Montez de Oca, & Zhang, 2020).

While the weight between 'difficulty' and 'artistry' has been adjusted so that the new judging system can work well, another issue emerged. This issue is with the equality between disciplines, as disciplines of artistic gymnastics are quite different from each other in terms of skills, compositional requirements, connection bonus, specific penalties, etc. Bučar, Čuk, Pajek, Karacsony, and Leskošek (2012) indicated that the vault and the floor exercise finals were sessions with the largest scope between high and low E-scores, but they also noted that more inspections should be made on this issue in future analyses of judging. Further research studies identified that the vault, compared to other WAG disciplines, is one on which deductions in execution tend to be minor (Kalinski, Atiković, Jelaska, & Milić, 2016; Kalinski, Jelaska, & Atiković, 2017; Kalinski, Padulo, Atiković, Milić, &

Jelaska, 2016), and one on which F-scores are significantly different from those on other apparatuses in WAG all-around qualification results (Atiković et al., 2020; Massidda & Calò, 2012).

Despite the aforementioned research that focused on the most important apparatus for WAG all around success, there is a paucity of literature that explains why the vault is scored much higher than other WAG events. Consequently, the aim of this research is to understand how the vault has become superior to other disciplines in WAG. We analyze the question by focusing on two major changes that are unprecedented in the history of artistic gymnastics: one is the replacement of the vaulting horse by the vaulting table in 2001, and the other is the introduction of open-ended rules in 2006 (Naundorf, Brehmer, Knoll, Bronst, & Wagner, 2008). We believe that the interplay of these two major forces led to the superiority of the vault over other WAG disciplines. By focusing on the perspective of key changes within this sport, we believe that this research has the potential to shed a new light on not only this important topic, but also broader trends in modern artistic gymnastics.

METHODS

Since gymnasts strive to achieve the best performances and highest scores at the Olympic Games (OG) (as cited in Čuk & Atiković, 2009; Kalinski, Atiković, et al., 2016), we sampled all women gymnasts who have participated in the Olympic Games since 2000, as that was when competitors started to perform only optional routines (i.e., OG2000, OG2004, OG2008, OG2012, OG2016, OG2020). There were 98 women gymnasts at each Olympics, except at OG2000 and OG2012 (97 women gymnasts). Altogether, 586 women gymnasts were analyzed from the past 6 Olympics. Further, WAG competitions at the Olympic Games include four sessions: Qualification (C-I), All-around Final (C-II),

Apparatus Finals (C-III) and Team Final (C-IV), which are held on different days. We analyzed performance scores in Qualification since this session can be considered the most important; it is the round where both teams and individual gymnasts compete to qualify for the finals (Kalinski, Atiković, et al., 2016). In our analysis, we used scores on the first vault since generally most competitors completed only one vault in world competitions (Kalinski, Atiković, et al., 2016).

We analyzed the D-score, E-score, and F-score for exercises performed by women gymnasts on all events (i.e., vault, uneven bars, balance beam and floor exercise) in Qualifications from OG2000 to OG2020. There were only F-scores at the OG2000 and OG2004 because exercises were judged under the traditional "perfect 10" system. Additionally, we dropped cases with "0" points from the statistical analysis. We retrieved these scores from the results book available at the Gymnastics Results website (https://gymnasticsresults.com/).

First, we calculated means and standard deviations of D-, E-, and F-scores. We presented our results in the "mean $\pm$ standard deviation" manner in Table 1. Then, we conducted the One-Way ANOVA to detect the mean difference of scores across four apparatuses, i.e., the vault (VT), uneven bars (UB), balance beam (BB), and floor exercise (FX). For the WAG Qualifications at the OG2000 and 2004, the means of F-scores on these apparatuses were analyzed. For data after 2004, we applied this analytical strategy to D-, E-, and F-scores. Each ANOVA with a statistically significant (p<0.05) F ratio indicated that at least one apparatus had a different mean score. We computed the eta square (η^2) as the overall effect size of the mean difference across four apparatuses. To further demonstrate which apparatus had a mean that was different from another three, we used the Levene's test result to guide our choice of the suitable multiple comparison test statistic. When the Levene's test was not statistically significant, we conducted

multiple mean comparisons by using the Tukey's B statistic. When the Levene test returned a statistically significant result, we performed the analysis again with the Games-Howell test. For each significant mean difference identified by either Tukey's B or Games-Howell statistic, we further calculated the Cohen's d value to assess the effect size. To follow the convention, we used eta square values of 0.01, 0.06, and 0.14 and d values of 0.20, 0.50, and 0.80 as thresholds of small, medium, and large effects, respectively (Cohen, 1988). All statistics was obtained from the IBM SPSS Statistics 22.0.

RESULTS

Differences in F-score means of WAG disciplines from OG2000 to OG2020

Table 1 shows the One-Way ANOVA with post hoc comparisons in F-score means, D-score means and E-score means for the past six Olympics (i.e. from OG2000 to OG2020). As far as the F-score mean is concerned, at OG2000, which was the first time compulsory routines were no longer required, there were no statistical differences of F-score mean between WAG events, although the F-score mean on both balance beam and particularly uneven bars looked a little bit higher than the vault and the floor exercise. The following Olympics in 2004 were not only the last time that competitions at the Olympics were judged under the traditional "perfect 10" system, but also the first time that the traditional vaulting horse was replaced by the new vaulting table. It can be seen that at the OG2004, the vault stands out from all WAG events and shows significantly higher results in terms of F-score mean than both the balance beam and the floor exercise, and is equal to the uneven bars. With the Cohen's d value, we know that the effect size of 0.83 (the mean difference between the vault and the balance beam in standard deviation units) and 0.50 (the mean difference between the vault and the floor exercise in standard deviation units) are

considered to be large and medium respectively.

Four years later at the OG2008, when the "perfect 10" scoring system gave way to the "open ended" system, however, significant differences only emerged between the vault and the floor exercise, with the medium effect size of 0.68. But it can be seen that at the OG2008, the F-score mean for the vault was quite a bit higher than for the other three WAG events vis-à-vis the previous Olympics (i.e., OG2000 and OG2004). At the OG2012, the vault pulled away in terms of the F-score mean, extending its lead to around 0.6 points, i.e., there were significant differences between the vault and the other three events (effect sizes of 0.56, 0.86 and 0.80 are considered to be medium and large respectively). Meanwhile, no significant differences existed between the uneven bars, the balance beam and the floor exercise at the 2012 Olympics. At OG2016 and OG2020, the F-score mean on the vault was significantly higher than on the other three WAG events, particularly at the OG2020 where the F-score mean for the vault was around 1.0 point higher than for other WAG events, and the effect sizes of 0.80, 1.36 and 1.47 were all large. Additionally, it should be noted that at the OG2016, the F-score mean for the uneven bars was significantly higher than both for the balance beam and the floor exercise, but not so at the OG2020. It can be concluded that basically the four WAG disciplines have not been equal for gymnasts in terms of obtaining high F-scores in the past six Olympics (eta square values for F-scores of all events were considered to be medium at OG2004 through to OG2012 and large at both OG2016 and OG2020). Among them, the vault came to be the one on which WAG gymnasts are more likely to obtain high F-scores after OG2000 (especially at the Olympics in 2012 and beyond). Although sometimes the uneven bars showed the momentum to rise in terms of the F-score mean, it lagged behind the vault most of the time. In the following section, we further examine the components of F-score (i.e., D-score and E-score) from OG2008 through to OG2020.

Differences in D-score means for WAG disciplines from OG2008 to OG2020

We investigated the D-score means (as well as E-score in the following section) for WAG disciplines since OG2008 where the F-score began to be calculated by using D-score and E-score under the new open-ended system. At OG2008, the D-score means on both the uneven bars and the balance beam are significantly higher than those on the vault and the floor exercise, and further scores in the floor exercise are significantly higher than on the vault. Consequently, the D-score mean for the vault is the lowest in comparison with the other three WAG events at the OG2008, which is similar to the maximum D-score differences between WAG disciplines in 2008 (Figure 1). At OG2012, the D-score means on all apparatuses but vault decreased mainly because the difficulties counted in routines on the uneven bars, the balance beam and the floor exercise decreased from 10 at OG2008 (FIG, 2006) to 8 at OG2012 (FIG, 2009), so that the D-score means for all WAG events approximated each other (only the uneven bars event was significantly different from the floor exercise). At OG2016, the D-score means on all WAG events increased only a little in comparison to those at the OG2012, with significant differences between the uneven bars, the vault and the floor exercise. At OG2020, despite a sharp fall of D-score means for all WAG events compared to OG2016 (especially the vault on which the maximum D-score was 6.0, which is much less than four years earlier), both the uneven bars and the balance beam scores were significantly higher than those on the vault and the floor exercise. In short, it is obvious that the maximum D-scores are different for each event from OG2008 to OG2020. Coupled with the differences for D-score means between WAG events, we can conclude that although the difficulty

value for elements on each apparatus varied at different Olympics due to the updates to the CoP, WAG disciplines are literally not equal: eta square values for D-scores for all events were considered to be small at both the OG2008 and the OG2012 and medium at the OG2016 and the OG2020. Gymnasts are more likely to achieve high D-scores on both the uneven bars and the balance beam relative to the vault and the floor exercise, although such chances seemed to decrease on the whole as Cohen's d value for each significant D-score mean difference showed a downward trend after OG2008.

Differences in E-score means for WAG disciplines from OG2008 to OG2020

When we refer to E-score means, the situation seems pretty much the reverse of how the D-score means work. At the OG2008, the E-score means for both the vault and the floor exercise are significantly higher than those for the uneven bars and the balance beam. Particularly the vault stands out where the E-score mean was 9.150 points with small standard deviation. Additionally, the maximum vault E-score reached 9.650 points (Figure 2). Such numbers appeared to large extent also at the following Olympic Games from 2012 to 2020 (eta square values for E-scores for all events were considered large from OG2008 to OG2020), although the E-score means and the maximum E-score for all WAG events dropped by different degrees after the OG2008, mainly due to stricter deductions for execution and artistry: e.g., apart from a deduction for a fall that was increased from 0.8 in 2006 to 1.0 in 2009, the FIG added the evaluation of artistry on both the balance beam and the floor exercise to the new scoring system in the 2009 CoP (FIG, 2009), and refined it afterwards in the Rio (FIG, 2015) and Tokyo cycle (FIG, 2017). Additionally, the vault rules of landing within the corridor down the center of the landing mat was introduced to the CoP in 2009, while on the uneven bars, more specific deductions for E-panel, like the angle of completion of

elements, were implemented in the London Olympic cycle (FIG, 2009). Regardless of the updated rules regarding execution in different WAG events, the vault has consistently remained the apparatus where the deductions for execution are the lowest among WAG events (the E-score means were almost above 8.8 points at the OG2008 and beyond) and had a small standard deviation (around 0.4 points from OG2008 to OG2020), which is congruent with the results of previous research (Atiković, Kalinski, Bijelić, & Vukadinović, 2011; Kalinski, Atiković, et al., 2016; Kalinski et al., 2017; Kalinski, Padulo, et al., 2016). The other three events typically score lower on E-score than the vault. This can be demonstrated by Cohen's d value for each significant E-score mean difference between the vault and the other three WAG events at the Olympics from 2008 to 2020 (Table 1), nearly all of which exceed one and a quarter standard deviation units, representing substantial effect sizes.

Moreover, we noted that among the four WAG events, the balance beam is the event on which women gymnasts are more challenged to perform with perfection, probably due to the anxiety caused by the risk of injury (Kolt & Kirkby, 1994; Sands, 2000), the instability of balance, and deductions for artistry (Kalinski, Padulo, et al., 2016). This is especially striking at the OG2020 where the E-score mean and the maximum E-score on the balance beam is less than 7.3 and 8.4 points respectively, which was the lowest since OG2008 when the new open-ended scoring system made its Olympic debut. Nevertheless, it should be noted that although the difficulty value (both the D-score mean and the maximum D-score) of the balance beam increased at OG2016 in comparison to OG2012, the E-score mean for the balance beam did not decrease correspondingly at the OG2016 (with fewer standard deviations relative to OG2012), although the maximum E-score decreased a little. This result presented a challenge to Kalinski (2017) who argued that the higher difficulty values on the

balance beam would lead to a lower score for execution, based on the assumption that the performance of more complex and difficult elements was more challenging. This suggests that further research on whether or not routines that comprise more complex and difficult elements lead to lower E-scores and consequently F-scores should be conducted.

Table 1
ANOVA results with post hoc comparisons.

Games	Contents	VT	UB	BB	FX	F	Eta square	Levene statistic	Sig.mean difference (Cohen's d)
OG 2000	N	83	84	81	83				
	F-score	9.197 ±0.32	9.329 ±0.51	9.267 ±0.42	9.126 ±0.54	2.57			
OG 2004	N	84	85	85	83				
	F-score	9.182 ±0.22	9.105 ±0.53	8.853 ±0.57	9.004 ±0.50	9.82 ***	0.06	16.41 ***	1-3(0.83), 1-4(0.50), 2-3(0.46)
OG 2008	N	82	84	84	82				
	D-score	5.461 ±0.45	6.189 ±0.66	6.167 ±0.47	5.696 ±0.42	43.52 ***	0.27	4.61 **	2-1(1.31), 2-4(0.91), 3-1(1.54), 3-4(1.07), 4-1(0.54)
	E-score	9.150 ±0.33	8.124 ±0.64	8.260 ±0.66	8.492 ±0.42	91.82 ***	0.35	12.22 ***	1-2(2.11), 1-3(1.79), 1-4(1.76), 4-2(0.70), 4-3(0.43)
	F-score	14.606 ±0.65	14.283 ±1.07	14.426 ±0.94	14.123 ±0.77	6.53 ***	0.04	4.07 **	1-4(0.68)
OG 2012	N	80	78	83	82				
	D-score	5.463 ±0.57	5.685 ±0.71	5.553 ±0.59	5.391 ±0.45	3.59 *	0.03	3.54 *	2-4(0.51)
	E-score	8.647 ±0.47	7.808 ±0.82	7.618 ±0.96	8.070 ±0.51	41.53 ***	0.22	13.58 ***	1-2(1.30), 1-3(1.43), 1-4(1.18), 4-3(0.62)
	F-score	14.092 ±0.85	13.492 ±1.31	13.167 ±1.29	13.405 ±0.88	13.46 ***	0.08	7.17 ***	1-2(0.56), 1-3(0.86), 1-4(0.80)
OG 2016	N	81	79	82	82				
	D-score	5.512 ±0.52	5.816 ±0.61	5.670 ±0.52	5.444 ±0.67	6.52 ***	0.05	0.99	2-1(0.54), 2-4(0.58)
	E-score	8.862 ±0.42	8.146 ±0.63	7.801 ±0.72	7.970 ±0.51	74.49 ***	0.33	11.69 ***	1-2(1.37), 1-3(1.87), 1-4(1.94), 2-3(0.51)

	F-score	14.356 ±0.72	13.963 ±1.12	13.464 ±1.11	13.297 ±1.39	19.61 ***	0.12	2.66 *	1-2(0.43), 1-3(0.97), 1-4(1.00), 2-3(0.45), 2-4(0.53)
	N	84	88	91	85				
OG 2020	D-score	5.055 ±0.55	5.350 ±0.85	5.279 ±0.53	5.034 ±0.51	5.72 **	0.04	5.86 **	2-1(0.43), 2-4(0.46), 3-1(0.42), 3-4(0.47)
	E-score	8.829 ±0.33	7.593 ±0.85	7.283 ±0.86	7.706 ±0.51	171.08 ***	0.42	20.61 ***	1-2(2.07), 1-3(2.54), 1-4(2.67), 4-3(0.62)
	F-score	13.863 ±0.72	12.869 ±1.74	12.559 ±1.18	12.695 ±0.86	43.08 ***	0.15	9.92 ***	1-2(0.80), 1-3(1.36), 1-4(1.47)

*Note: VT=Vault, UB=Uneven bars, BB=Balance beam, FX=floor exercise; * P<0.05, ** P<0.01, *** P<0.001;*

In Multiple Comparisons, 1=Vault; 2=Uneven bars; 3=Balance beam; 4=floor exercise; "1-3" shows that the mean of scores on vault is significantly higher than that of balance beam at 0.05 level, etc;

We carried out Games-Howell test for all the D-, E-, and F-scores of the four WAG events from OG2004 to OG2020, except the D-scores at OG2016 for which Tukey's B is used;
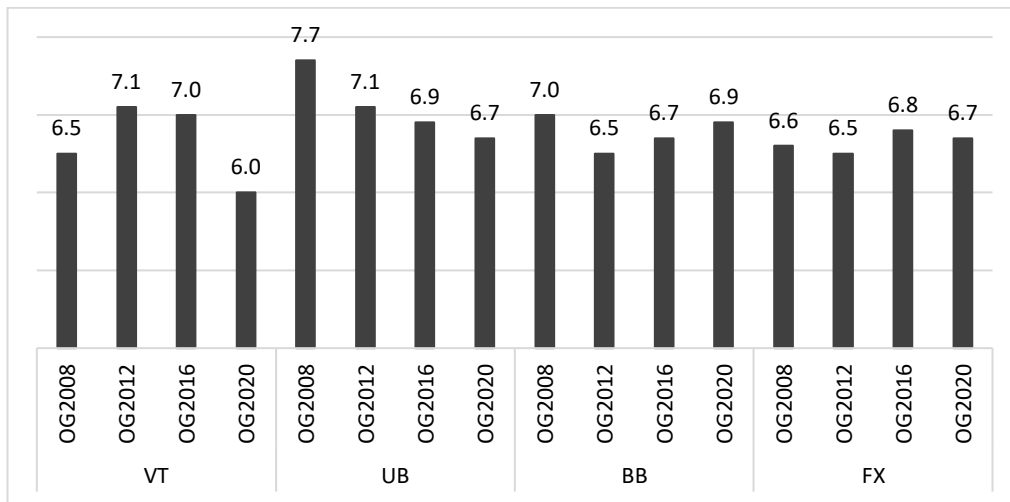

Figure 1. Maximum D-score on each apparatus in WAG Qualification at Olympics 2008-2020.
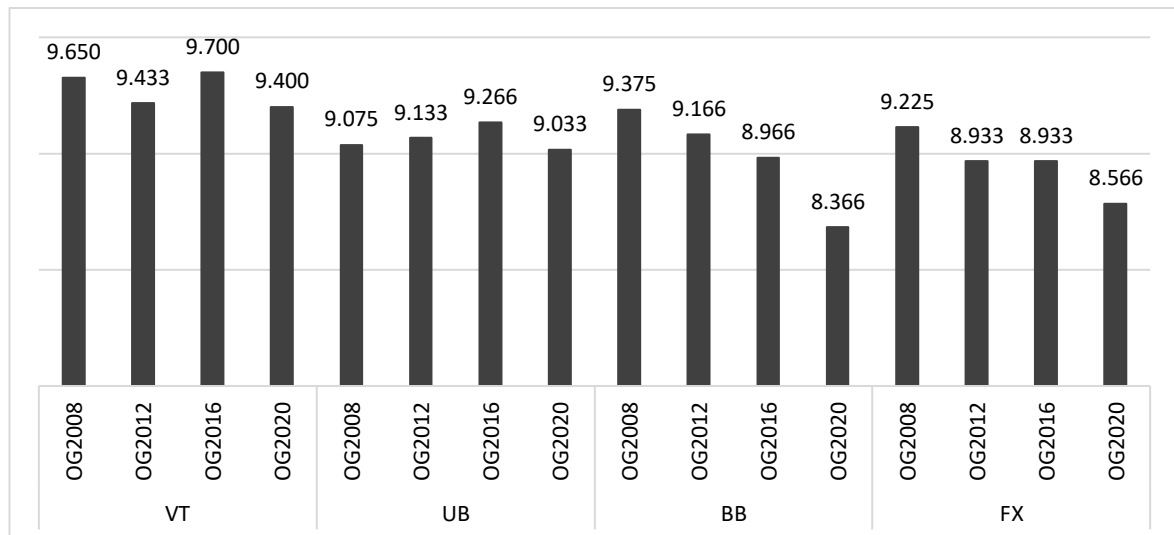
Figure 2. Maximum E-score on each apparatus in WAG Qualification at Olympics 2008-2020.

DISCUSSION

The aim of this study is to investigate why the vault has stood out since OG2004 relative to the uneven bars, balance beam and floor exercise in WAG. The vault includes only one element, which is distinct from the uneven bars, the balance beam and the floor exercise where the content (D-score) includes the highest 8 difficulties, compositional requirements, connection values and a bonus. Scholars argue that compared to the uneven bars, balance beam and floor exercise that require multiple skills and combinations, gymnasts can obtain higher scores on the vault since it is judged on a single skill alone (Atiković et al., 2011; Atiković et al., 2020). Such proposition is, however, contradicted by the results at the OG2000, where the superiority of the vault over other WAG events in terms of high scores barely existed. It can be seen that at the OG2000 the F-score mean for the vault is a little lower than for the uneven bars and balance beam, although no significant differences existed between these events. This result can confirm the point proposed by Kalinski, Padulo, et al. (2016) that the vault is a highly demanding apparatus and can never be considered easy simply because it includes only one element (Čuk & Atiković, 2009).

At the following Olympics in 2004, where the competition ran under the same "perfect 10" scoring system as in 2000, however, the vault distinguished itself with F-score means significantly higher than both on the balance beam and the floor exercise. The question, hence, is why there is a significantly higher F-score mean for the vault than for the other WAG events except on uneven bars four years later? As we know, among many factors, modifying an apparatus will promote new elements and development of routines on the apparatus. For example, modifications endorsed by the FIG to the apparatuses in the 1970s, such as wider-set bars, carpeted springboards, padded beam, and "double sprung" floor enabled gymnasts to perform new moves on the bars, somersault on the beam, and tumble to greater heights in the floor exercise in the following decades (Cervin, 2015). In this sense, we believe that the introduction of the new vaulting table in 2001 played a crucial role.

The replacement of the traditional vaulting horse with the vaulting table in 2001 by the FIG represents the largest single change in artistic gymnastics in recent times (Irwin & Kerwin, 2009). Although the new table was introduced by the FIG for safety (Kalinski et al., 2017) and is identical to the traditional horse in height, the design and construction of the new table

enabled a higher run-up velocity (Milčić, Živčić, & Krističević, 2019; Schärer, Lehmann, Naundorf, Taube, & Hübner, 2019), facilitated more anatomically functional position of the arms during the support phase, which allows for a more effective transfer of horizontal kinetic energy to vertical and angular one for the second flight phase (Čuk & Ferkolj, 2012). Moreover, scholars indicate that the larger and flatter surface area of the new table have not only perceptually and materially changed the vaulting technique in handspring forward vault (Irwin & Kerwin, 2009), but also made Yurchenko (round-off) and Tsukahara (handspring with ¼ - ½ turn) much easier to complete (Sands & McNeal, 2002). Consequently, we can conclude that the new apparatus has influenced performance on the vault (Jackson, 2010), although it remains to be determined if the new table has facilitated more advanced vaults (Irwin & Kerwin, 2009). Therefore, we believe that the introduction of the new vaulting table by the FIG in 2001 with the aim of improving safety has also increased the probability of getting higher scores on the vault for women gymnasts since OG2004.

In 2006, the FIG reworked the scoring system in response to the issue of fairness under the traditional "perfect 10" judging system (Kerr & Obel, 2015). Critics have since argued that the open-ended scoring system has led to a loss of artistry as gymnasts achieved top scores with acrobatic movements (Kerr & Obel, 2015), since under the new rules the sport saw a great increase in difficulty levels from 2006 on (Leskošek, Čuk, & Pajek, 2013; Thornton, 2010). To prevent the trend of "sacrificing execution for difficulty" (Leskošek et al., 2013), the FIG has been weighting performance quality in the open-ended CoP since 2009. Particularly in WAG, as mentioned above, apart from the highest difficulties in a routine (except on the vault) that were reduced to 8 and a fall deducted by 1.0, the FIG added the evaluation of artistry on both the balance beam and the floor exercise to the new scoring system in the 2009 CoP, and refined it in the Rio and Tokyo cycles. As a result, the sport has seen a shift from difficulty to elegance and artistry since 2009, which is a drastic change at the institutional level (He et al., 2020).

The new open-ended scoring system used to modify the balance between artistry and difficulty in scoring coupled with the adoption of the vaulting table intended to promote safety but unintentionally it increased the superiority of the vault over other WAG events. Although significant differences only existed between the vault and the floor exercise chiefly owing to much lower D-scores for the vault versus other three WAG events (particularly in comparison with uneven bars and balance beam) at the OG2008, the vault came to be dominant in terms of F-score mean among the WAG events ever since the 2012 Olympics. We believe that there are two reasons for this: one is the D-score means for all events but the vault dropped sharply at the OG2012 because the number of difficulties counted in a routine decreased from 10 to 8, which rendered all WAG events more even in terms of difficulty value. Also, stricter deductions for execution and artistry carried out in the open-ended scoring system after OG2008, when coupled with the performance advantages of the new vault table, made the vault score higher than other events in WAG.

It can be said that the interplay of the two big changes brought about what the sociologist Robert Merton would call "unanticipated consequences" (Merton, 1936). The FIG worked to introduce a new vaulting table, ostensibly for safety reasons, but this equipment served to encourage greater height and force generated in the event, thus making the vault superior over the uneven bars, balance beam and floor exercise in the light of the "perfect performance" gymnasts can present. Furthermore, when the ceiling for the "perfect 10" was removed in 2006, the

"perfect performance" the gymnast can assume on the vault was maximized by the new scoring philosophy, particularly when greater weight was attached to execution and artistry after OG2008. That is, where the FIG intended to address safety and fairness in gymnastics competitions, it unintentionally promoted the superiority of the vault over other events. Suffice to say, this unanticipated consequence may constitute the most important element (i.e., imbalance of disciplines) in modern women's artistic gymnastics.

CONCLUSIONS

This study marks a tentative step towards explaining why the vault seems to stand out among WAG events. Our conclusion shows that the strength of vault scores resulted from the introduction of the new vaulting table in 2001 and the new open-ended CoP in 2006, as well as further modifications in the balance between artistry and difficulty since 2009. Although these two significant changes carried out by the FIG in the early 21st century aimed to improve safety of the vault and fairness of judging, the interplay of the two significant changes promoted the vault to become the most powerful event in women's artistic gymnastics. Predictably, this unanticipated consequence may constitute the most important element (i.e., imbalance of disciplines) in the process of reification of key values (i.e., safety, artistry and difficulty) in modern women's artistic gymnastics. Moreover, as the imbalance of disciplines will definitely impact a variety of other aspects of international gymnastics (e.g., talents selection and training, judging in competitions, strategies of competitors and national teams, to name a few), we suggest research regarding such issues should continue due to the rapid progress of the sport.

REFERENCES

Atiković, A., Kalinski, S. D., Bijelić, S., & Vukadinović, N. A. (2011). Analysis results judging world championships in men's artistic gymnastics in the London 2009 year. *Sportlogia, 7*(2), 95-102. doi:10.5550/sgia.110702.en.095A

Atiković, A., Kalinski, S. D., & Čuk, I. (2017). Age trends in artistic gymnastic across World Championships and the Olympic Games from 2003 to 2016. *Science of gymnastics journal, 9*(3), 251-263.

Atiković, A., Kamenjašević, E., Mujanović, A. N., Užičanin, E., Tabaković, M., & Ćurić, M. (2020). Differences between all-around results in women's artistic gymnastics and ways of minimizing them. *Baltic journal of health and physical activity, 12*(3), 80-91. doi:10.29359/BJHPA.12.3.08

Barker-Ruchti, N. (2009). Ballerinas and pixies: a genealogy of the changing female gymnastics body. *International journal of the history of sport, 26*(1), 45-62. doi:10.1080/09523360802500089

Bučar, M., Čuk, I., Pajek, J., Karacsony, I., & Leskošek, B. (2012). Reliability and validity of judging in women's artistic gymnastics at University Games 2009. *European journal of sport science, 12*(3), 207-215.

Cervin, G. (2015). Gymnasts are not merely circus phenomena: influences on the development of women's artistic gymnastics during the 1970s. *International journal of the history of sport, 32*(16), 1929-1946. doi:10.1080/09523367.2015.1124859

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J: L. Erlbaum Associates.

Čuk, I., & Atiković, A. (2009). Are disciplines in all-around men's artistic gymnastics equal. *Sport Scientific & Practical Aspects, 6*(1/2), 8-13.

Čuk, I., & Ferkolj, S. M. (2012). Changes in technique of handspring double salto forward tucked performed on horse and vaulting table. *Acta kinesiologiae*

Universitatis Tartuensis, 13, 20. doi:10.12697/akut.2008.13.02

Cuk, I., & Karacsony, I. (2004). *Vault : methods, ideas, curiosities, history*. Ljubljana: ŠTD Sangvincki.

FIG. (2006). 2006-2008 Code of point (women's artistic gymnastics). Retrieved from https://gymnasticsresults.com/technical/code-of-points/

FIG. (2009). 2009-2012 Code of point (women's artistic gymnastics). Retrieved from https://gymnasticsresults.com/technical/code-of-points/

FIG. (2015). 2013-2016 Code of point (women's artistic gymnastics). Retrieved from http://www.gymnastics.sport/site/rules/rules.php

FIG. (2017). 2017-2020 Code of point (women's artistic gymnastics). Retrieved from http://www.gymnastics.sport/site/rules/rules.php

He, J., Montez de Oca, J., & Zhang, L. (2020). Why Team China suffered its worst performance in artistic gymnastics at Rio 2016? *International Journal of Sports Science & Coaching, 15*(5-6), 696-705. doi:10.1177/1747954120939347

Irwin, G., & Kerwin, D. G. (2009). The influence of the vaulting table on the handspring front somersault. *Sports Biomechanics, 8*(2), 114-128. doi:10.1080/14763140902745027

Jackson, M. I. (2010). *The mechanics of the table contact phase of gymnastics vaulting*. ProQuest Dissertations Publishing, (Dissertation/Thesis)

Kalinski, S. D. (2017). Elite female gymnasts at two Olympic Games-quality or luck? *Science of gymnastics journal, 9*(2), 153-164.

Kalinski, S. D., Atiković, A., Jelaska, I., & Milić, M. (2016). Performance analysis of female gymnasts' vault in elite competitions from 2008 to 2015. *Science of gymnastics journal, 8*(2), 109-123.

Kalinski, S. D., Jelaska, G., & Atiković, A. (2017). Elite female vault finals from 2008–2016. *Acta Kinesiologica, 11*(1), 62-66.

Kalinski, S. D., Padulo, J., Atiković, A., Milić, M., & Jelaska, I. (2016). Olympic Games (Beijing) analysis: the performance key. *Acta Kinesiologica, 10*(1), 48-55.

Kerr, R. (2006). The impact of Nadia Comaneci on the sport of women's artistic gymnastics. *Sporting Traditions, 23*(1), 87–102.

Kerr, R., Barker-Ruchti, N., Nunomura, M., Cervin, G., & Schubring, A. (2018). The role of setting in the field: the positioning of older bodies in the field of elite women's gymnastics. *Sociology (Oxford), 52*(4), 727-743. doi:10.1177/0038038516674676

Kerr, R., & Obel, C. (2015). The disappearance of the perfect 10: evaluating rule changes in women's artistic gymnastics. *The International Journal of the History of Sport, 32*(2), 318-331. doi:10.1080/09523367.2014.974031

Kolt, G. S., & Kirkby, R. J. (1994). Injury, anxiety, and mood in competitive gymnasts. *Perceptual and motor skills, 78*(3), 955-962. doi:10.1177/003151259407800351

Leskošek, B., Čuk, I., & Pajek, M. B. (2013). Trends in E and D scores and their influence on final results of male gymnasts at European Championships 2005–2011. *Science of gymnastics journal, 5*(1), 29-38.

Massidda, M., & Calò, C. M. (2012). Performance scores and standings during the 43rd Artistic Gymnastics World Championships, 2011. *Journal of sports sciences, 30*(13), 1415-1420. doi:10.1080/02640414.2012.710759

Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American sociological review, 1*(6), 894-904. doi:10.2307/2084615

Milčić, L., Živčić, K., & Krističević, T. (2019). Differences in vault run-up velocity in elite gymnasts. *Science of gymnastics journal, 11*(2), 201-207.

Naundorf, F., Brehmer, S., Knoll, K., Bronst, A., & Wagner, R. (2008). *Development of the velocity for vault runs in artistic gymnastics for the last decade.* Paper presented at the XXVI International Conference on Biomechanics in Sports.

Pajek, M. (2018). Individual apparatus results of female all around Olympic champions. *Science of gymnastics journal, 10*(3), 357-368.

Sands, W. A. (2000). Injury prevention in women's gymnastics. *Sports medicine (Auckland), 30*(5), 359-373. doi:10.2165/00007256-200030050-00004

Sands, W. A., & McNeal, J. R. (2002). Some guidelines on the transition from the old horse to the new table. *Technique, 22*(1), 22.

Schärer, C., Lehmann, T., Naundorf, F., Taube, W., & Hübner, K. (2019). The faster, the better? Relationships between run-up speed, the degree of difficulty (D-score), height and length of flight on vault in artistic gymnastics. *PloS one, 14*(3), e0213310-e0213310. doi:10.1371/journal.pone.0213310

Thornton, A. (2010). A fascinating look at scoring trends. Retrieved from https://www.american-gymnast.com/a-fascinating-look-at-scoring-trends/ .

Corresponding author:

Jun He
Nanjing Normal University
No. 1 Wenyuan Road Qixia district,
Nanjing, Jiangsu, China 210023
Tel.: +86 15261850615
Fax: +86 (25)85891033
E mail: hejun@njnu.edu.cn