THE RELATIONSHIP BETWEEN THE PERCENTAGE OF BODY FAT AND JUDGING IN GERMAN WHEEL GYMNASTICS

Johanna Weber¹, ²

¹Neurocognition and Action – Biomechanics, Bielefeld University, Germany
²Institut für Sportwissenschaft, Christian-Albrechts-University of Kiel, Germany

Original article DOI:10.52165/sgj.13.3.425-437

Abstract
In some sports which focus on aesthetic aspects, such as figure skating, gymnastics and dancing, there is a strong demand for athletes to be and look lean. This demand could lead to pressure on the athletes and cause health problems. It must be clarified whether this is also the case in wheel gymnastics. For this purpose, the percentage of body fat and competitive results of 203 wheel gymnasts (183 female, 20 male, age 21.2 ± 11.9 and 16.8 ± 4.9 respectively; body fat percentage 14.5 ± 3.4 and 8.0 ± 3.7 respectively) were assessed cross-sectionally and tested for correlation between the percentage of body fat and competitive results. Furthermore, their body fat percentage was compared to that of athletes from other aesthetic sports, and it was investigated whether judges might be influenced by percentage of body fat. For this purpose, technical difficulties in training and competition were compared for gymnasts with different percentages of body fat, at the same time taking into account the rating of judges’ performance as perceived the gymnasts and their satisfaction with their results. The difference between the technical difficulty during training and competition correlates positively with the body fat percentage (p ≤ .010, r = .268). Gymnasts who are a) not content with judges’ ratings; b) content with their own performance, and c) have a higher percentage of body fat, significantly differed from gymnasts where fewer than three of these parameters were true regarding deductions for technical difficulty (p ≤ .000, η² = .323).

Keywords: wheel gymnastics, body fat percentage, judges, competitive results.

INTRODUCTION
Wheel gymnastics originated in Germany in 1925 when Otto Feick built a two-rimmed wheel in which a person can stand while the wheel itself is moving (Sebesta, 2002). In current wheel gymnastics, there are four different disciplines: vault, straight line with/ without music, spiral and, more recently, cyr (Kauther, Rummel, Hussmann, Lendemans, Wedemeyer & Jaeger, 2015). In wheel gymnastics, points for overall merit are given to the athlete as a sum of points for composition, technical difficulty, and execution. The points earned for technical difficulty and composition of a routine are added to the remainder of the execution value. Deductions for the execution value are made if the athlete performs a movement technically correctly but inaccurately (Deutscher Turner-Bund (DTB), 2008). Literature states that wheel gymnastic disciplines demand strength, endurance, flexibility as well as core, explosive and jumping strength,
coordination of movement, concentration, emotional control, perception, stress resistance and self-esteem, as well as aesthetic expression (Hundrieser, 2012; Weyermann, 2016).

Mies (1994) states that wheel gymnastics can be counted among the aesthetic sports, that require, according to Potter, Lavery and Bell (1996), athletes to look slim. In aesthetic sports, such as artistic and rhythmic gymnastics, performance is related to thinness and BMI or percentage of body fat (Harris & Greco, 1990; Claessens, Lefevre, Beunen & Malina, 1999; Avilla-Carvalho, Klientrou, da Luz Palomero & Lebre, 2013; Bacciotti, Baxter-Jones, Gay & Maia, 2017) for female athletes, and this can apply to male athletes as well (Gurd & Klientrou, 2003).

Athletes in aesthetic sports display lower body fat values than non-athletes or athletes from other sports (Georgopoulos, Markou, Theodoropoulou, Berndot, Leglise & Vagenakis, 2002; Parm, Saar, Pärna, Jürimäe, Maasalu, Neissaar & Jürimäe, 2011; Galetta, Franzoni, D’Alessandro, Piazza, Tocchini, Fallahi et al., 2015). Bacciotti et al. (2018) suggested that gymnasts should be selected according to the typical physical prototype. This view, however, must be contested, as the human body changes during puberty (Matthys, 2012) and a selection decision based on these factors may not be valid for long. Psychological factors may also be a valid criterion for selection (Gonçalves, Rama & Figueiredo, 2012; Moesch, Hauge & Wiekmann, 2013). The question is whether sports clubs should select athletes according to what judges like, and whether judges should judge performance rather than physical appearance.

This leads to the question whether wheel gymnastics as an aesthetic sport is subject to the same restrictions as, e. g., rhythmic (D’Alessandro, Morelli, Evangelisti, Galetta, Franzoni, Lazzeri, Piazza & Cupisti, 2007) and artistic gymnastics (Bacciotti et al., 2017; Bacciotti et al., 2018), dancing (Potter et al., 1996) and figure skating (Jonnalagadda, Ziegler & Nelson, 2004, amongst others) when it comes to the aesthetic aspects of the athlete. In the abovementioned disciplines, athletes are required to look lean in order to achieve reasonable competitive results; there is a proven connection between the aesthetic aspect and competitive results (Bacciotti et al., 2017; Bacciotti, Baxter-Jones, Gay & Maia, 2018). Athletes in several gymnastic disciplines display low body fat percentages (see Supplementary Table).

So far, there has been no scientifically based study to determine whether wheel gymnastics should be counted among aesthetic sports. While Samuelsen (2003) claims that anthropometric factors are irrelevant for performance in wheel gymnastics because the wheels come in different sizes, Rummel (2016) names BMI as relevant for wheel gymnastics and claims that it is comparable to that in ski-jumping, where body weight is a performance-limiting factor due to its biomechanical aspects (Müller, Groschl, Müller & Sudi, 2006). Female gymnasts show a lower BMI than males and female wheel gymnasts’ BMIs are in the lowest percentile within the German population (Kromeyer-Hauschild, Wabitsch, Kunze, Geller, Geiß, Hesse et al., 2001). Some gymnasts’ BMI is below the z-scores recommended by the World Health Organization (De Onis, Onyango, Borghi, Syiam, Nishida & Siekmann, 2007). Low BMI is mostly prevalent in amateur wheel gymnasts according to Rummel (2016).

When considering aesthetic sports, athletes with a lower percentage of body fat show better competitive results in gymnastics (Vandorpe, Vandenriessche, Vaeyens, Pion, Lefevre, Philippaerts et al., 2011) and it is possible that too much value is attributed to thinness. Schek (2001, p. 22) calls this phenomenon “anorexia athletica”. If judges are subconsciously influenced by the gymnasts’ appearance, this could result in biased judging. According to Plessner &
Haar (2006), sports performance judgements are prone to bias. Findlay & Ste-Marie (2004) state that in figure skating expectations of the judges can show in the ratings. According to the authors, judges might, for example, associate low body fat percentage with a higher motivation for extensive practice and discipline and award marks accordingly, thus judging not the performance but the psychological characteristics supposedly displayed by the physical appearance of the gymnasts. It has to be stated that judging bias is present in gymnastic disciplines in relation to several factors, e. g., team order (Plessner, 1999, amongst others) and nationality (Leskošek, Čuk, Pajek, Forbes & Bučar-Pajek, 2012, amongst others); the percentage of body fat might also be one of them.

This probably also extends to younger athletes and might create a certain pressure: young athletes, not only from aesthetic sports, display a lower percentage of body fat than non-athletes (Granacher & Borde, 2017), which can lead to young athletes being underweight (Stokić, Srdić & Barak, 2005; Camargo, Gomez-Campos, Cossio-Bolaños, Barbeta, Arruda & Guerra-Junior, 2014), dissatisfied with their body image (Borrione, Battaglia, Fiorillli, Moffa, Tsopani, Piazza et al., 2013; de Oliveira, de Oliveira, de Pinho Gonçalves, Valentim Silva, Roquetti Fernandes & Fernandes Filho, 2017) and/ or developing other health-problems in order to achieve better results in competition. Their potential issues include pathologic eating behaviour (Schwidergall, Weimann, Witzel, Mölenkamp, Brehl & Böles, 1998), which can lead to amenorrhoea or malnutrition (Schek, 2001; Klentrou & Plyley, 2003; Schevchenko, Abramov, Gibson & Omar, 2008), delayed pubertal development and growth (Bayo, 2001), anorexia and bulimia. Athletes in aesthetic sports are at a higher risk of developing eating disorders than athletes in other fields, as found in a literature review by Schek (2001). In addition to creating pressure, the need to look thin can also cause an early dropout of promising athletes (Schek, 2001). A literature review shows that the body fat percentage of wheel gymnasts is similar to that of athletes in other aesthetic sports (see supplementary material).

These findings suggest that an overrating of the aesthetic aspect of athletes could be present in wheel gymnastics as well, and can thus endanger healthy development of young competitors. It is true that some athletes might receive low marks because they are overweight and technically not capable of performing difficult elements, but at the same time it is also possible that low marks are given due to the aesthetic aspect despite a flawless technical performance. If an overrating of aesthetic aspects takes place in wheel gymnastics, athletes and especially young athletes must be protected.

The main question whether judges rate gymnasts according to their body fat percentage and therefore whether wheel gymnastics has to be viewed as an aesthetic sport will be answered by researching the following, more detailed questions:

1. Is there a connection between body fat percentage and differences regarding deductions for planned difficulty in the straight line discipline (correlation)?

2. Does body fat percentage differ significantly between athletes with varying deductions for planned difficulty in the straight line discipline (one-way ANOVA)?

3. Are there differences in the judging for athletes with different percentages of body fat performing in the straight line discipline (three-way ANOVA)?

The following hypotheses are examined: It is possible that judges are too strongly taking into account the physical appearance of wheel-gymnasts which might show in a) unjustified deductions for technical difficulty for less slim gymnasts.
(gymnasts with a higher percentage of body fat might perform the difficulty safely during training, but are not given marks accordingly during competition even though they demonstrate the difficulty elements); b) a positive correlation between difficulty safely performed during training and difficulty earned during competition and percentage of body fat, and c) differences in judging for athletes with different percentages of body fat.

METHODS

Measurements included 203 voluntary participants of the German Gymnastics Federation (Deutscher Turner-Bund/DTB), section gym wheel. Informed consent was obtained from all participants. The age ranged from 6 to 58 for female (N = 178, age average = 21.17 ± 11.91) and from 7 to 27 for male (N = 20, age average = 16.84 ± 4.90) gymnasts.

Skinfolds were recorded using a calliper. Competitive results, performance during training, evaluation of a recent competition performance, gender, age, and competitive level were obtained using a specially developed questionnaire. All values were recorded at major national competitions in 2018.

Body fat percentage was calculated using the Siri equation (1956) for calculating body fat using three skinfolds for female gymnasts (Jackson, Pollock & Ward, 1980). For male gymnasts, Siri (1956) and Jackson & Pollock (1978) equations were used to calculate percentage of body fat using three skinfolds and two circumferences. Due to different compositions of body tissue, female and male athletes require individual calculation methods (Jackson & Pollock, 1978; Jackson et al., 1980). Due to different compositions of body tissue, female and male athletes require individual calculation methods (Jackson & Pollock, 1978; Jackson et al., 1980).

There was a significant difference in percentage of body fat for male and female gymnasts (female: % body fat average = 14.54 ± 3.4; male: % body fat average = 8.00 ± 3.74) which is in accordance with earlier findings (see Table 3). Since there were not enough male gymnasts, the results for them were not evaluated further except for a short comparison to previous studies.

Since the sample displays a broad age-range, it was necessary to rule out any age effects. Age did not correlate in any way with deductions for difficulty or overall merit. There was no age-related difference in athletes who were content or not content with judges’ ratings or their own performance. No difference in age was found between athletes who were in the highest quartile of body fat percentage, were content with their own performance, not content with the ratings and had deductions for difficulty, and all other athletes.

The questionnaire asked for age, gender, straight line difficulty (technical merit) achieved during training, straight line difficulty achieved at the recent
competition, self-rating of own performance at the recent competition and rating of the judges’ performance who judged the athlete. The difficulty difference or planned difficulty was calculated as the difference between technical difficulty achieved during training minus technical difficulty achieved during competition. In German competitive wheel gymnastics, athletes are required to hand in a difficulty card before competition, stating what difficulty they were able to perform during training and are therefore intending to perform in competition. Usually, gymnasts hand in cards that show which difficulty was safely performed during training. Cards are prepared in cooperation with the coaches and it is a common practice to compose the card realistically. For this study, values from those difficulty cards were used to calculate the difference between difficulty safely performed during training and difficulty earned during competition.

To answer the first research question, correlations between the difference in technical difficulty during training and in competition (difficulty difference = difficulty_training − difficulty_competition) on the one hand and body fat percentage on the other hand were calculated via Spearman correlation with correlation levels >0.1 (weak), >0.3 (moderate) and >0.5 (strong).

To address the second research question, age controlled one-way ANOVA was performed for:

Testing for differences in body fat percentage between athletes who earned lower difficulty scores during competition than during training and athletes who received the same technical difficulty during both training and competition.

Testing for differences regarding deductions for planned technical difficulty between those in the quartile with highest percentage of body fat and those in the lower three quartiles.

To address the third research question, age controlled three-way ANOVA was performed for:

Testing for differences in the dependent factor “deduction to planned difficulty” using the following independent factors: “not content with judging performance”, “content with own performance” and “ranked within the quartile with the highest percentage of body fat”.

Finally, to address the third research question, age controlled one-way ANOVA was performed for:

Testing for differences between athletes who had different deductions for planned difficulty and discrepancies in the levels of satisfaction with the judges’ and their own performance.

The criterion level for significance was set at $p<0.05$ and the trend significance at $p<0.10$. The effect size was evaluated with $\eta^2$ (Eta partial squared), where $0.01<\eta^2<0.06$ constitutes a small effect, $0.06<\eta^2<0.14$ constitutes a medium effect and $\eta^2>0.14$ constitutes a large effect (Cohen, 1988).

Representativity was calculated according to Rinne (2008) with confidence interval 95% and delta_max = 3.45%. The tester had a retest-reliability of Cronbach’s alpha = 0.987 for the measurement of skin folds.

Statistical analysis was performed in SPSS, version 25 (SPSS, Inc., Chicago, IL).

RESULTS

Regarding research question 1, there was a weak correlation for the difference in technical difficulty during training and competition (difficulty difference) and percentage of body fat ($r = .268; p < .010$, Figure 1), while elimination of one outlier with more than 25% of body fat did not change significance nor correlation level.
Figure 1. Correlation between deductions for planned difficulty and percent body fat.

Table 1
Differences regarding percentage of body fat in respect of deductions for difficulty and deduction for difficulty for athletes in different quartiles body fat percentage.

<table>
<thead>
<tr>
<th>Differences</th>
<th>N</th>
<th>$X \pm SD$</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Body fat by deduction for difficulty for female athletes</td>
<td>35</td>
<td>15.55 ± 3.41</td>
<td>0.050</td>
<td>0.066</td>
</tr>
<tr>
<td>(deduction vs. no deduction)</td>
<td>56</td>
<td>14.08 ± 3.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deduction for difficulty by % of body fat for female athletes*</td>
<td>25</td>
<td>0.40 ± 0.51</td>
<td>0.035</td>
<td>0.077</td>
</tr>
<tr>
<td>(highest quartile vs. other quartiles*)</td>
<td>62</td>
<td>0.17 ± 0.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Limits at 12.48, 14.41 and 16.47 %
Table 2  
*Differences regarding deduction for difficulty in respect of percentage of body fat, judges’ rating and performance self-evaluation.*

<table>
<thead>
<tr>
<th>Differences</th>
<th>N</th>
<th>$X ± SD$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deduction for planned difficulty for gymnasts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Not content with judges, content with own performance, in the highest quartile of body fat %” vs. “Not content with judges, content with own performance, not in the highest quartile of body fat %”</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>$0.67 ± 1.15$</td>
<td>$0.000$</td>
<td>$0.323$</td>
</tr>
<tr>
<td>Descriptive statistics (percentage of body fat)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the highest quartile of body fat percentage</td>
<td>44</td>
<td>$18.96 ± 2.26$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In the lower three quartiles of body fat percentage</td>
<td>133</td>
<td>$16.45 ± 2.24$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not content with judges, content with performance, deduction for planned difficulty (all true) vs. Two or less of the above true</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>$15.67 ± 2.89$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>87</td>
<td>$14.65 ± 3.42$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3  
*Percentage of body fat in relation to rating of judging/ performance and deduction for planned difficulty.*

<table>
<thead>
<tr>
<th>Ratings/ Difference</th>
<th>N</th>
<th>$X ± SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judges ok, performance not ok, no deduction</td>
<td>8</td>
<td>$12.73 ± 3.18$</td>
</tr>
<tr>
<td>Judges not ok, performance not ok, deduction</td>
<td>4</td>
<td>$13.70 ± 2.11$</td>
</tr>
<tr>
<td>Judges ok, performance ok, no deduction</td>
<td>41</td>
<td>$14.14 ± 3.14$</td>
</tr>
<tr>
<td>Judges not ok, performance ok, deduction</td>
<td>3</td>
<td>$15.64 ± 3.53$</td>
</tr>
<tr>
<td>Judges ok, performance not ok, deduction</td>
<td>15</td>
<td>$15.73 ± 2.73$</td>
</tr>
<tr>
<td>Judges ok, performance ok, deduction</td>
<td>12</td>
<td>$15.99 ± 4.64$</td>
</tr>
<tr>
<td>Judges not ok, performance ok, no deduction</td>
<td>4</td>
<td>$16.86 ± 4.20$</td>
</tr>
</tbody>
</table>

Judges ok/not ok: Gymnast was content/not content with difficulty ratings; Performance ok/not ok: Gymnast was content/not content with own performance regarding technical difficulty; Deduction/no deduction: Difficulty was not rated as planned/difficulty was rated as planned.

When considering research question 2, the differences in percentages of body fat between athletes with or without deduction for planned difficulty is present in our sample (Table 1). Values differ significantly between athletes with higher and lower body fat in deductions for planned difficulty. Athletes received higher deductions for their planned difficulty when they had a higher percentage of body fat, as indicated by the ANOVA testing for differences in the amount of difficulty points lost when comparing training and competition within the highest quartile and the three lower quartiles (Table 1).
For research question 3, athletes, who at the same time a) were pleased with their own performance, b) did not agree with the rating of the judges and c) displayed a higher percentage of body fat, received a significantly higher deductions for planned difficulty compared to athletes who were dissatisfied with ratings but satisfied with their own performance and not in the highest quartile of body fat percentage. However, it has to be stated that due to further dividing into subgroups in the three-way ANOVA the, sample size for these two subgroups is quite small (Table 2).

Additionally, it is apparent that athletes with different levels of satisfaction with their own and the judges’ performance as well as different deductions for difficulty scores showed characteristic percentages of body fat, but only by trend (p < .072; $\eta^2 = .148$, Table 3).

DISCUSSION

The correlation between deduction for planned difficulty and body fat percentage suggests that judges may have deducted more difficulty points from less slim gymnasts (Figure 1). Judges may be biased against athletes with higher percentage of body fat, as has already been shown that they are prone to bias in terms of other factors, e. g., team order (Plessner, 1999) or nationality (Leskošek et al., 2012).

The results in Table 1 also lead to the assumption that judges could have been influenced by the physical appearance of the athletes, since they deducted more difficulty points from athletes with a higher percentage of body fat.

Furthermore, results in Table 2 suggest that athletes with a higher percentage of body fat were subject to higher deductions for difficulty despite having performed the difficulty safely during training and being content with their performance during competition. However, for a more accurate statement on the matter, a larger sample size would be required.

Table 3 shows that athletes who made errors in their routine during competition received their targeted difficulty scores from the judges anyway when their body fat percentage was lower and did not get it if their body fat percentage was higher. This might suggest that the judges favoured slimmer gymnasts since athletes who received their difficulty as planned had a rather low percentage of body fat. The athletes who did not mess up but still received deductions for their difficulty in comparison to their training may have also been subject to bias due to their higher percentage of body fat. Motivational factors which have been tested within the same sample for another project did not correlate significantly with deductions for difficulty. More detailed research might be necessary to test for connections between, e. g., age, percentage of body fat, self-esteem, and other psychological factors.

Our results suggest that judges may be prone to award points according to the physical appearance of the gymnasts. Athletes who performed well and were content with the judging performance despite deductions for their difficulty had a higher percentage of body fat (Table 3). This might be due to the fact that less slim athletes had already become used to being underrated. The highest percentage of body fat can be seen in athletes who were not content with judging performance and nevertheless earned their difficulty as planned while being content with their own performance. It is possible that they received deductions for other factors, like overall merit, composition or execution, caused dissatisfaction. This can unfortunately not be clarified, since the composition value is not shown separately during competition and most athletes are not judged for their execution and overall merit during training. This makes it impossible to compare these three values in the light of the difference between competition and training.
When comparing wheel gymnasts to athletes from other aesthetic sports, especially gymnastics, it is evident that German wheel gymnasts display quite low body fat values in comparison to other athletes (see Supplementary Table). While body fat percentage in athletes in other gymnastic disciplines ranges between 9 and 22% for females and between 8 and 11% for males, wheel gymnasts in this study displayed values of 14.53% for females and 8.00% for males on average. This suggests that the aesthetic aspect might be important in wheel gymnastics. Thus, this study is not concurrent with the assumption of Samuelsen (2003) who stated, without having conducted a corresponding study, that it is irrelevant for performance if a gymnast does not have ideal body proportions.

The results of our study match those of, e.g., Schwidergall et al. (1998) and results from previous studies, because, as expected, females were leaner than males. 8.0% body fat of male wheel gymnasts is (although displaying broader variation) in line with the findings of Šibanc et al. (2000) and Gurd & Klentrou (2003), who further state that this has no effect on physical and pubertal development of male gymnasts. The lack of male participants as well as the small sample size in general have to be considered a strong limitation of this study.

CONCLUSIONS

It must be taken into account that the present study could only produce reliable findings for female German wheel gymnasts. Analyzing male athletes separately was not possible due to too few cases, but descriptive statistics hint that a similar effect as for females might be present also for males. To make statements about male wheel gymnasts and provide more specifics for female wheel gymnasts, a larger sample is necessary. The study should also be repeated in other countries and at international tournaments.

Although biomechanical aspects can largely be ruled out as the cause of athletes earning lower difficulty scores during competition than during training, psychological factors could be the reason that athletes would not do their best in a competition, although they were able to perform at the intended difficulty level when training. Therefore, our findings have to be interpreted carefully. Nevertheless, they suggest that the aesthetic aspect is important in wheel gymnastics.

In summary, the abovementioned findings lead to the conclusion that physical appearance is overrated in wheel gymnastics and that the sport indeed is an aesthetic sport. Although the percentage of body fat can influence gym wheel performance via its biomechanical aspects, our study suggests that it most probably also influences the performance of judges. Female gymnasts are subject to underrating if they have a higher body fat percentage despite performing well. Judges should be educated about this fact so that they do not put undue pressure on young female athletes in particular, as this may lead to early dropout and psychological or physical harm to gymnasts.

REFERENCES


Granacher, U., & Borde, R. (2017). Effects of sport-specific training during the early stages of long-term athlete development on physical fitness, body composition, cognitive, and academic...


Corresponding author:

Johanna Weber
Neurocognition and Action - Biomechanics
Bielefeld University
Universitätsstraße 25
33615 Bielefeld, Germany
or
Institut für Sportwissenschaft
Christian-Albrechts-University of Kiel
Olshausenstr. 74
24098 Kiel, Germany

Phone: 0170 41 000 10
Email: j.weber_@hotmail.de

Article received: 16.2.2021
Article accepted: 23.7.2021