Journal of Athletic Enhancement

Research Article

A SCITECHNOL JOURNAL

Intra and Inter-rater Reliability of the Volodalen® Scale to Assess Aerial and Terrestrial Running Forms

Patoz A[1]*, Gindre C[1], Mourot L[2,3] and Lussiana T[4]

Abstract

Running form being a global system defined by several biomechanical parameters, it is of major interest to assess it using a global method. For this purpose, the Volodalen® scale was developed. This scale, based on five items, attributes a global subjective score (V®score) to the running form of individuals and allows their classification along an aerial-terrestrial continuum. As no study has yet reported the reliability of such scale, the aim of this paper was to evaluate its intra- and inter-rater reliability. Thirty-six runners ran two 10-min running trials. Runners were classified according to their V®score by two experts and one novice raters. Relative and absolute reliability, and systematic bias were determined by intra-class correlation coefficient (ICC), coefficient of variation (CV), and statistically significant difference, respectively. Regarding the global V®score, high intra-rater reliability for the expert (CV=6.1 ± 7.0%, ICC=0.940, and *p*-value=0.864) and high inter-rater reliability for both novice (CV=6.6 ± 6.5%, ICC=0.945, and *p*-value=0.248) and expert (CV=6.8 ± 5.7%, ICC=0.950, and *p*-value=0.405) raters were involved. However, several subcomponents of the V®score reported poor inter-rater reliability. The Volodalen® scale is a reliable tool to assess global running forms whatever the degree of expertise whereas the subjective assessment of a single parameter of the V®score is rater-dependent.

Keywords

Running form; Biomechanics; Reliability; Validation; Subjective scale; Coaching; Training

Introduction

Running technique is highly variable between runners where the associated running form is a global and dynamic system with several interconnected biomechanical parameters [1]. As typical examples, McMahon et al. [2] tested the "Groucho running", a running form with excessive knee flexion and associated with increased ground contact time and step length together with decreased flight time and vertical oscillation of the body. On the other hand, Arendse et al. [3] investigated "Pose running", characterized by mid- to fore-foot striking, short contact times and step lengths, and less knee flexion during stance. However, when taken separately, biomechanical parameters can lead to misunderstandings. For instance, Daoud et al. [4] showed that mid- or fore-foot runners reduce the occurrence of injuries compared to rear-foot strikers, while no statistically significant difference in the injury rate between rear-, mid-, or fore-foot runners has been reported in three large-scale epidemiological studies [5-7]. Therefore, it is of major interest to use multicomponent methods instead of single parameter analyses.

One possibility to assess the global running form is to use the duty factor, a biomechanical parameter that takes into account simultaneously the ground contact time and the swing phase [8,9]. However, a device is always needed to obtain the necessary temporal informations (for instance a force platform or a motion capture system). Therefore, the required measurement is not always easy to perform, especially in outdoor conditions. As subjective evaluation is frequently used by coaches in the sports field [10], another option is provided by the subjective scale developed by Gindre et al. [11], i.e. the Volodalen® scale. This scale is based on five subjective visual observations of the running form: vertical oscillation of the head, antero-posterior motion of the elbows, pelvis position at ground contact, foot position at ground contact, and strike pattern. It has proven to reflect quantifiable objective parameters [12]. Being practical in nature, the Volodalen® scale allows to classify runners, according to their running form, along a continuum. With this scale, runners can be divided into two categories term aerial (AER) and terrestrial (TER). The TER runners rely more on a forward propulsion strategy whereas the AER runners are more capable to store and release elastic energy [13]. Despite the usage of different biomechanical strategies, these two groups of runners are not showing any difference in their running economy [14].

According to Lussiana et al. [12], the runners were classified along the aerial-terrestrial continuum by two researchers with several years of experience in using the subjective rating score. However, it has been shown for functional movement screen (FMS) that subjective assessment does not necessarily necessitate a training for novices [15,16]. Indeed, inter-rater reliability analyzed using the weighted kappa statistic demonstrated excellent agreement for novices [15]. Moreover, intra-rater reliability assessed by intra-class correlation coefficient (ICC) reported as good regardless of previous experience with FMS [17,16]. Despite the growing number of studies using the Volodalen® scale, none has yet reported its intra- and inter-rater reliability regarding its usage by experts and novices. Determining the psychometric properties of the Volodalen® scale for classifying runners along the aerial-terrestrial continuum could guide current users into the correctness of their subjective appreciation, facilitate between-study comparison, and help potential users to estimate the advantages and drawbacks of using this subjective scale.

Our aim was to evaluate the intra- and inter-rater reliability of the Volodalen® scale. First, we hypothesized that intra-rater evaluation of the Volodalen® scale would be reliable. Second, as observed for FMS [15,17], we hypothesized that expert versus novice and between experts evaluation of the Volodalen® scale would also be reliable.

Materials and Methods

Subjects and experimental procedure

Thirty-six runners with at least three years of regular track/road training (age: 32.7 ± 9.4 years, height: 176.9 ± 5.9 cm, body mass: 70.1

***Corresponding author:** Aurélien Patoz, Volodalen, Swiss SportLab, Aigle, Switzerland, Tel: +41 (0)79 535 07 16; E-mail: aurelien@volodalen.com

Received: May 03, 2019 **Accepted:** June 13, 2019 **Published:** June 20, 2019

± 7.5 kg, and maximal aerobic speed: 17.3 ± 1.1 km/h) participated in this study. The university's institutional review board approved the study protocol prior to participant recruitment (CPP: 2014-A00336-41), which was conducted in accordance with the latest amendments of the Declaration of Helsinki. Each participant completed two experimental sessions interspersed by one month. Participants did not change their training program within this one-month period. The second session was matched for daily time with respect to the first one and occurred under similar weather conditions. During both sessions, they performed a 10 min running trial on a running track at a self-selected speed (velocity between 2.5 and 3.5 m/s). Participants ran alone and were only asked to run at their preferred speeds.

Subjective assessment of running gait

During the first session, three raters, which were runners and running coaches, independently focused on the overall movement pattern of participants. Amongst the raters, two were experts with more than 10 years of experience (S_1E_1 and S_1E_2) using the Volodalen® scale whereas the third one was a novice (S_1N) and used this scale for the first time. The novice received a one-hour training session on the usage of the Volodalen® scale by one of the two experts. The raters paid attention to five key elements, i.e. vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E) [see 11, 12, for details]. Each element was scored from one to five, leading to a global subjective score (V®score) that represents the running form of participants. This V®score ultimately allows the classification of runners into two different categories: TER (V®score ≤ 15) or AER (V®score>15) group. During the second session, one of the two experts (S_2E_1) focused once more on the running form of participants, scored each of the five key elements to obtain a V®score, and reclassified the runners accordingly.

Statistical analysis

Descriptive statistics are presented using mean ± standard deviation (s.d.). Normality of data has been verified using Shapiro-Wilk tests. To detect systematic bias [18], non-parametric Wilcoxon signed-rank tests were used when normality failed and bilateral Students t-tests for paired samples otherwise. Although useful, this statistical assessment should not be used on its own as an assessment of reliability, as this statistic provides no indication of random variation between tests [18].

To assess the relative intra- and inter-rater reliability, ICC was computed. These coefficients were computed as [ICC=1-(SEM/SD)²], where SEM is the standard error in measurement and SD is the mean between the s.d. obtained from the evaluation of two different raters or by the same rater twice. The SEM was computed as [SEM=s.d. (between–trial difference in measures)/√2]. To quantify the absolute reliability, the difference in means (Δ, in raw units and %) and the coefficient of variation (CV) were computed. The latter is defined as the mean of the coefficient of variation of each sub-pair of data.

On the basis of commonly used thresholds, the relative reliability was considered poor, fair, and good when the corresponding ICC values were <0.4, 0.4-0.75 and ≥ 0.75, respectively [19,20]. The absolute reliability was considered adequate, moderate, and inadequate when the corresponding CV values were ≤ 10%, 10-20%, and >20%, respectively [21-23]. Statistical analysis was done using a customized script in R 3.5.0 (The R Foundation for Statistical Computing, Vienna, Austria) and Microsoft Excel 2016 (Microsoft Corp., Redmont, WA, USA) with a level of significance set at $\alpha \leq 0.05$.

Results

The subjective assessment S_1E_1 led to AER and TER groups composed of N_{AER}=18 and N_{TER}=18, respectively. These two groups were used as the reference when S_1E_1 was compared to S_2E_1, S_1N, and S_1E_2.

Intra-rater reliability

When comparing S_1E_1 and S_2E_1, a total of 34 matches (94.4%) were obtained between the two different subjective ratings, there were 100.0% and 88.9% concordance for AER and TER groups, respectively, which means two subjects were transferred from TER to AER group. Even though there were two additional runners in the AER group of S_2E_1, as the 18 AER runners of S_1E_1 were included in the 20 AER runners of S_2E_1, this led to an AER concordance of 100.0%. Means and s.d. of S_1E_1 and S_2E_1 rating scores are depicted in Figure 1 for each item of the Volodalen® scale together with the V®score. Intra-rater reliability, determined by ICC, Δ (raw units and %), CV, and systematic bias is reported in Table 1 for each item of the scale together with the V®score.

Overall, the V®score demonstrated good and adequate relative and absolute reliability, respectively, except for the AER group where relative reliability was fair. As for each item that constitutes the V®score, good relative reliability is obtained when considering all participants (ALL) whereas AER and TER groups demonstrated fair to good relative reliability. Absolute reliability was moderate to adequate when considering ALL as well as AER and TER groups.

Expert versus novice reliability

The S_1N subjective rating gave 35 matches (97.2%) with respect to the S_1E_1 rating. The AER and TER concordances were 100.0% and 94.4%, respectively, which means one subject was transferred from TER to AER group. Even though there was one additional runner in the AER group of S_1N, as the 18 AER runners of S_1E_1 were included in the 19 AER runners of S_1N, this led to an AER concordance of 100.0%. Means and s.d. of S_1E_1 and S_1N rating scores are depicted in Figure 2 for each item of the Volodalen® scale together with the V®score. Expert versus novice reliability is reported in Table 2.

Overall, the V®score demonstrated good and adequate relative and absolute reliability, respectively, except for the AER group where relative reliability was fair. As for each item that constitutes the V®score, good relative reliability is obtained when considering ALL whereas AER and TER groups demonstrated fair to good relative reliability, except for the antero-posterior motion of the elbows (B) of both AER and TER groups and the pelvis position at ground contact (C) of AER group which reported as poor. Absolute reliability was moderate to adequate when considering ALL as well as AER and TER groups.

Between experts reliability

The S_1E_2 subjective rating gave a total of 35 matches (97.2%) with respect to the S_1E_1 rating. There were 100.0% and 94.4% concordance for TER and AER groups, respectively, which means one subject was transferred from AER to TER group. Even though there was one additional runner in the TER group of S_1E_2, as the 18 TER runners of S_1E_1 were included in the 19 TER runners of S_1E_2, this led to a TER concordance of 100.0%. Means and s.d. of experts' rating scores are depicted in Figure 3 for each item of the Volodalen® scale together with the V®score. Between experts reliability is reported in Table 3.

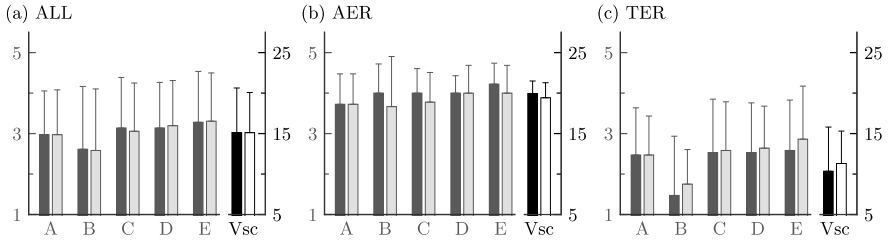

Figure 1: Means and standard deviations (error bars) of each item of the Volodalen® scale: vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E), together with the V®score (Vsc). Parameters were assessed twice by the same expert rater: S_1E_1 (dark gray and black boxes) and S_2E_1 (light gray and white boxes) for (a) all (ALL), (b) aerial (AER), and (c) terrestrial (TER) runners.

Table 1: Relative (intra-class correlation coefficient, ICC) and absolute (difference in means, Δ and coefficient of variation, CV) reliability, and systematic bias (*p*-value) of each item of the Volodalen® scale: vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E), together with the V®score. Parameters were assessed twice by the same expert rater (S_1E_1 and S_2E_1) for all (ALL), aerial (AER), and terrestrial (TER) runners.

Parameters		A	B	C	D	E	V® score
ALL	ICC	0.857	0.861	0.782	0.888	0.857	0.940
	Δ	0.00 (0.0%)	0.03 (1.1%)	0.08 (2.7%)	-0.06 (-1.8%)	-0.03 (-0.8%)	0.03 (0.2%)
	CV (%)	8.1 ± 16.3	12.2 ± 18.4	11.4 ± 17.6	5.8 ± 9.9	9.9 ± 15.7	6.1 ± 7.0
	p-value	1.000	0.842	0.594	0.565	0.830	0.864
AER	ICC	0.896	0.618	0.418	0.625	0.760	0.661
	Δ	0.00 (0.0%)	0.33 (8.3%)	0.22 (5.6%)	0.00 (0.0%)	0.22 (5.3%)	0.78 (3.9%)
	CV (%)	2.2 ± 6.5	15.0 ± 18.4	7.9 ± 12.6	5.7 ± 8.4	4.2 ± 8.2	4.0 ± 4.7
	p-value	1.000	0.145	0.330	1.000	0.072	0.057
TER	ICC	0.586	0.600	0.723	0.827	0.771	0.852
	Δ	0.00 (0.0%)	-0.28 (-22.7%)	-0.06 (-2.4%)	-0.11 (-4.9%)	-0.28 (-11.9%)	-0.72 (-7.0%)
	CV (%)	14.0 ± 20.8	9.4 ± 18.6	15.0 ± 21.3	5.8 ± 11.4	15.7 ± 19.3	8.2 ± 8.4
	p-value	1.000	0.089	0.821	0.424	0.152	0.108

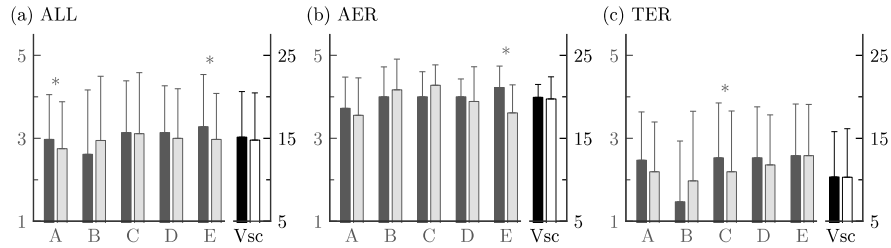

Figure 2: Means and standard deviations (error bars) of each item of the Volodalen® scale: vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E), together with the V®score (Vsc). Parameters were assessed by an expert (S_1E_1, dark gray and black boxes) and a novice (S_1N, light gray and white boxes) rater for (a) all (ALL), (b) aerial (AER), and (c) terrestrial (TER) runners. ∗ indicates a significant difference (systematic bias) between expert and novice raters.

Overall, the V®score demonstrated good and adequate relative and absolute reliability, respectively, except for the AER group where relative reliability was fair. As for each item that constitutes the V®score, good relative reliability is obtained when considering ALL whereas AER and TER groups demonstrated fair to good relative reliability. Absolute reliability was moderate to adequate when considering ALL as well as AER and TER groups.

Discussion

In agreement with our hypotheses, the Volodalen® scale is a reliable tool to assess global running forms whatever the degree of expertise in the usage of the Volodalen® scale. Therefore, the classification of AER and TER runners, according to their running form, can be performed by novices and experts. Nonetheless, discrepancies can exist between expert and novice raters when considering each item of the V®score separately. This highlights the importance to assess the running form using a multicomponent approach. In this study, intra- and inter-rater absolute reliability defined by CV values (2.2-17.6%) are in line with those obtained for biological system (10-15%) [23]. The use of several statistical parameters is recommended for quantifying the reliability of measures [18]. We obtained that almost all statistical indicators, except ICC for AER group which were reported as fair, involved high intra- and inter-rater reliability for the V®score.

When considering each item that defines the V®score, a high intra-rater reliability was also involved because no statistical parameters reported poor, inadequate, or statistically significant

Table 2: Relative (intra-class correlation coefficient, ICC) and absolute (difference in means, Δ and coefficient of variation, CV) reliability, and systematic bias (*p*-value) of each item of the Volodalen® scale: vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E), together with the V®score. Parameters were assessed by an expert (S_1E_1) and a novice (S_1N) rater for all (ALL), aerial (AER), and terrestrial (TER) runners. Significant differences and poor relative reliabilities between expert and novice raters are depicted in italics.

Parameters		A	B	C	D	E	V® score
ALL	ICC	0.857	0.810	0.868	0.806	0.861	0.945
	Δ	0.22 (7.5%)	-0.33 (-12.8%)	0.03 (0.9%)	0.14 (4.4%)	0.31 (9.3%)	0.36 (2.4%)
	CV (%)	11.2 ± 15.7	14.3 ± 19.6	12.0 ± 17.4	11.7 ± 15.4	10.5 ± 12.3	6.6 ± 6.5
	p-value	*0.036*	0.053	0.835	0.263	*0.008*	0.248
AER	ICC	0.704	*0.091*	0.612	*0.297*	0.714	0.481
	Δ	0.17 (4.5%)	-0.17 (-4.2%)	-0.28 (-6.9%)	0.11 (2.8%)	0.61 (14.5%)	0.44 (2.2%)
	CV (%)	8.3 ± 10.9	13.0 ± 13.5	6.6 ± 10.5	12.4 ± 11.3	11.8 ± 9.8	6.3 ± 5.9
	p-value	0.299	0.524	0.120	0.627	*0.001*	0.477
TER	ICC	0.719	*0.392*	0.826	0.768	0.833	0.940
	Δ	0.28 (12.5%)	-0.50 (-40.9%)	0.33 (14.6%)	0.17 (7.3%)	0.00 (0.0%)	0.28 (2.7%)
	CV (%)	14.1 ± 19.2	15.7 ± 24.6	17.4 ± 21.3	11.0 ± 18.9	9.1 ± 14.4	7.0 ± 7.3
	p-value	0.073	*0.031*	*0.041*	0.233	1.000	0.311

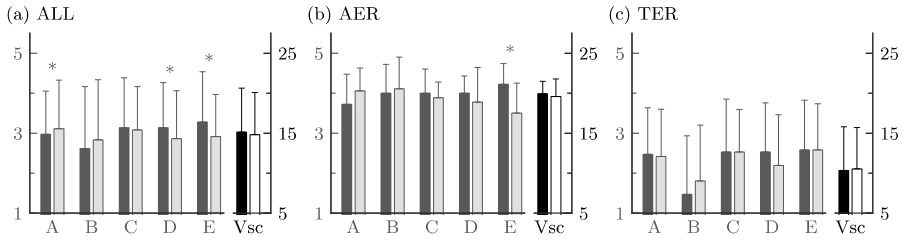

Figure 3: Means and standard deviations (error bars) of each item of the Volodalen® scale: vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E), together with the V®score (Vsc). Parameters were assessed by a first (S_1E_1, dark gray and black boxes) and a second (S_1E_2, light gray and white boxes) expert rater for (a) all (ALL), (b) aerial (AER), and (c) terrestrial (TER) runners. ∗ indicates a significant difference (systematic bias) between expert raters.

Table 3: Relative (intra-class correlation coefficient, ICC) and absolute (difference in means, Δ and coefficient of variation, CV) reliability, and systematic bias (*p*-value) of each item of the Volodalen® scale: vertical oscillation of the head (A), antero-posterior motion of the elbows (B), pelvis position at ground contact (C), foot position at ground contact (D), and strike pattern (E), together with the V®score. Parameters were assessed by two different expert raters (S_1E_1 and S_1E_2) for all (ALL), aerial (AER), and terrestrial (TER) runners. Significant differences are depicted in italics.

Parameters		A	B	C	D	E	V® score
ALL	ICC	0.802	0.925	0.832	0.860	0.804	0.950
	Δ	-0.14 (-4.7%)	-0.22 (-8.5%)	0.06 (1.8%)	0.28 (8.8%)	0.36 (11.0%)	0.33 (2.2%)
	CV (%)	12.9 ± 17.9	10.3 ± 15.0	10.6 ± 17.2	10.0 ± 14.7	12.5 ± 15.3	6.8 ± 5.7
	p-value	0.263	*0.036*	0.644	*0.014*	*0.007*	0.405
AER	ICC	0.635	0.642	0.627	0.685	0.547	0.551
	Δ	-0.33 (-9.0%)	-0.11 (-2.8%)	0.11 (2.8%)	0.22 (5.6%)	0.72 (17.1%)	0.61 (3.1%)
	CV (%)	8.2 ± 9.6	8.0 ± 9.3	6.7 ± 10.0	6.9 ± 12.9	14.4 ± 13.8	5.1 ± 5.6
	p-value	*0.041*	0.530	0.484	0.203	*0.002*	0.376
TER	ICC	0.536	0.679	0.654	0.577	0.812	0.888
	Δ	0.06 (2.5%)	-0.33 (-27.3%)	0.00 (0.0%)	0.33 (14.6%)	0.00 (0.0%)	0.06 (0.5%)
	CV (%)	17.6 ± 22.9	12.6 ± 19.1	14.6 ± 21.8	13.2 ± 15.9	10.6 ± 16.9	8.4 ± 5.5
	p-value	0.821	*0.020*	1.000	*0.041*	1.000	0.871

for relative reliability, absolute reliability, and systematic bias, respectively. Moreover, the subjective intra-rater ICC values (0.418-0.940) obtained here were in line with the objective ones obtained for the coordinates of ten anatomical landmarks on a frontal and a dorsal body posture photograph (0.66-1.00) [24].

As for expert versus novice (S_1E_1 and S_1N) and between experts (S_1E_1 and S_1E_2) comparisons, we observed a moderate to adequate absolute reliability and a fair to good relative reliability for the V®score. However, several parameters reported a poor relative reliability as well as a systematic bias when comparing S_1E_1 and S_1N. Systematic bias was also observed for several parameters when comparing the two experts (S_1E_1 and S_1E_2). As the scale of a single-item (1-5) is five times smaller than the scale of the V®score (5-25), the impact of a discrepancy in a single item of the scale is much bigger than its impact on the global V®score. In addition, scoring a parameter such as the antero-posterior motion of the elbows (B) seemed harder than the other parameters. A possible explanation could be a more important focus on some parameters than on others. For instance, more attention could be given to the strike pattern (E) due to the existence of important debates on what the optimal strike pattern is.

Moreover, in the case of the strike pattern (E), a fixed reference is given by the ground. This fixed reference can help in the scoring of this parameter compared to the one of the antero-posterior motion of the elbows (B) for which no fixed reference exists. Despite everything, the variations of the five key parameters of the Volodalen® scale are small (few centimeters). These only small variations emphasize the need for the raters to be runners and running coaches, i.e. they are already "expert" in terms of human observation.

Nonetheless, our results suggested that the V®score is highly reliable. A similar result was obtained when evaluating the inter-rater reliability of the brief psychiatric rating scale [25]. The authors observed that the trends between professional groups were not large and not always related to degree of training. Another study [16] evaluated FMS with raters of different educational background and experience (two-hour training session for novice). Their ICC for individual FMS movements ranged from 0.3 to 0.98, which are in line with our values for each item of the V®score (0.091-0.925). Moreover, they also observed that inter-rater reliability was good regardless of education or previous experience with FMS. Finally, Minick et al. [15] and Onate et al. [17] showed excellent agreements between novice and expert raters for FMS. Taken together, these results emphasized the importance to consider the global rating score when classifying runners into two different groups with the Volodalen® scale, and not a sub-item in isolation.

A limitation of the present study is that intra-rater reliability was evaluated using two different running sessions. Therefore, as runners could have slightly changed their running form and preferred running speed, for instance due to fatigue, a small bias could be present. Nevertheless, this possibility is dampened by the high expertise of the runners in our cohort and by the non-modification of their training program within the one-month period. To reduce this bias, a possible solution would be to use video recordings. However, due to the fact videos are taken from a fixed point and are shorts in time, the rating task can even be more difficult.

Conclusion

On the basis of traditional thresholds, the relative (i.e. ICC) and absolute (i.e. CV) reliability of the V®score classified as good (fair for AER group) and adequate, respectively, for both intra- and inter-rater comparisons. These findings suggest that this scale is practically useful for classifying individuals, according to their running form, in two different running categories and can be used by novices and experts in the usage of the Volodalen® scale. When each item of the V®score is taken separately, the intra- and inter-rater absolute reliability classified as moderate to adequate, whereas the relative reliability reported as poor to good. Furthermore, systematic bias shown by statistically significant differences were also observed for the five key parameters when comparing different raters. Hence, the present results emphasized the importance to assess the running form using a global approach. Therefore, we suggest users who would like to use the Volodalen® scale to not use only one of its subcomponents to assess running forms.

Acknowledgments

This study was supported by the University of Franche-Comté (France) and the Exercise Performance, Health, and Innovation Platform of Besançon. The authors warmly thank the subjects for their time and cooperation.

Conflicts of Interest

C.G. is the originator of the Volodalen® scale. However, this paper does not constitute endorsement of the scale by the other authors and stems completely from a PhD research project undertaken at the Franche-Comté University by T.L.

References

1. Dickinson MH, Farley CT, Full RJ, Koehl MA, Kram R, et al. (2010) How animals move: An integrative view. Science 288: 100-106.
2. McMahon TA, Valiant G, Frederick EC (1987) Groucho running. J Appl Physiol (1985) 62: 2326-2337.
3. Arendse RE, Noakes TD, Azevedo LB, Romanov N, Schwellnus MP, et al. (2004) Reduced eccentric loading of the knee with the pose running method. Med Sci Sports Exerc 36: 272-277.
4. Daoud AI, Geissler GJ, Wang F, Saretsky J, Daoud YA, et al. (2012) Foot strike and injury rates in endurance runners: A retrospective study. Med Sci Sports Exerc 44: 1325-1334.
5. Kleindienst FJ (2003) Gradierung funktioneller sportschuhparameter am laufschuh. Shaker, Aachen 234-235 (In German).
6. Walther M (2003) Vorfusslaufen schutzt nicht vor uberlastungsproblemen. Orthopadieschuhtechnik 6: 34 (In German).
7. Goss DL, Gross MT (2012) Relationships among self-reported shoe type, footstrike pattern, and injury incidence. US Army Med Dep J 25-30.
8. Bonnaerens S, Fiers P, Galle S, Aerts P, Frederick EC, et al. (2019) Grounded running reduces musculoskeletal loading. Med Sci Sports Exerc 51: 708-715.
9. Lussiana T, Patoz A, Gindre C, Mourot L, Hébert-Losier K (2019) The implications of time on the ground on running economy: Less is not always better. J Experi Bio 222: jeb192047.
10. Moreno FJ, Reina R, Luis V, Sabido R (2002) Visual Search Strategies in Experienced and Inexperienced Gymnastic Coaches. Percept Mot Skills 95: 901-902.
11. Gindre C, Lussiana T, Hébert-Losier K, Mourot L (2016) Aerial and terrestrial patterns: A novel approach to analyzing human running. Int J Sports Med 37: 25-29.
12. Lussiana T, Gindre C, Mourot L, Hébert-Losier K (2017) Do subjective assessments of running patterns reflect objective parameters? Eur J Sport Sci 17: 847-857.
13. Lussiana T, Gindre C (2016) Feel your stride and find your preferred running speed. Biology Open 5: 45-48.
14. Lussiana T, Gindre C, Hébert-Losier K, Sagawa Y, Gimenez P, et al. (2017) Similar Running Economy With Different Running Patterns Along the Aerial-Terrestrial Continuum. Int J Sports Physiol Perform 12: 481-489.
15. Minick KI, Kiesel KB, Burton L, Taylor A, Plisky P, et al. (2010) Interrater reliability of the functional movement screen. J Strength Cond Res 24: 479-486.
16. Smith CA, Chimera NJ, Wright NJ, Warren MJ (2013) Interrater and intrarater reliability of the functional movement screen. J Strength Cond Res 27: 982-987.
17. Onate JA, Dewey T, Kollock RO, Thomas KS, Van Lunen BL, et al. (2012) Real-time intersession and interrater reliability of the functional movement screen. J Strength Cond Res 26: 408-415.
18. Atkinson G, Nevill AM (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 26: 217-238.
19. Portney LG, Watkins MP (2009) Foundations of clinical research: Applications to practice. (3rd edtn) Pearson Prentice Hall, Upper Saddle River, New Jersey.
20. Rosner B (2010) Fundamentals of biostatistics. (7th edtn) Cengage Learning, Boston, Massachusetts
21. Hébert-Losier K, Martyn Beaven C (2014) The MARS for Squat, Countermovement, and Standing Long Jump Performance Analyses. J Strength Cond Res 28: 1849-1857.
22. Holsgaard Larsen A, Caserotti P, Puggaard L, Aagaard P (2007) Reproducibility and relationship of single-joint strength vs multi-joint strength and power in aging individuals. Scand J Med Sci Sports 17: 43-53.
23. Stokes M (1985) Reliability and repeatability of methods for measuring muscle in physiotherapy. Physio Pract 1: 71-76.

24. Zonnenberg AJJ, VanMaanen CJ, Elvers JWH, Oostendorp RAB (1996) Intra/Interrater Reliability of Measurements on Body Posture Photographs. J Crani Sleep Pract 14: 326-331.

25. Flemenbaum A, Zimmermann RL (1973) Inter- and Intra-Rater Reliability of the Brief Psychiatric Rating Scale. Psychol Reports 33: 783-792.

Author Affiliation

Top

[1]*Volodalen, Swiss SportLab, Aigle, Switzerland*

[2]*Research Unit EA 3920, Prognostic Markers and Regulatory Factors of Cardiovascular Diseases and Exercise Performance, Health, Innovation platform, University of Franche-Comté, Besançon, France*

[3]*Tomsk Polytechnic University, Tomsk, Russia*

[4]*Volodalen, Research and Development Department, Chavéria, France*