Wingate Anaerobic Test Reliability on the Velotron With Ice Hockey Players

Ryan F. Bringhurst,[1] Dale R. Wagner,[1] and Sarah Schwartz[2]

[1]Kinesiology and Health Science Department, Utah State University, Logan, Utah; and [2]Psychology Department, Utah State University, Logan, Utah

Abstract

Bringhurst, RF, Wagner, DR, and Schwartz, S. Wingate anaerobic test reliability on the velotron with ice hockey players. *J Strength Cond Res* 34(6): 1716–1722, 2020—This study evaluated the test-retest reliability of the Wingate Anaerobic Test (WAnT) performed on a Velotron electromagnetically braked cycle ergometer for power-trained athletes and assessed whether a familiarization trial was necessary to achieve high test-retest reliability. Twenty-one male ice hockey players (age 23.5 ± 4.7 years, mass 86.3 ± 16.6 kg, height 180.9 ± 7.4 cm) from a collegiate club team (Club = 10) and a recreational league (Rec = 11) performed three 30-second WAnTs within 2 weeks and with at least 24 hours between visits. Mean power (MP), anaerobic capacity, peak power (PP), anaerobic power, maximum revolutions per minute, and fatigue index were assessed. Resistance was 8.5% of the participant's body mass. The effect of time on power output was moderated ($p < 0.001$, $\eta_p^2 = 0.24$) such that a significant increase was observed after a practice trial, but not between subsequent trials for the Club players; no practice effect was observed among Rec players. Extremely high reliability ($ICC_{1,1}$) was found between trials after excluding the practice trial (MP = 0.973, anaerobic capacity = 0.975, PP = 0.957, and anaerobic power = 0.890). Club players achieved higher outputs despite no significant differences in body size or age compared with Rec players. Ice hockey players performing the 30-second WAnT on the Velotron electromagnetically braked cycle ergometer had highly reliable data, and using a familiarization trial is recommended to increase reliability and achieve higher power outputs.

Key Words anaerobic performance, cycle ergometer, muscular power, test-retest reliability, power-trained athlete

Address correspondence to Dr. Dale R. Wagner, dale.wagner@usu.edu.

34(6)/1716–1722

Journal of Strength and Conditioning Research

Introduction

The Wingate Anaerobic Test (WAnT) was developed over 40 years ago (2), yet it is still a popular test among coaches and exercise scientists. The 30-second WAnT is a maximal effort test, performed on a cycle ergometer, to measure lower-body anaerobic power. Other anaerobic tests can measure peak power (PP); these tests include the vertical jump test, standing long jump test, and Bosco repeated jumps (23). These tests, although easier to perform, are not scaled for body mass. By contrast, resistance is applied as a percentage of the subject's body mass during the WAnT (3,8,11,20), which allows for between-subject comparison of PP of athletes who differ in size and compete in different sports (1,4,6–8,20). Some of the benefits of the WAnT include monitoring chronic responses to lower-body training, creating reference norms of athletes for coaches and trainers, and assessing changes in fatigue index (FI) (23). The WAnT has been acknowledged by many as the primary method for measuring anaerobic power in study samples ranging from recreationally trained individuals to elite, power-trained athletes (1,4,6,7,11–13,20,22,23).

Dotan (7), one of the researchers who developed the WAnT in the 1970s, noted the worldwide acceptance of the test as a research and fitness-diagnostic tool yet acknowledged that technological advances such as the advent of electromagnetically braked cycle ergometers make the WAnT "ripe for an overhaul." The WAnT was originally developed using a mechanically braked ergometer (ME), so one area of uncertainty is the reliability of the WAnT when the test is performed on an electromagnetically braked ergometer (EE). A high degree of reliability for a test method is crucial because small, but meaningful, changes in performance due to an experimental manipulation cannot be detected without high test-retest reliability (9).

In a review of the WAnT, Bar-Or (3) reported test-retest reliability correlation coefficients ranging from 0.89 to 0.98 and noted that mean power (MP) tended to be somewhat more reliable than PP. The studies included in Bar-Or's review were performed with MEs. However, reliability studies of anaerobic tests using the Velotron EE are lacking. Several researchers have performed studies comparing ME with EE, including a meta-analysis of WAnT methodology

(1,3,7–9,11,12,20). One researcher stated that the ME, because of having to overcome a friction force, may underestimate power output measurements (1). Also, it is easier to administer the WAnT on an EE because less examiners are needed compared with conducting this test on an ME (7).

To our knowledge, only Astorino and Cottrell (1) have evaluated the test-retest reliability of the WAnT performed on the Velotron EE. They reported moderately high to high test-retest reliability for MP (intraclass correlation coefficient [ICC] = 0.90) and PP (ICC = 0.70); however, their sample consisted of primarily recreationally active men and women who were not specifically power trained. As shown in previous research, there is a difference in power output between anaerobically trained vs. nonanaerobically trained athletes (8,20), as well as between men and women (1,8,9,23). Therefore, because the subjects in the Astorino and Cottrell study were not anaerobically trained, and included both men and women, conclusions drawn are not valid for a power-trained sample. A practice effect was also not considered in their test-retest study, and a practice trial is recommended when testing anaerobic power (4,9,11–14,21,22). Thus, the current study is the first to assess test-retest reliability of the WAnT on an EE with power-trained athletes.

As primarily an anaerobic test, the WAnT is more useful and applicable to athletes who are anaerobically trained and compete in an anaerobic sport rather than recreationally trained individuals. Ice hockey is a sport with a high anaerobic demand, and previous investigators have demonstrated that the WAnT is highly related to on-ice skating performance in both collegiate and youth hockey athletes (10,15–19). However, these investigators used MEs for the test. Peak power during a WAnT is derived from a 5-second average on an ME, but it is recorded instantaneously on a Velotron EE (1). Therefore, despite a substantial amount of previous WAnT reference values for ice hockey players, there are no published standards for this athletic population when tested on a Velotron EE. The purpose of this study was to evaluate the test-retest reliability of the WAnT performed on a Velotron EE for power-trained athletes and to determine whether a familiarization trial was necessary to achieve high test-retest reliability. An additional objective was to compare the power outputs of ice hockey players from a nationally ranked collegiate club team to those from an adult recreation league. We hypothesized that high reliability of the WAnT would be achieved on the Velotron EE for all participants, and that the power outputs of the ice hockey players in our sample would exceed those of previously published reports because of the instantaneous measurement of the EE compared with the 5-second average measurement of the ME.

Methods

Experimental Approach to the Problem

An observational approach with repeated measurements was used for this study. Each participant visited the exercise physiology laboratory 3 times within 2 weeks and with at least 24 hours between visits. This design allowed for determining reliability of the EE and ascertaining any practice effect using repeated-measures analysis of variance (ANOVA).

Subjects

Initially, 25 ice hockey players from the Utah State University club team (Club) and 20 ice hockey players from an adult recreation league (Rec) were invited to participate in the study. A total of 11 from the Club completed a consent form, but only 10 completed all 3 trials. A total of 12 Rec players completed a consent form, but only 11 players completed all 3 trials.

Demographic characteristics of the sample are in Table 1. All participants were at least 18 years of age. Each participant was informed of the benefits and risks of the investigation before signing an institutionally approved informed consent document. In addition, a Physical Activity Readiness Questionnaire (PAR-Q) was completed before participation. The study was approved by the University's Institutional Review Board (Utah State University, protocol #7406).

Table 1. Demographics of study sample.*†

Groups	Total	Club	Rec	sig
n	21	10	11	
Age (yrs)	23.5 ± 4.7	22.0 ± 1.1	24.8 ± 6.3	0.180
Body mass (kg)	86.3 ± 16.6	85.4 ± 11.6	87.1 ± 20.7	0.821
Height (cm)	180.9 ± 7.4	181.7 ± 8.0	180.1 ± 7.2	0.629
Yrs exp‡	13.4 ± 5.9	17.2 ± 3.5	10.0 ± 5.5	0.002‡
BMI ($kg \cdot m^{-2}$)	26.3 ± 4.0	25.7 ± 1.9	26.7 ± 5.4	0.573

*BMI = body mass index [kilograms (kg) per meters squared (m^2)]; n = sample size; Yrs exp = years' experience.
†Values in cells represent mean ± *SD*, significance based on independent group *t*-test, Levene's test used to determine whether to assume equality of variance.
‡Significant difference ($\alpha < 0.05$).

Procedures

All participants performed the 30-second WAnT each visit on the EE Velotron Dynafit Pro cycle ergometer (RacerMate, Seattle, WA, USA) with a 62-tooth chainring. During the initial visit, height was measured using a wall-mounted stadiometer (Seca 216; Seca Corp., Ontario, CA, USA), and mass was measured using a digital scale (Seca 869; Seca Corp.,). Each participant self-selected seat height, seat setback, handlebar height, and reach on the Velotron and were recorded and used for the subsequent trials. To try to improve reliability and avoid a "practice effect" (4,9,11–14,21,22), the first visit served as a practice trial to familiarize each subject with the Velotron and the 30-second WAnT test. To not influence or bias their effort, participants were not informed of the purpose of the first trial. Previous investigators have recommended that a practice trial is necessary for reliable WAnT data (4,11,14,22).

In an effort to minimize extraneous influences, each participant completed his 3 trials at the same time of day (±1 hour), and they confirmed that no strenuous workouts beyond their normal training occurred in the 24 hours that preceded each testing session. In addition, participants were instructed to maintain their normal diet and not eat within 1 hour of each testing session. Furthermore, the barometric pressure and temperature within the laboratory were similar across all trials. Five of the 10 Club players completed their test sessions during the hockey season, whereas the other 5 did their tests immediately after season. All Rec players completed their sessions in season.

Each visit lasted a maximum of 15 minutes and consisted of (a) review of testing procedures, (b) 5-minute warm-up at a resistance of 75 W and a cadence of 60–100 rpms, (c) 3-minute rest before test start, (d) 30-second WAnT, and (e) cool down until subject's heart rate had returned to 120 $b \cdot min^{-1}$. Strong verbal encouragement was given throughout the 30-second protocol and was similar for all trials and all participants. The WAnT was performed with a resistance of 8.5% body mass because previous investigators have determined this to be the optimal load when testing power-trained male athletes (6,8). The test was performed using a traditional flying start with the participants given a 20-second warm-up followed by 6 seconds of acceleration to achieve maximal rpms before the load was applied and the 30-second WAnT commenced (1).

Statistical Analyses

All analyses were performed using SPSS version 24 (IBM, Inc., Armonk, NY, USA). Statistical significance was accepted at $p \leq 0.05$ unless otherwise stated. Mean power was defined as average power output in Watts (W) over the 30-second test and anaerobic capacity as MP per kilogram of body mass ($W \cdot kg^{-1}$). Peak power was defined as the highest instantaneous power output achieved in Watts (W) and anaerobic power as PP per kilogram of body mass ($W \cdot kg^{-1}$). Maximum revolutions per minute (RPM_{max}) was the highest instantaneous pedaling cadence, and FI was calculated as FI = [(PP − Min Power)/PP] × 100, where Min Power is minimum power. The preliminary trial (Prelim) was the practice trial, and the subsequent trials were labeled trial 1 (T1) and trial 2 (T2). Mean and *SD* were calculated for MP, anaerobic capacity, PP, anaerobic power, RPM_{max}, and FI.

Before statistical analysis, assumptions were tested: Shapiro-Wilk test for normality and Mauchly's test for sphericity. Two-way mixed design ANOVA tested for differences in mean with time (Prelim, T1, T2) as the within-subject factor and team type (Club vs. Rec) as the between-subject factor. This was accomplished by including the main effects for time and team type, as well as their interaction. In the presence of a significant interaction, mean were compared with pairwise *t*-tests using the Bonferroni correction for multiple comparisons. A previous power analysis estimated a total sample of 28 would provide 80% power to detect a moderate ($f = 0.25$) interaction effect in the presence of moderate correlation ($r = 0.50$) between repeated measures.

Test-retest reliability was evaluated using the statistical methods recommended by Hopkins et al. (9) and Weir (21). These include evaluation of the intraclass

Table 2. Reliability with and without Prelim Trial for all variables.*†‡

	All 3 trials		Without prelim trial			
	η_p^2	$ICC_{1,1}$ (sig)	$ICC_{1,1}$ (sig)	*SEM*	MD	CV (%)
MP	0.234	0.829 (<0.0001)	0.973 (<0.0001)	16.98	47.00	13.3
Anaerobic capacity	0.235	0.836 (<0.0001)	0.975 (<0.0001)	0.19	0.55	13.8
PP	0.249	0.847 (<0.0001)	0.957 (<0.0001)	45.15	125.15	17.4
Anaerobic power	0.246	0.344 (0.124)	0.890 (<0.0001)	0.54	1.49	11.1
RPM_{max}	0.242	0.350 (0.118)	0.890 (<0.0001)	6.50	18.02	11.2
FI	0.031	0.701 (0.0005)	0.797 (0.0003)	3.18	8.81	11.7

*ICC = intraclass correlation coefficient; MD = minimal difference; CV = coefficient of variation; MP = mean power; PP = peak power; RPM = revolutions per minute; FI = fatigue index.
†η_p^2 = partial eta squared for the interaction between time and team in the RM ANOVA.
‡ICC > 0.75 was considered good, *SEM* and MD calculated using ICC without prelim trial using total *SD* of all participants for T1 and T2.

correlation coefficient ($ICC_{1,1}$ one-way random), *SEM* [$SEM = SD\sqrt{(1-ICC)}$], minimal difference ($MD = SEM \times 1.96 \times \sqrt{2}$), and coefficient of variation (CV).

Results

The ICCs between T1 and T2 were very high and significant for all variables, and, except for PP, the CVs were between 11.1 and 13.8% (Table 2). Normality was determined for all variables except FI. The 2-way mixed design ANOVA of PP used the Greenhouse-Geisser correction for degrees of freedom ($\varepsilon = 0.732$) because the assumption of sphericity was violated, Mauchly's $\omega = 0.635 \sim \chi^2(2) = 8.182$, $p = 0.017$. The interaction between time and team type was found to be significant, F (1.465, 27.833) = 6.286, $p = 0.010$, $\eta_p^2 = 0.249$. Visual inspection of Figure 1 (panel A) reveals the mean for Club members increased from Prelim to T1 but remained stable between T1 and T2, whereas the Rec players remained constant across all 3 trials at a PP similar to the Prelim for Club players. Post hoc analysis concluded Club players increased by an average of 226.1

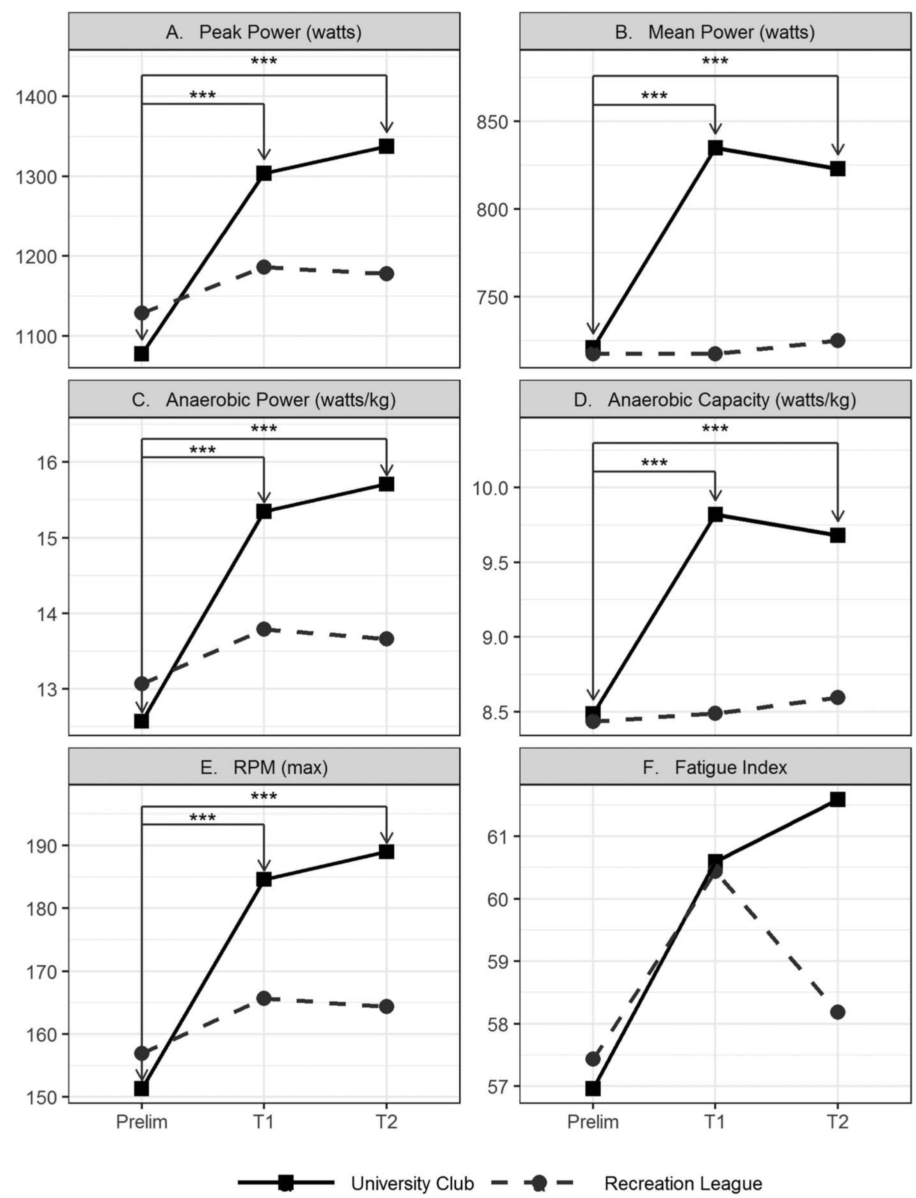

Figure 1. Mean from the 2-way mixed design ANOVA models of each of the 6 outcome measures. ***Indicates Bonferroni-corrected post hoc p value < 0.001. RPM = revolutions per minute.

Table 3. Descriptive Wingate (WAnT) data from trial 1.*†

	Total	Club	Rec	sig
n	21	10	11	
MP (W)	773 ± 107	834 ± 103	718 ± 79	0.008‡
Anaerobic capacity ($W\cdot kg^{-1}$)	9.12 ± 1.26	9.82 ± 0.73	8.49 ± 1.34	0.012‡
PP (W)	1,242 ± 206	1,303 ± 163	1,186 ± 231	0.197
Anaerobic power ($W\cdot kg^{-1}$)	14.50 ± 1.54	15.35 ± 1.37	13.79 ± 1.33	0.016‡
RPM_{max}	175 ± 18.6	184 ± 16.5	166 ± 16.0	0.015‡
FI	60.5 ± 7.4	60.6 ± 7.3	60.4 ± 6.8	0.960

*MP = mean power; PP = peak power; RPM = revolutions per minute; FI = fatigue index.
†*n* = sample size. Value in cells represents M ± *SD*, significance based on independent group *t*-test without controlling for multiple comparisons; Levene's test used to determine whether to assume equality of variance.
‡Significant difference ($\alpha < 0.05$).

W (*SE* = 51.11) in PP from Prelim to T1, ($p = 0.001$), but PP did not increase from T1 to T2, ($p = 0.760$). Peak power did not increase in Rec players across the 3 trials ($p = 0.763$). Similar analyses were conducted for MP, anaerobic capacity, anaerobic power, RPM_{max}, and FI with nearly identical interactive effects. The only exception was for FI, in which none of the pairwise post hoc comparisons reached significance (Figure 1).

Post hoc comparisons using pairwise *t*-tests showed no significant difference between T1 and T2 for any outcomes but did, however, reveal the Prelim values to be significantly less than values obtained during T1 and T2 trials among club players for all variables other than FI. Table 3 displays the summary statistics for T1; as there were no significant differences between T1 and T2, T2 data are not included. Results further demonstrate that MP, anaerobic capacity, anaerobic power, and RPM_{max} measurements were significantly higher for Club players compared with Rec players (Table 3). The Club players achieved these higher power outputs despite no significant differences with the Rec players regarding body size or age (Table 1). The only descriptive factor that was significantly different between the 2 groups was years of experience ($\alpha = 0.002$), with a higher average for Club players.

Discussion

The purpose of this study was to evaluate the test-retest reliability of the WAnT performed on a Velotron EE with power-trained athletes. Furthermore, we aimed to determine whether a practice trial was necessary for achieving high reliability. In addition, descriptive power output data specific to ice hockey players were obtained. This was important because previous WAnT data for this athletic group were gathered from tests performed on MEs (10,15–19); this is the first study to report WAnT data for ice hockey players using an EE.

Based on current findings, the Velotron racermate EE is a reliable method for testing anaerobic power. Previous researchers have described WAnT reliability using MEs (3,7,14), but reliability studies using an EE are limited. To our knowledge, only one (1) has evaluated the test-retest reliability of the WAnT using the Velotron; however, their sample consisted of recreationally active men and women who were not specifically power trained. Also, a practice effect was not considered in their test-retest study. They reported ICCs of 0.70 and 0.90 for anaerobic power and anaerobic capacity, respectively. These are slightly less than the ICCs observed in this study between trials T1 and T2. They reported small MDs of 0.44 $W\cdot kg^{-1}$ for anaerobic power and 0.11 $W\cdot kg^{-1}$ for anaerobic capacity. However, our sample produced substantially higher power outputs than the participants in the Astorino and Cottrell study, which may contribute to the larger MDs in this study. Despite similar peak cadences of 181 rpm for Astorino and Cottrell's participants and 175 rpm for our athletes, the anaerobic power of 14.5 $W\cdot kg^{-1}$ of the hockey players in this study was substantially greater than the 9.7–9.8 $W\cdot kg^{-1}$ of the recreationally active participants in the Astorino and Cottrell study. Their CVs of 13.7% for anaerobic power and 8.9% for anaerobic capacity are comparable to ours of 11.1 and 13.8%. However, it is an inequitable comparison because of sex and sample size differences between studies. In addition, the current study was of power-trained athletes, whereas the study of Astorino and Cottrell included only recreationally active participants. The high ICCs and low MDs reported by Astorino and Cottrell are at least partially attributed to the lack of anaerobic power observed in their participants, which averaged nearly 5 $W\cdot kg^{-1}$ less than those in the current study. With substantially smaller power outputs, it is logical that variability would also be less, resulting in high test-retest reliability and low MDs.

Having the athletes complete a practice trial clearly improved the test-retest reliability of the entire sample. However, it is interesting that this improvement was observed primarily in the more powerful Club players and not in the Rec players. Astorino and Cottrell (1) achieved high reliability without a practice trial in their sample of recreationally active participants. By contrast, several research teams that have tested the reliability of WAnT using ME have reported that a practice trial is necessary (4,14). Barfield et al. (4) described improvements of 14 and 6% for anaerobic power and anaerobic capacity, respectively, for 2

WAnTs performed by non–specifically-trained college-aged men separated by a week. Similarly, Ozkaya (14) reported improvements in anaerobic power of 20% and anaerobic capacity of 6% for repeat WAnTs of 20 men who were participating in power sports. Our findings of improved reliability with a practice trial agree with the recommendations of many others to include a practice trial when testing anaerobic power (4,9,11–14,21).

Club ice hockey players had significantly higher MP, anaerobic capacity, anaerobic power, and RPM_{max} compared with the Rec players. The Club team participated in the national championship tournament for this level of play (American Collegiate Hockey Association). Thus, given their higher level of competition, higher training intensity, and more years of experience, it is not surprising that they had higher WAnT data than the Rec players of similar age and body mass index. The power output data of both the Club and Rec players in this study is considerably higher than the power output data reported for similar WAnT studies performed on EE (1,12); however, these researchers reported their participants as being "physically active" rather than power trained. Compared with other power-trained men, the anaerobic power and anaerobic capacity of 13.8 and 8.5 $W \cdot kg^{-1}$, respectively, for the Rec players is almost identical to the 13.9 and 8.5 $W \cdot kg^{-1}$ recently reported for CrossFit-trained men who performed the WAnT on the same Velotron (5).

Compared with data of other ice hockey players, the PP and anaerobic power of the athletes in this study were comparable with the PP (1,306 W) and anaerobic power (14.7 $W \cdot kg^{-1}$) reported for members of an NCAA Division I national runner-up team (16) and higher than that reported for other NCAA Division I players (PP = 1,112 W) (15) and Division III players (anaerobic power = 11.35 $W \cdot kg^{-1}$) (10). Even the Rec players had PP and anaerobic power higher than the under-20 Polish National Team of 1,031 W and 12.97 $W \cdot kg^{-1}$, respectively (18). How are these high-power outputs possible for participants that clearly have less skill and realistically less anaerobic power than national team athletes? All the ice hockey comparison studies cited (10,15,16,18) were performed on ME. As Micklewright et al. (12) described, results obtained on an EE are not comparable with those from an ME because of mechanical differences between ergometers. Such differences include the inertia of the flywheel, load applied mechanically vs. electronically, and PP identified as the highest value attained during the test on the Velotron rather than a 5-second average on an ME (1). Wingate Anaerobic Test reference values for male power athletes are available for tests performed on an ME (6,23). However, these reference values are not applicable for tests performed on an EE. There are no reference values or normative tables for WAnT data from an EE, but the data provided in this study provide some reference point for future studies of power-trained men tested on a Velotron EE.

In conclusion, ice hockey players performing the 30-second WAnT on the Velotron EE had highly reliable data. Reliability increased with a familiarization trial for all participants. The Club players had greater power outputs than the Rec players. Also, power outputs increased significantly after the familiarization trial for Club team, but not for the Rec players. Unfortunately, normative WAnT data specific to tests conducted on EE do not yet exist, and comparing WAnT data from an EE to normative data created from tests conducted on ME is not appropriate.

Practical Applications

A practice trial is warranted when trying to obtain anaerobic power outputs of power-trained athletes. Some athletes will be able to achieve a higher pedaling cadence, and consequently higher power outputs, after a familiarization trial. A single practice trial followed by 1 WAnT is sufficient because multiple WAnTs do not result in higher power outputs; after the initial familiarization trial, the WAnT is reliable. In addition, athletes and coaches should not evaluate power outputs obtained on a Velotron EE against WAnT norms. The WAnT norms were created from tests performed on an ME with a 5-second average to calculate the power output. The Velotron software reports PP as the highest observed value rather than the highest 5-second average. Norms for WAnT performed on an EE do not yet exist, but data from this study can be used as reference values for hockey coaches who are using an EE rather than an ME for Wingate testing.

Acknowledgments

The authors thank the hockey athletes who participated in this study and the Kinesiology and Health Science Department at Utah State University who provided a small financial incentive for the study participants. The authors have no conflicts of interest to report. The results of this study do not constitute endorsement of the product by the authors or the NSCA.

References

1. Astorino, TA and Cottrell, T. Reliability and validity of the Velotron Racermate cycle ergometer to measure anaerobic power. *Int J Sports Med* 33: 205–210, 2012.
2. Ayalon, A, Inbar, O, and Bar-Or, O. Relationships among measurements of explosive strength and anaerobic power. *Int Ser Sport Sci Biomech* 4: 572–577, 1974.
3. Bar-Or, O. The Wingate anaerobic test: An update on methodology, reliability, and validity. *Sports Med* 4: 381–394, 1987.
4. Barfield, JP, Sells, PD, Rowe, DA, and Hannigan-Downs, K. Practice effect of the Wingate anaerobic test. *J Strength Cond Res* 16: 472–473, 2002.
5. Clark, NW, Wagner, DR, and Heath, EM. Influence of Velotron chainring size on Wingate anaerobic test. *J Sci Med Sport* 21: 202–206, 2018.
6. Coppin, E, Heath, EM, Bressel, E, and Wagner, DR. Wingate anaerobic test reference values for male power athletes. *Int J Sports Physiol Perform* 7: 232–236, 2012.
7. Dotan, R. The Wingate anaerobic test's past and future and the compatibility of mechanically versus electro-magnetically braked cycle-ergometers. *Eur J Appl Physiol* 98: 113–116, 2006.

8. Dotan, R and Bar-Or, O. Load optimization for the Wingate anaerobic test. *Eur J Appl Physiol* 51: 409–417, 1983.

9. Hopkins, WG, Schabort, EJ, and Hawley, JA. Reliability of power in physical performance tests. *Sports Med* 31: 211–234, 2001.

10. Janot, JM, Beltz, NM, and Dalleck, LD. Multiple off-ice performance variables predict on-ice skating performance in male and female Division III ice hockey players. *J Sports Sci Med* 14: 522–529, 2015.

11. Malone, JK, Blake, C, and Caulfield, B. Test-retest reliability of the 30-sec Wingate cycle test in a trained male cohort. *Isokinet Exerc Sci* 22: 251–258, 2014.

12. Micklewright, D, Alkhatib, A, and Beneke, R. Mechanically versus electromagnetically braked cycle ergometer: Performance and energy cost of the wingate anaerobic test. *Eur J Appl Physiol* 96: 748–751, 2006.

13. Noreen, EE, Yamamoto, K, and Clair, K. The Reliability of a simulated uphill time trial using the Velotron electronic bicycle ergometer. *Eur J Appl Physiol* 110: 499–506, 2010.

14. Ozkaya, O. Familiarization effects of an elliptical all-out test and the Wingate test based on mechanical power indices. *J Sports Sci Med* 12: 521–525, 2013.

15. Peterson, BJ, Fitzgerald, JS, Dietz, CC, Ziegler, KS, and Ingraham, SJ. Division I hockey players generate more power than Division III players during on- and off-ice performance tests. *J Strength Cond Res* 29: 1191–1196, 2015.

16. Potteiger, JA, Smith, DL, Maier, ML, and Foster, TS. Relationship between body composition, leg strength, anaerobic power, and on-ice skating performance in Division I men's hockey athletes. *J Strength Cond Res* 24: 1755–1762, 2010.

17. Power, A, Faught, BE, Przysucha, E, McPherson, M, and Montelpare, W. Establishing the test-retest reliability and concurrent validity for the repeat ice skating test (RIST) in adolescent male ice hockey players. *Meas Phys Educ Exerc Sci* 16: 69–80, 2012.

18. Roczniok, R, Maszczyk, A, Czuba, M, Stanula, A, Przemyslaw, P, and Gabrys, T. The predictive value of on-ice special tests in relation to various indexes of aerobic and anaerobic capacity in ice hockey players. *Hum Mov* 13: 28–32, 2012.

19. Skowronek, T, Socha, T, Roczniok, R, and Socha, S. The predictive value of various anaerobic capacity indices in relation to specific on-ice performance tests in ice hockey players. *Life Sci J* 10: 2828–2832, 2013.

20. Vargas, NT, Robergs, RA, and Klopp, DM. Optimal loads for a 30-s maximal power cycle ergometer test using a stationary start. *Eur J Appl Physiol* 115: 1087–1095, 2015.

21. Weir, JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 19: 231–240, 2005.

22. Zavorsky, GS, Murias, JM, Gow, J, Kim, DJ, Poulin-Harnois, C, Kubow, S, and Lands, L. Laboratory 20-km cycle time trial reproducibility. *Int J Sports Med* 28: 743–748, 2007.

23. Zupan, MF, Arata, AW, Dawson, LH, Wile, AL, Payn, TL, and Hannon, ME. Wingate anaerobic test peak power and anaerobic capacity classifications for men and women intercollegiate athletes. *J Strength Cond Res* 23: 2598–2604, 2009.