

Research article

Analysis of Pushing Forces During the Water Polo Eggbeater: Reliability and Validity of a Novel Approach

Félix Croteau^{1,2,3}✉, Elouan Lanoé^{1,4} and Julien Clément^{1,2,4}

¹ Institut national du sport du Québec, Montréal, Canada; ² Water Polo Canada, Montréal, Canada; ³ McGill University, School of Physical and Occupational Therapy, Montréal, Canada; ⁴ École de Technologie Supérieure, Département de Génie des systèmes, Montréal, Canada

Abstract

Water polo players benefit from greater odds of success when maintaining their tactical position against their opponents. This study evaluated the reliability and validity of a water-based resistance test to replicate this skill. Thirty-three water polo players participated in this study (19 males and 14 females, 14 from senior and 19 from junior national teams). Data were collected during two regular training sessions, separated by one week, using a load cell to instrument a weight stack resistance setup on the pool deck. Performance parameters such as mean force, maximum force, mean peak force and total impulse were defined with custom Python scripts. Test-retest reliability was assessed using intra-class correlations (ICC3,1). Group comparisons were explored between male and female players. Level of significance was set at $p < 0.05$. The reliability findings were high to very high for the mean force, maximum force, mean peak force, inter-stroke range, and total impulse (ICC 0.85 - 0.93, $p < 0.01$). Group comparisons showed significantly greater values in male players for these variables ($p < 0.01$, ES = 1.05 - 9.36) with large to very large effect sizes. However, there was no significant difference in endurance measured between sexes ($p = 0.88$, ES = 0.04). This study presents a methodology with satisfactory metrological qualities for field applications using simple and affordable equipment. The testing apparatus presented in this study can readily be replicated in a variety of training environments by practitioners working with water polo teams. Coaches can use this approach to evaluate individual player progress or to compare performance across a group of water polo players.

Key words: Aquatic, strength, swimming, performance analysis.

Introduction

Water polo player positions include goalkeepers and various field positions, such as drivers, centers, and defenders (Smith, 1998). The last two provide a key tactical contribution to the team by occupying the space immediately in front of the goalkeeper cage, which yields a much higher scoring chance than shooting from the periphery (Perazzetti et al., 2023a). A recent study of over 5000 goals from NCAA matches concluded that for every meter further from the center position, the odds of scoring decreased by 29% (Gullikson et al., 2020). Therefore, the ability to maintain this position against the push-back from the defensive team separates successful center forwards from the others (Botonis et al., 2019). From a defensive perspective, the ability to push these invading players away from the goalkeeper cage is likewise a key tactical advantage (Lupo

et al., 2016; Perazzetti et al., 2023b). Unfortunately, land-based measures of lower extremity dynamic strength show very low correlation with water-based abilities in water polo players (Platanou, 2005). Indeed, water polo players produce this type of movement with a different action than dryland jumping using the eggbeater kick, by “alternating the circular and continuous movements of the legs, producing an upward force and maintaining players afloat in a vertical position” (Uljevic et al., 2013). Therefore, sport-specific evaluations are required.

While the ability to push opponents is closely associated with players’ strength, there are limited sport-specific approaches to assess these qualities in the daily training environment (Sanders, 1999a). Often, the description of leg strength relies on the ability to jump high out of the water (Gobbi et al., 2013; Platanou and Varamenti, 2011) or to swim across the pool in an upright position with the arms out of the water as quickly as possible (McCluskey et al., 2010). Alternately, players have been tested by measuring time to swimming upright exhaustion with gradual increasing weighed vests (Melchiorri et al., 2015). This method allowed more rigor in the execution of the task and showed that expert players were able to sustain these loads for significant more time than non-expert water polo players ($p < 0.01$, effect size 3.5 - 6.0). However, the setup for this experiment requires the players to be upright and are not specific to the ability to push an opponent in the water in a horizontal plane. Instead, other authors have described stationary swimming tests where athletes swim away from a load cell while being attached with elastic tethers to the edge of the pool (Muniz-Pardos et al., 2019). The amount of resistance provided with elastic tethers in such a setup is not continuous, as it increases proportionally to the distance away from the starting point. Therefore, analyses of the forces expressed by participants against this type of resistance cannot inform coaches about the cyclical patterns employed to accomplish the push (Bratuša and Dopsaj, 2016; Dopsaj et al., 2003). Nevertheless, it has been used successfully in swimmers to evaluate mean force and maximum force in 10-, 30- or 60-seconds challenges (Dopsaj, 2010; Dopsaj et al., 2003; Joaquim Baratto de Azevedo et al., 2021). In water polo, a comparable approach was used with 28 youth players who were attached at the end of a non-elastic rope to a fixed point and instructed to perform 10-second maximal swim tests using the lower body only (Stirn et al., 2014). This method provides a resulting force curve to analyze for every individual with which the

authors further proposed that alternating eggbeater motion yielded greater average force compared with simultaneous motion ($128 \pm 26\text{N}$ vs $111 \pm 22\text{N}$). The greatest maximal forces were recorded with the simultaneous motion however ($244 \pm 37\text{N}$ vs $189 \pm 36\text{N}$), which has implications for using the optimal movement based on the duration of the task. This approach still cannot directly discriminate between participants' performance at pushing the same amount of resistance.

In training, coaches also use the resistance provided by a stationary weight stack or containers filled with water tied with a long rope to provide a constant resistance against the players in the water (Muniz-Pardos et al., 2019). In theory, this type of resistance would simulate more closely the sustained task requirements of pushing an opponent in the water. Consequently, we propose a method to measure the forces generated while performing an eggbeater kick by instrumenting this type of apparatus. By embedding the instrument within training equipment, coaches can readily repeat these measurements and evaluate the changes obtained from each training cycle (Abernethy et al., 1995).

The main goal of this study was to develop and validate a method to assess sport-specific pushing forces in water polo players. The hypotheses related to these metrological properties were that: (1) reliability should be sufficient for the appropriate implementation of this procedure in the daily training environment (intra-class correlation >0.80) and (2) male players would demonstrate greater strength compared to the female players.

Methods

Subjects

A total of 33 water polo players participated in this study. The male group included 19 players (3 international level and 16 national level), whereas the female level group included 14 players (11 international and 3 national level)

(McKay et al., 2022). The mean age, height and weight were 19.5 ± 2.8 years, $187.1 \pm 5.1\text{cm}$ and $89.1 \pm 13.1\text{Kg}$ for the male group (23.4 ± 3.6 years, $174.0 \pm 5.5\text{cm}$ and $75.2 \pm 12.7\text{Kg}$ for the female group). Anthropometric measurements were obtained prior to participation in the study in accordance with the standards from the International Society for Advancement of Kinanthropometry (International Society for Advancement of, 2001). All participants trained more than five days per week, reported no training restrictions or injury, and were part of either the junior or senior Canadian national water polo teams. Data was collected as part of regular training activities, and players provided their written consent for anonymized data to be included in research. The protocol was approved by the ethics board of École de Technologie Supérieure de Montréal in accordance with the principles of the Declaration of Helsinki (case number H20230401).

Procedures

One weight stack (Adjustable dual pulley system, Atlantis Strength Inc., Canada) with a four-pulley system was instrumented with a waterproof load cell (STS-1000, Chatillon Ametek, Florida, USA) with a sampling rate of 100 Hz and a maximum capacity of 1000 lbf. This portable and cost-efficient measurement device could be attached to any other resistance equipment used by a water polo club. A non-elastic rope of 10 m was attached to the load cell and connected to a thick rubber loop at its extremity (Figure 1). The alignment of the cord was such that the slope was near 0° , as previous authors have shown that varying the slope angle can change the resulting performance measures (Joaquim Baratto de Azevedo et al., 2021). In each instance, the tests were filmed with an underwater view with a Canon VIXIA camera model HFG20 at 60 Hz frame rate fixed to a metal pole on a rolling cart. The camera was positioned 1.5 m perpendicularly away from the participants at a depth of 1.0 m underwater. The video footage was added to provide experimental context, but not analyzed for further kinematic insights.

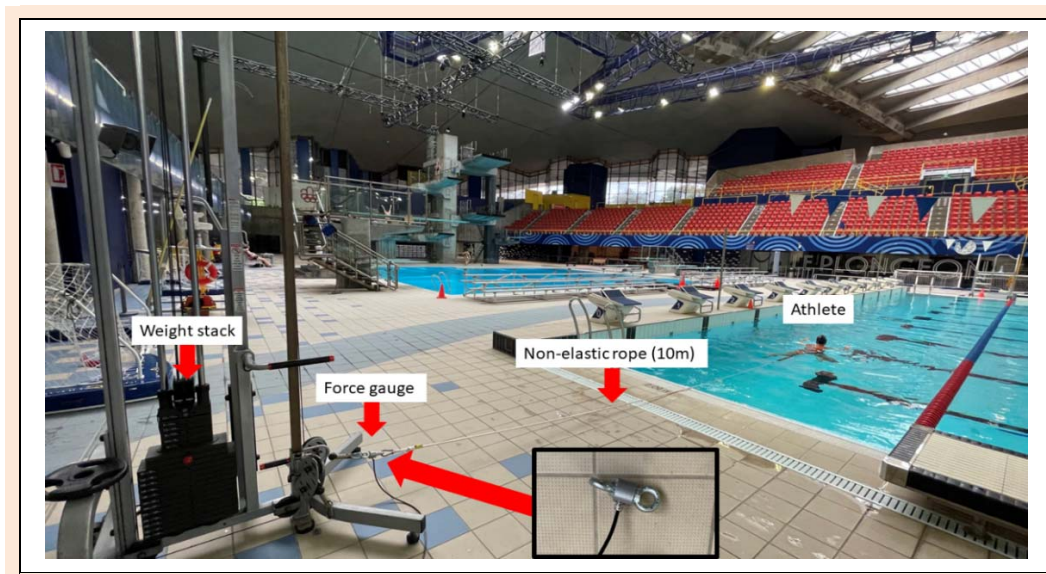


Figure 1. Experimental setup with weight stack resistance attached to the load cell to record force from participants.

Data was collected during regular training sessions after the players performed a typical water polo warmup (freestyle swimming first, then throwing activities for approximately 10 minutes and 15 minutes respectively). A subset of 27 players were tested a second time with a target of seven days between measurements (six players were unavailable for re-testing due to injury or travel commitments, 4 male and 2 female). The reliability subset of 27 players included 12 international players and 15 national level players. On each occasion, participants began the test in the water with the strap worn across the chest and the rope near its taut position, with the arms facing forward and their hands out of the water. The players' bodies were in a horizontal plane, parallel with the surface of the water, facing directly away from the weight stack (Figure 1). Coaches provided the mean weight resistance used by both the male and female teams in training, which was used as a standard resistance for each sex. A sex-specific weight resistance was chosen to reduce the bias from known differences in strength between male and female water polo players (Croteau et al., 2021).

Once in position, players were instructed to “push as far as they can without backing down” for ten seconds against a standard weight stack resistance (54.4Kg for women and 81.6Kg for men) using an alternating eggbeater kick with their hands outside of the water. The examiner used a commercial chronometer from a smartphone and the countdown was shouted out loud for the participants to hear. This duration was chosen based on coaches' experience of the typical time during which players are required to maintain position in front of the goalkeeper's cage, as well as the physiological parameters for strength evaluation in water polo (Keiner et al., 2020). This was reflected in studies of time-motion analysis of water polo as well, where the mean contact time between male players was $9.8 \pm 3.4s$ (Platanou, 2004; Platanou and Geladas, 2006). It is also the same duration reported in similar studies of water polo players, therefore allowing comparison with previous authors (Dopsaj, 2010; Joaquim Baratto de Azevedo et al., 2021; Yanagi, 1995).

Data processing

Raw data were extracted from the load cell with the ForceTest 3.1 software (Ametek, USA) and analyzed using a custom Python script. Data were filtered using the Savitzky-Golay algorithm (window = 21 and polynomial = 2). Force peaks were identified with the function *find_peaks* (prominence = 0, distance = 20, height = Fset-point/2, threshold = None, width = 10) included in *ScyPi* library. The start point was defined as the moment when the force became greater than half of the force set point. This removed the portion of the task where the elastic components of the setup were deformed (i.e., cord and rubber loop). The end of the measure was defined at the last force peak before the cessation of the experiment. Sometimes players performed further kicks after the end of the test or instead stopped too soon. To prevent these extra kicks to bias the results, we arbitrarily chose, among the last three force peak, the last one that did not differ by more than $\pm 20\%$ from the previous force peak. Thus, we eliminated force peaks with an excessive variation (e.g., a double kick

at the end of the task) and included only relevant strokes in the test measures. The force set point corresponds to the actual resistance from the weight stack that the athletes must overcome after considering the pulley system. It is expressed in Newtons from the resistance in pounds and then divided by four (number of pulleys in this weight stack):

$$F_{set\ point\ (N)} = \frac{F_{resistance\ (lbs)} \times 0.4535923\ (kg/lbs) \times 9.81\ (m/s^2)}{4}$$

The filtered signal was used to calculate overall mean force, maximum force, mean peak force, inter-stroke force range, mean stroke duration and variability, number of strokes, total impulse, endurance index, and time to first peak. The endurance index represents a ratio between the total impulse performed during the task and an ideal scenario in which a player would have been able to maintain the set point force for the entire test duration.

Statistical analysis

Reliability was assessed through intra-class correlation (ICC3,1) with a two-way mixed effects and absolute agreement (Currell and Jeukendrup, 2008). Reliability was described as very high ($ICC > 0.90$), high ($0.70 < ICC < 0.89$), or moderate ($0.50 < ICC < 0.69$) (Plichta and Kelvin, 2012). Bland-Altman plots were drawn to illustrate bias and limits of agreement between the two sessions. The standard error of measurement (SEM) (Joaquim Baratto de Azevedo et al., 2021) was computed as:

$$SEM = SD \times \sqrt{1 - ICC}$$

Construct validity was assessed by comparing players from different sexes (male and female) (McKay et al., 2022) using independent t-tests and effect size (ES) calculations with Hedge's *g* correction for the group mean scores (Hays and Reeve, 2008). The same comparisons were also made with results relative to body weight to reduce systematic bias from participant physical size (Croteau et al., 2021). These calculations were only performed on variables that minimally showed high to very high intra-class correlations to limit spurious findings. Effect sizes were interpreted as trivial (<0.2), small (<0.5), moderate (<0.8) or large (≥ 0.8) (Cohen, 2013). Normality of the data distribution was assessed for each variable using Kernel density plots and the Shapiro-Wilk test (Evans, 1996). The complete statistical analysis was conducted in R (version 4.1.0) (Team, 2022). All results are expressed with 95%CI and significance level set to $p = 0.05$.

Results

On average, the second test was performed 8.0 ± 3.1 days after the initial session. Overall mean force, maximum force, mean peak force, inter-stroke force range, total impulse, and endurance index showed a high level of reliability ($ICC > 0.74$, $p < 0.01$) (Table 1). Bland-Altman

analyses showed small bias between repetitions for these variables, with an average relative bias of 4.05% (Figure 2). SEM are further described in Table 1 to provide absolute measures of error for each variable. Compared to the female players, the male group was composed of a significantly greater number of national level players ($\chi^2(1) = 10.56$, $p < 0.01$), and therefore was significantly younger ($W = 50$, $p < 0.01$) but also both taller ($W = 275$, $p < 0.01$) and heavier ($W = 216$, $p < 0.01$). Group comparisons iden-

tified significantly greater values of large to very large effect size in male players for mean force, maximum force, mean peak force, inter-stroke range and total impulse ($p < 0.01$, $ES = 1.05-9.36$). However, the endurance index showed no difference between both groups ($p = 0.88$, $ES = 0.04$). All comparisons remained significant when evaluating relative strength measurements ($p < 0.01$ to 0.04), however effect sizes between groups were smaller ($ES = 0.71 - 1.22$). See Table 2 for full list of comparisons.

Table 1. Test-retest reliability analysis (two-way mixed effects, absolute agreement) and Bland-Altman results

Variable	ICC3,1	95%CI	Sig	SEM	Bias	Relative bias (%)	Lower LOA	Upper LOA
Overall mean force (N)	0.91	[0.85-0.95]	0.00	7.81	9.71 (10.68)	6.66	-11.23	30.64
Max Force (N)	0.93	[0.88-0.96]	0.00	7.93	3.98 (11.15)	2.25	-17.87	25.83
Mean peak force (N)	0.92	[0.86-0.96]	0.00	8.31	9.18 (11.40)	5.83	-13.16	31.51
Inter-stroke range (N)	0.85	[0.74-0.91]	0.00	4.47	-0.13 (6.37)	0.60	-12.62	12.36
Mean stroke duration (s)	0.38	[0.10-0.60]	0.01	0.03	2.85×10^{-3} (0.04)	0.06	-0.09	0.08
Stroke duration variability (s)	0.44	[0.18-0.65]	0.00	0.05	-0.02 (0.07)	18.17	-0.16	0.12
Total number of strokes	0.46	[0.20-0.66]	0.00	1.31	-0.19 (1.84)	1.03	-3.79	3.42
Total impulse (Ns)	0.88	[0.79-0.93]	0.00	85.75	74.76 (121.84)	5.41	-164.04	313.56
Endurance index (%)	0.74	[0.58-0.85]	0.00	0.06	0.03 (0.08)	3.57	-0.12	0.19
Time to first peak (s)	0.41	[0.14-0.62]	0.01	0.12	0.03 (0.17)	5.31	-0.29	0.36
Rate of force development (N/s)	0.25	[0.00-0.50]	0.08	1335.5	532.8 (1895.7)	32.84	-3182.9	4248.5

ICC = intra-class correlation, CI = confidence interval, Sig=significance level, SEM = standard error of measurement, LOA = limit of agreement, N = Newtons, s = seconds

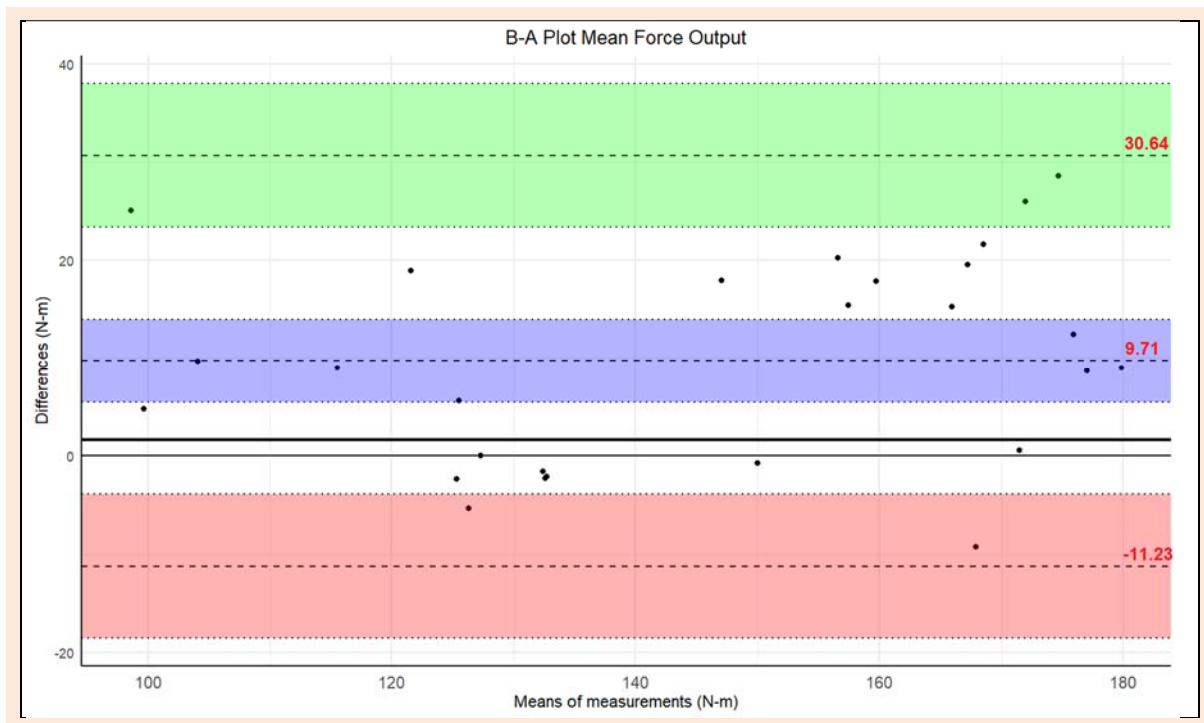


Figure 2. Bland and Altman diagram for mean peak force (values indicate the bias, upper limit of agreement, and lower limit of agreement)

Table 2. Group comparisons between sexes.

Variable	Male mean (SD)	Female mean (SD)	Significance (effect size)	Male relative mean (SD)	Female relative mean (SD)	Significance (effect size)
Overall mean force (N)	170.2 (12.2)	122.1 (10.0)	<0.01 (4.23)	1.94 (0.28)	1.66 (0.28)	<0.01 (1.00)
Max Force (N)	203.9 (6.9)	143.8 (5.8)	<0.01 (9.36)	2.33 (0.32)	1.96 (0.29)	<0.01 (1.22)
Mean peak force (N)	183.3 (14.2)	130.5 (12.1)	<0.01 (3.96)	2.09 (0.30)	1.77 (0.30)	<0.01 (1.07)
Inter-stroke range (N)	25.6 (11.3)	15.5 (6.6)	<0.01 (1.05)	0.29 (0.13)	0.21 (0.08)	0.01 (0.76)
Total impulse (Ns)	1578 (182)	1171 (88)	<0.01 (2.71)	18.05 (3.21)	15.93 (2.61)	0.04 (0.71)
Endurance index (%)	84.8 (11.1)	85.3 (10.7)	0.88 (0.04)	n/a	n/a	n/a

N = Newtons, s = seconds, SD = standard deviation

Discussion

The main objective of this study was to evaluate reliability and validity of a standardized testing approach to strength in water polo players. These objectives were achieved with the development of novel parameters specific to the performance of a ten-second resisted eggbeater task. However, factors describing the onset of the task were not sufficiently reliable to help evaluate player performances. The outcome variables of mean force, maximum force, mean peak force, inter-stroke range and total impulse were useful to distinguish between male and female players.

In the current study, the male players ($170 \pm 12\text{N}$ mean force and $204 \pm 7\text{N}$ maximum force) demonstrated greater force compared with the female players ($122 \pm 10\text{N}$ mean force and $144 \pm 6\text{N}$ maximum force) (Figure 3). These female values are greater than those of 60 - 112N reported by Yanagi et al (1995) in a group of 15 female Japanese water polo players (Yanagi, 1995). The testing regimen evaluated vertical eggbeater performance in their study, which may explain the lower values. However, the male values resemble those reported previously by Dopsaj et al (2010) ($140 \pm 21\text{N}$ mean force and $191 \pm 36\text{N}$ maximum force) in 14 senior level male water polo players attached to a PVC rope (Dopsaj, 2010) as well as Stirn et al (2014) ($128 \pm 26\text{N}$ mean force and $189 \pm 36\text{N}$ maximum force) in 28 youth male water polo players aged 14-16 years old with elastic tether) (Stirn et al., 2014). In contrast, Abad et al (2022) have recently reported higher values for mean ($389 \pm 70\text{N}$) and maximum forces ($672 \pm 135\text{N}$) in a group of 32 professional Brazilian water polo players aged 22.2 ± 4.4 years in a non-elastic rope tethered 10-second test (Abad et al., 2022). The range of values found here may be explained partly by age group differences, but it suggests that there may also still be persistent differences in testing methods (Taylor, 2001).

This current study differed from previous methods by imposing a fixed resistance on the water polo players executing the task as opposed to being tethered with elastic tubing or non-elastic ropes (Dopsaj, 2010; Dopsaj et al., 2003; Stirn et al., 2014). Reliability of these previous methods have shown good to very good results across a variety

of swimming protocols and populations (Kjendlie and Thorsvald, 2005; Nagle Zera et al., 2021). Most often, authors found that mean force estimates were more reliable than maximum force means, with lower coefficients of variation (5.4 - 8.9% vs 11.6 - 14.7%, $n = 13$) (Taylor, 2001), greater internal consistency (Cronbach's $\alpha = 0.869 - 0.995$) (Dopsaj et al., 2003; Kjendlie and Thorsvald, 2005), and greater intra-class correlations (ICC = 0.975 vs 0.861, $p < 0.001$, $n = 19$) (Nagle Zera et al., 2021). In the current study, the reliability coefficients for overall mean force and maximum force productions showed comparable results (ICC = 0.91 [0.85 - 0.95] and 0.93 [0.88 - 0.96] respectively), suggesting that within-subject variations are small across repeated sessions for these performance variables (Weir, 2005). However, time-based variables such as mean stroke duration, the standard deviation of mean stroke duration (Joaquim Baratto de Azevedo et al., 2021), the total number of strokes and the time to first peak showed poor correlation across the two testing sessions (ICC = 0.38 - 0.46).

The results suggest that the early phase of the test is prone to measurement error because of the hysteresis of the system (Psycharakis et al., 2011). This was observed on video analysis, where the elastic components of the system yield a certain amount of deformation at the onset of the test. Consequently, time to first peak was not dependable between testing sessions (ICC = 0.41, 95%CI = 0.14 - 0.62) in this study, which in turn made the estimation of a rate of force development entirely unreliable (ICC = 0.25, relative bias = 32.8%) (Koo and Li, 2016). The difficulty to obtain reliable values for these brief time windows are also reported by previous authors (Dopsaj, 2010).

Additionally, the current study also developed variables that characterize the ability of water polo players to maintain a high level of force output throughout the entire 10 seconds of the test. The first is an endurance index similar to what has been explored previously (Figure 4) (Morouço et al., 2012), whereas the second is a measure of total impulse (force x time). Of these, the most reproducible characteristic was impulse (ICC = 0.88, 95%CI = 0.79 - 0.93), which can readily compare overall performance across participants using different techniques

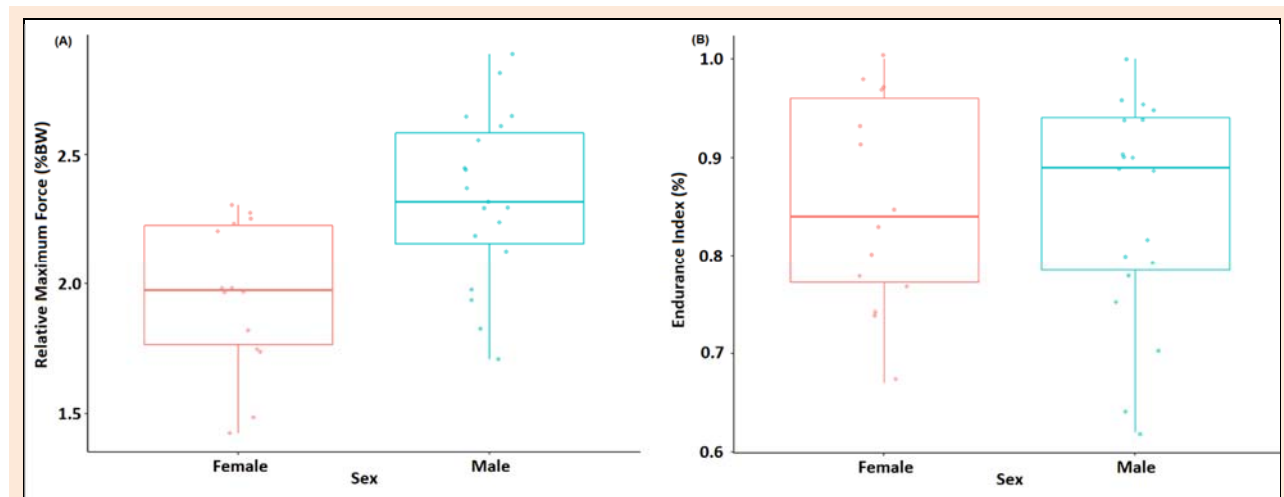


Figure 3. Comparison of (A) maximum strength relative to body weight between males and females, and (B) endurance index between males and females.

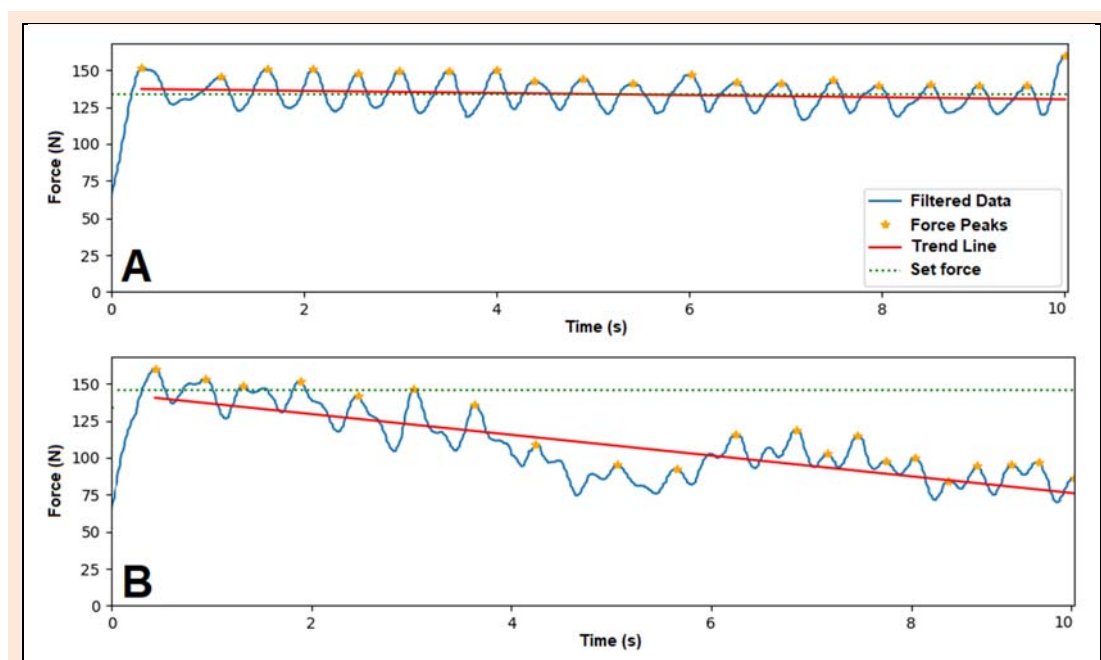


Figure 4. Force recordings during the pushing task with example for high (A) and low (B) endurance indices.

to achieve the task goal. This information is most useful to coaches, as it can identify players whose technique relies on sudden bursts of force as opposed to a sustained force. Technical recommendations can directly result from these observations and improve performance.

The next objective of this study was to evaluate the validity of the test by comparing male and female players. Indeed, the male group showed significantly greater values for mean force, maximum force, mean peak force, inter-stroke range and total impulse ($p < 0.01$). Furthermore, these results showed a large to very large effect size across all five variables as well ($ES\ g = 1.05 - 9.36$). Conversely, the endurance indices were not significantly different between the two groups ($p = 0.88$). This is consistent with expected strength differences between sexes (Bartolomei et al., 2021). Unfortunately, no other comparison of similar water-based strength tests is available between groups of water polo players. Instead, previous authors have explored construct validity for similar tests by measuring the correlation with swimming performances (Currell and Jeukendrup, 2008; Morouço et al., 2014). Overall, mean and maximum force production showed moderate to strong correlation with swimming performance across shorter distances ($r = 0.57 - 0.82$, $p < 0.01$) (Joaquim Baratto de Azevedo et al., 2021; Nagle Zera et al., 2021). Altogether, these findings consistently demonstrate that pushing tasks are a valid method to assess strength performance in water polo players, with a reasonable relationship to ecological game situations (Melchiorri et al., 2020).

Limitations

The data for this study was collected as part of regular training sessions. This was done to use the same weight stack resistances that coaches use to train the participants. However, the consequence is that not all players began the testing process after having done the same warm-up. Given that this warm-up was self-paced, some began the test after choosing to perform more energetic routines. This may

have biased the results, with some players having greater energetic reserves at the outset of the test.

The choice to position the hands forward and out of the water was made by the research team to standardize the test and compare results across athletes. However, match duels are also highly influenced by using the upper body, which was difficult to regulate in the current study. Additionally, the initial discussions with the coaching experts revealed that many alternative testing positions would be accessible with this technology, but the research team chose a protocol that could show adequate reliability as opposed to perfect ecological validity.

The target of testing all players one week apart was maintained for most participants. This period was preferred to a shorter comparison time because the weekly training sessions differ between days but repeat on a weekly cycle. However, six players were unavailable to repeat the testing because of academic commitments, and others were not present at the training session seven days later. Therefore, some testing was done on the closest available days to the retest date. This may have led to greater bias between sessions for those participants, but the sample is too small to ascertain. Future research may control for potential differences in training load between testing sessions using player-reported measures of rating of perceived exertion (RPE) (Croteau et al., 2023; Lupo et al., 2014).

The sample is comprised of players from both international and national levels, however the sex distribution between the groups is not balanced. The international players included significantly more females ($p < 0.01$) and were significantly older than the national level players ($p < 0.01$). Nevertheless, the use of standard resistance for males that was different than females was used to control for sex and decrease its effect on the outcomes. The participants volunteered for this study, and at the time of data collection, a greater portion of senior national team female players were present. A more balanced distribution of player levels would be optimal to evaluate the effect of

player experience on these strength variables. Moreover, a larger sample could explore whether player position show consistent differences in performance.

Finally, future research could focus on evaluating the kinematic parameters collected with video analysis to investigate the relationship with kinetic variables described in this study. This would allow comparison of these findings with studies based primarily on analysis of eggbeater movement characteristics (Platanou, 2006; Sanders, 1999b).

Conclusion

The study successfully evaluated the reliability and validity of a water polo strength test against a fixed resistance. This study is original compared to previous work thanks to the addition of more kinetic parameters derived from the analysis of the force outputs. This fixed-resistance setup provides a more ecological condition for the players to demonstrate how they would succeed in resisting the same opponent. The instruments required are portable and cost-effective. The strength variables developed to assess the outcomes showed high to very high reliability, whereas time-based variables only showed small to moderate reliability. Validity was explored through known groups, where male players demonstrated higher strength, but not superior endurance to the female players. The practical applications of this testing apparatus are readily available for coaches and sport scientists working with water polo teams. Similar load cells can be attached to a weight stack such as done here, or to a resistance provided by water jugs on a pulley. The data are simple to obtain, and the parameters that were explored in this study can inform on both individual progress or to make comparisons across players. The specific analysis of these biomechanical markers can also serve to evaluate the changes made by modifying technical execution of the eggbeater motion: changing the speed of execution, the range of motion at the hips, knees or ankles, as well as using an alternating or simultaneous kick to push the resistance.

Acknowledgements

The authors want to thank the players who agreed to take part in this study and the Innovations 4 Gold (I4G) initiative from Own the Podium for supplying funding for this project's realization. Thank you to the coaches from the Canadian Olympic teams for providing key insights about the worthwhile variables to explore. There is no conflict of interest. The present study complies with the current laws of the country in which it was performed. The datasets generated and analyzed during the current study are not publicly available but are available from the corresponding author, who was an organizer of the study.

References

- Abad, C., Lara, J., Sousa, B., Oliveira, A., Planas, R., Rossin, R. and Avallone, A. (2022) Characterization and relationships among anthropometrical and physical fitness performances in professional Brazilian water polo players. *Serbian Journal of Sport Sciences* **13**, 19-28.
- Abernethy, P., Wilson, G. and Logan, P. (1995) Strength and power assessment. Issues, controversies and challenges. *Sports Medicine* **19**, 401-417. <https://doi.org/10.2165/00007256-199519060-00004>
- Bartolomei, S., Grillone, G., Di Michele, R. and Cortesi, M. (2021) A Comparison between Male and Female Athletes in Relative Strength and Power Performances. *Journal of Functional Morphology and Kinesiology* **6**. <https://doi.org/10.3390/jfkm6010017>
- Botonis, P.G., Toubekis, A.G. and Platanou, T.I. (2019) Physiological and Tactical On-court Demands of Water Polo. *Journal of Strength & Conditioning Research* **33**, 3188-3199. <https://doi.org/10.1519/JSC.0000000000002680>
- Bratuša, Z. and Dopsaj, M. (2016) The effect of various leg kick techniques on the vertical jump among water polo players. *Facta Universitatis Series Physical Education and Sport*. 419-430.
- Cohen, J. (2013) *Statistical power analysis for the behavioral sciences*. 2 edition. London, England: Routledge. <https://doi.org/10.4324/9780203771587>
- Croteau, F., Gaudet, S., Briand, J. and Clément, J. (2023) Case study of IMU loads and self-reported fatigue monitoring of water polo goalkeepers preparing for the Olympic games. *Frontiers in Sports and Active Living* **5**. <https://doi.org/10.3389/fspor.2023.1198003>
- Croteau, F., Robbins, S.M. and Pearsall, D. (2021) Hand-Held Shoulder Strength Measures Correlate With Isokinetic Dynamometry in Elite Water Polo Players. *Journal of Sport Rehabilitation* 1-4. <https://doi.org/10.1123/jsr.2020-0277>
- Currell, K. and Jeukendrup, A. (2008) Validity, reliability and sensitivity of measures of sporting performance. *Sports Medicine* **38**, 297-316. <https://doi.org/10.2165/00007256-200838040-00003>
- Dopsaj, M. (2010) Pulling force characteristics of 10s maximal tethered eggbeater kick in elite water polo players: a pilot study. *Biomechanics and Medicine in Swimming XI*.
- Dopsaj, M., Matković, I., Thanopoulos, V. and Okičić, T. (2003) Reliability and validity of basic kinematics and mechanical characteristics of pulling force in swimmers measured by the method of tethered swimming with maximum intensity of 60 seconds. *Physical Education and Sport* **10**, 11-22.
- Evans, J.D. (1996) *Straightforward Statistics for the Behavioral Sciences*. Pacific Grove: Brooks/Cole Publishing Company.
- Gobbi, M., D'Ercole, C., D'Ercole, A. and Gobbi, F. (2013) The components of the jumps in expert and intermediate water polo players. *Journal of Strength & Conditioning Research* **27**, 2685-2689. <https://doi.org/10.1519/JSC.0b013e318240ebf1>
- Gullikson, J., Mayberry, J.K., Gale, L.R. and Killick, L. (2020) Not throwing away my shot: an analysis of shot features in men's collegiate water polo. *Journal of Performance Analysis in Sport* **20**, 240-253. <https://doi.org/10.1080/24748668.2020.1741915>
- Hays, R.D. and Reeve, B.B. (2008) Measurement and Modeling of Health-Related Quality of Life. In: *International Encyclopedia of Public Health*. Ed: Heggenhougen, H.K. Oxford: Academic Press. 241-252. <https://doi.org/10.1016/B978-012373960-5.00336-1>
- International Society for Advancement of, K. (2001) *International Standards for Anthropometric Assessment*. Potchefstroom, South Africa: International Society for the Advancement of Kinanthropometry.
- Joaquim Baratto de Azevedo, O., Knierim Correia, C., Soares Pereira, G., Prado, L.S., Roesler, H., Pereira, S.M. and Ruschel, C. (2021) Effect of three different set-up conditions on the propulsive force measures, reliability, and ecological validity during front crawl tethered-swimming. *Journal of Performance Analysis in Sport* **21**, 1081-1100. <https://doi.org/10.1080/24748668.2021.1974182>
- Keiner, M., Rahse, H., Wirth, K., Hartmann, H., Fries, K. and Haff, G.G. (2020) Influence of Maximal Strength on In-Water and Dry-Land Performance in Young Water Polo Players. *Journal of Strength & Conditioning Research* **34**, 1999-2005. <https://doi.org/10.1519/JSC.0000000000002610>
- Kjendlie, P.-L. and Thorsvald, K. (2005) A tethered swimming power test is highly reliable. *Portuguese Journal of Sport Sciences* **6**, 231-233.
- Koo, T.K. and Li, M.Y. (2016) A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine* **15**, 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Lupo, C., Capranica, L., Cugliari, G., Gomez, M.A. and Tessitore, A. (2016) Tactical swimming activity and heart rate aspects of youth water polo game. *Journal of Sports Medicine and Physical Fitness* **56**, 997-1006.
- Lupo, C., Capranica, L. and Tessitore, A. (2014) The validity of the session-RPE method for quantifying training load in water polo. *International Journal of Sports Physiology and Performance* **9**, 656-660. <https://doi.org/10.1123/ijsp.2013-0297>

- McCluskey, L., Lynskey, S., Leung, C.K., Woodhouse, D., Briffa, K. and Hopper, D. (2010) Throwing velocity and jump height in female water polo players: performance predictors. *Journal of Science & Medicine in Sport* **13**, 236-240. <https://doi.org/10.1016/j.jsams.2009.02.008>
- McKay, A.K.A., Stellingwerff, T., Smith, E.S., Martin, D.T., Mujika, I., Goosey-Tolfrey, V.L., Sheppard, J. and Burke, L.M. (2022) Defining Training and Performance Caliber: A Participant Classification Framework. *International Journal of Sports Physiology and Performance* **17**, 317-331. <https://doi.org/10.1123/ijsp.2021-0451>
- Melchiorri, G., Viero, V., Tancredi, V., Bianco, R.D. and Bonifazi, M. (2020) Actual Playing Time of Water Polo Players in Relation to the Field Position. *Journal of Human Kinetics* **73**, 241-249. <https://doi.org/10.2478/hukin-2019-0148>
- Melchiorri, G., Viero, V., Triossi, T., Tancredi, V., Galvani, C. and Bonifazi, M. (2015) Testing and Training of the Eggbeater Kick Movement in Water Polo: Applicability of a New Method. *Journal of Strength & Conditioning Research* **29**, 2758-2764. <https://doi.org/10.1519/JSC.0000000000000946>
- Morouço, P.G., Marinho, D.A., Keskinen, K.L., Badillo, J.J. and Marques, M.C. (2014) Tethered swimming can be used to evaluate force contribution for short-distance swimming performance. *Journal of Strength & Conditioning Research* **28**, 3093-3099. <https://doi.org/10.1519/JSC.0000000000000509>
- Morouço, P.G., Vilas-Boas, J.P. and Fernandes, R.J. (2012) Evaluation of adolescent swimmers through a 30-s tethered test. *Pediatric Exercise Science* **24**, 312-321. <https://doi.org/10.1123/pes.24.2.312>
- Muniz-Pardos, B., Gomez-Bruton, A., Matute-Llorente, A., Gonzalez-Aguero, A., Gomez-Cabello, A., Gonzalo-Skok, O., Casajus, J.A. and Vicente-Rodriguez, G. (2019) Swim-Specific Resistance Training: A Systematic Review. *Journal of Strength & Conditioning Research* **33**, 2875-2881. <https://doi.org/10.1519/JSC.0000000000003256>
- Nagle Zera, J., Nagle, E.F., Nagai, T., Lovalekar, M., Abt, J.P. and Lephart, S.M. (2021) Tethered Swimming Test: Reliability and the Association With Swimming Performance and Land-Based Anaerobic Performance. *Journal of Strength & Conditioning Research* **35**, 212-220. <https://doi.org/10.1519/JSC.0000000000002501>
- Perazzetti, A., Dopsaj, M., Mandorino, M. and Tessitore, A. (2023a) Assessment of the Offensive Play in Elite Water Polo Using the Team Sport Assessment Procedure (TSAP) over an Entire Competitive Season. *Journal of Functional Morphology and Kinesiology* **8**. <https://doi.org/10.3390/jfmk8030130>
- Perazzetti, A., Dopsaj, M. and Tessitore, A. (2023b) Technical and tactical evaluation of ball possession in international youth water polo matches using the Team Sport Assessment Procedure (TSAP) instrument. *Montenegrin Journal of Sports Science and Medicine* **12**, 3-9. <https://doi.org/10.26773/mjssm.230901>
- Platanou, T. (2004) Time-motion analysis of international level water polo players. *Journal of Human Movement Studies* **46**, 319-331.
- Platanou, T. (2005) On-water and dryland vertical jump in water polo players. *Journal of Sports Medicine & Physical Fitness* **45**, 26-31.
- Platanou, T. (2006) Simple 'in-water' vertical jump testing in water polo. *Kinesiology* **38**, 57-62.
- Platanou, T. and Geladas, N. (2006) The influence of game duration and playing position on intensity of exercise during match-play in elite water polo players. *Journal of Sports Sciences* **24**, 1173-1181. <https://doi.org/10.1080/02640410500457794>
- Platanou, T. and Varamenti, E. (2011) Relationships between anthropometric and physiological characteristics with throwing velocity and on water jump of female water polo players. *Journal of Sports Medicine & Physical Fitness* **51**, 185-193.
- Plichta, S.B. and Kelvin, E.A. (2012) *Munro's Statistical Methods for Health Care Research*. Wolters Kluwer Health/Lippincott Williams & Wilkins.
- Psycharakis, S.G., Paradisis, G.P. and Zachariogiannis, E. (2011) Assessment of accuracy, reliability and force measurement errors for a tethered swimming apparatus. *International Journal of Performance Analysis in Sport* **11**, 410-416. <https://doi.org/10.1080/24748668.2011.11868560>
- Sanders, R.H. (1999a) Analysis of the Eggbeater Kick Used to Maintain Height in Water Polo. *Journal of Applied Biomechanics* **15**, 284-291. <https://doi.org/10.1123/jab.15.3.284>
- Sanders, R.H. (1999b) A Model of Kinematic Variables Determining Height Achieved in Water Polo Boosts. *Journal of Applied Biomechanics* **15**, 270-283. <https://doi.org/10.1123/jab.15.3.270>
- Smith, H.K. (1998) Applied physiology of water polo. *Sports Medicine* **26**, 317-334. <https://doi.org/10.2165/00007256-199826050-00003>
- Stirn, I., Strmecki, J. and Strojnik, V. (2014) The examination of different tests for the evaluation of the efficiency of the eggbeater kicks. *Journal of Human Kinetics* **41**, 215-226. <https://doi.org/10.2478/hukin-2014-0049>
- Taylor, S.R.L.A.S.G.M.D.P.M. (2001) Reliability of force production in tethered freestyle swimming among competitive age-group swimmers. *Communications to the Annual Conference of the British Association of Sport and Exercise Sciences (BASES)* 3-14.
- Team, R.C. (2022). R: A language and environment for statistical computing. Vienna, Austria.
- Uljevic, O., Spasic, M. and Sekulic, D. (2013) Sport-specific motor fitness tests in water polo: reliability, validity and playing position differences. *Journal of Sports Science and Medicine* **12**, 646-654. 2013. <https://pubmed.ncbi.nlm.nih.gov/24421723>
- Weir, J.P. (2005) Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength & Conditioning Research* **19**, 231-240. <https://doi.org/10.1519/00124278-200502000-00038>
- Yanagi, H.A.K.S.T.H.A. (1995) Vertical forces exerted during the eggbeater kick in water polo. *XVth Congress of the International Society of Biomechanics* 1016-1017.

Key points

- Using a standard resistance allows to readily compare force output across different players that must do the same task in a water polo match.
- This study presents a method that can readily be implemented using training equipment with minimal costs.
- A list of six variables showed high to very high reliability in describing performance: overall mean force, mean peak force, maximal force, inter-stroke range, total impulse, and a new proposed endurance index.

AUTHOR BIOGRAPHY



Félix CROTEAU

Employment

Physiotherapist at Institut National du Sport du Québec

Degree

PT, PhD

Research interests

Injury prevention, training load, sport biomechanics

E-mail: fcroteau@insquebec.org



Élouan LANOË

Employment

Student at Grenoble INP Génie Industriel

Degree

Undergraduate Eng.

Research interests

Sport sciences

E-mail: elouan.lanoe@grenoble-inp.org

**Julien CLÉMENT****Employment**

Associate Professor at École de technologie supérieure

Degree

Eng., MScA, PhD

Research interests

Sport biomechanics, Training load, Performance analysis

E-mail: julien.clement@etsmtl.ca

✉ Félix Croteau

Institut national du sport du Québec, 4141 avenue Pierre de Coubertin, H4V 1N7, Montréal, Québec, Canada