Reproducibility of an aerobic endurance test for nonexpert swimmers

Adalberto Veronese da Costa¹,²
Manoel da Cunha Costa³
Daniel Medeiros Carlos¹
Luis Marcos de Medeiros Guerra¹,²
Antônio José Silva²
Tiago Manoel Cabral dos Santos Barbosa²,⁴

¹Department of Physical Education, Bioscience Laboratory of Human Kinetics, Rio Grande do Norte State University, Mossoró, Brazil; ²Sport Sciences Trás-os-Montes e Alto Douro University, Research Center in Sport, Health and Human Development, Vila Real, Portugal; ³Superior School of Physical Education, Human Performance Laboratory, Pernambuco State University, Recife, Brazil; ⁴National Institute of Education, Nanyang Technological University, Singapore

Background: This study aimed to verify the reproduction of an aerobic test to determine nonexpert swimmers’ resistance.

Methods: The sample consisted of 24 male swimmers (age: 22.79 ± 3.90 years; weight: 74.72 ± 11.44 kg; height: 172.58 ± 4.99 cm; and fat percentage: 15.19% ± 3.21%), who swim for 1 hour three times a week. A new instrument was used in this study (a Progressive Swim Test): the swimmer wore an underwater MP3 player and increased their swimming speed on hearing a beep after every 25 meters. Each swimmer’s heart rate was recorded before the test (BHR) and again after the test (AHR). The rate of perceived exertion (RPE) and the number of laps performed (NLP) were also recorded. The sample size was estimated using G*Power software (v 3.0.10; Franz Faul, Kiel University, Kiel, Germany). The descriptive values were expressed as mean and standard deviation. After confirming the normality of the data using both the Shapiro–Wilk and Levene tests, a paired t-test was performed to compare the data. The Pearson’s linear correlation (r) and intraclass coefficient correlation (ICC) tests were used to determine relative reproducibility. The standard error of measurement (SEM) and the coefficient of variation (CV) were used to determine absolute reproducibility. The limits of agreement and the bias of the absolute and relative values between days were determined by Bland–Altman plots. All values had a significance level of P < 0.05.

Results: There were significant differences in AHR (P = 0.03) and NLP (P = 0.01) between the 2 days of testing. The obtained values were r > 0.50 and ICC > 0.66. The SEM had a variation of ±2% and the CV was <10%. Most cases were within the upper and lower limits of Bland–Altman plots, suggesting correlation of the results. The applicability of NLP showed greater robustness (r and ICC > 0.90; SEM < 1%; CV < 3%), indicating that the other variables can be used to predict incremental changes in the physiological condition of swimmers.

Conclusion: The Progressive Swim Test for nonexpert swimmers produces comparable results for noncompetitive swimmers with a favorable degree of reproducibility, thus presenting possible applications for researching the physiological performance of nonexpert swimmers.

Keywords: swimming, physical evaluation, resistance training, health

Introduction

The assessment of cardiorespiratory fitness has been very useful for diagnosing the health and fitness in a variety of populations (eg, types of physical activity or age), assuming that a favorable heart condition is related to performance improvement and longevity.¹,² The evaluation of sports performance by researchers generally uses methodology designed for elite athletes. Our study was limited by the absence of new methodologies designed for the general public practicing the sport for a better lifestyle.³ This study outlines a new aerobic endurance test for nonexpert adult
swimmers who can perform swimming techniques, especially freestyle (front-crawl), and who are not affiliated with the national swimming organization nor are former participants in regional, national, and international championships.3

Specific sports utilize laboratory tests under controlled temperature, relative air humidity, intensity of treadmill exercise, and cycle ergometers to determine the maximal and submaximal level of relative oxygen consumption (VO2). This controlled environment is a test “field” that approaches the conditions of a real situation in a specific sport.4,5 In swimming, it is possible to identify specific tests for aerobic endurance that are valid for athletes.6-8

Recent research relating to competitive swimming has produced data on the performance, energetics, and biomechanics of athletes that can be used to answer questions about the effects of exercise intensity.6,10,11 However, these variables need to be assessed in the fitness levels of nonexpert swimmers3,12 as the biomechanics of the physical patterns become more efficient through regular exercise and practice, less energy is spent on the activity and the swimmers become healthier.4,13,14

Aerobic swimming tests were traditionally designed to control the intensity of exercise, using either the time or distance to be executed by the swimmer.15,16 Thus, there are testing intervals (those in which the intensity of the exercise is performed at a moderate or strong pace, with rest breaks between one repetition and another as in the two Mader test speeds) and continuous tests (which measure the greatest distance swum in a given time interval such as 30 minutes [T30]).17 However, it is believed that the most suitable type of test when working with nonexpert swimmers is a progressive test, or one in which every lap swum with the front-crawl technique increases in swimming intensity in relation to the execution time.18 Studies aimed at demonstrating this progressive method are best measured within a competitive environment,18-20 which facilitates the execution of a water test by nonexpert swimmers without concern for mistakes regarding the pace of a specific swimmer – a factor that might limit continuous tests and intervals for all these swimmers. Thus, this study questioned the reliability of a new progressive test, on the basis that evaluating its effectiveness will enable further research to be conducted to validate the test by direct methods.21

In view of the fact that some water tests are outside commercial reality in many swimming schools, this study resorted to new alternative methods that evaluate swimmers through indirect tests. An adaptation of the water progressive running test as proposed by Léger and Boucher22 allowed this study to conduct a low-cost, easily applicable aerobic test to evaluate the exertion of nonexpert swimmers. It was possible to assess the ability of the swimmers through the largest number of laps swum.

Therefore, this study aimed to verify the reproducibility of a test for aerobic endurance by nonexpert swimmers. The hypothesis is that the results are highly replicable.

Materials and methods

Subjects

This comparative descriptive study comprised 24 nonexpert male swimmers in the town of Mossoró, Rio Grande do Norte, northeast of Brazil, aged from 18 to 30 years (age: 22.79 ± 3.90 years, weight: 74.72 ± 11.44 kg, height: 172.58 ± 4.99 cm, body fat percentage: 15.19% ± 3.21%). Swimmers with good technique who used the front-crawl swimming style attended lessons at least three times a week and swam for at least 1 hour per session over a distance of approximately 800 meters. Those included in the criteria did not take any food supplements or conduct any type of physical activity 24 hours before the tests. The exclusion criteria for the study included swimming athletes affiliated to the Brazilian Confederation of Aquatic Sports, athletes who had been affiliated to this federation 3 years before the survey, or nonexpert swimmers or participants who indicated any type of debilitating illness (eg, flu, fever, or any type of injury). The health care professionals who evaluated the tests were two swimming teachers with an academic background in physical education who had a minimum of 3 years of professional experience, together with a less-experienced assistant teacher.

Not unlike many swimming schools around the world, the schools that agreed to participate in the survey had a higher number of adults in learning phase for this research program, unlike those who met the inclusion criteria of this study. Thus, this study did not prevent quantitative reproduction of a new test for this category of swimmers, which can be observed in other studies.15,23-25

New test for aerobic endurance for nonexpert swimmers as a test of reproducibility

The Progressive Swim Test corresponds to a series of 400 meters in a 25-meter pool, based on the world record for men (03’32”57; www.fina.org/H2O/) at this distance in a short-course pool. The world-record time was registered by the International Federation of Amateur Swimming in Berlin, Germany in 2009. A beeping sound indicates the swimming rhythm, occurring at a decrease in partial time of 1 second for each performed lap, with a beep given at the end of the first
lap at the time of 28″30 (ie, a rate of 0.88 m/s) and for the 16th at 13″30; (ie, a rate of 1.88 m/s) (Table 1).

The start of each progressively more intense lap was indicated by MP3-format sound signals, using the subaquatic device SwiMP3 v2 (Finis Inc, Livermore, CA) attached to the silicon strip of the swimmer’s goggles. The direct transfer of sound vibrations from the swimmer’s jaw bone to their ear provided exceptional sound clarity. The sounds were also played through a microsystem outside the pool, synchronized with the subaquatic MP3, as an aid to the evaluators monitoring the tests. The beeps increased in frequency for each 25 meters swum, causing an increase in speed during a repetition of 400 meters using the front-crawl swimming technique.

One week before testing, the evaluators familiarized the swimmers with the procedures that would be implemented by providing an illustrative video on how the test would be conducted. After week 1, the tests were initiated. Each swimmer, before starting the test, performed between 50 and 100 meters with the equipment, as a warm-up exercise.

The swimmer remained in the pool until the synchronization of the MP3 and the microsystem was made by the main evaluator. Following an announcement (“Attention swimmer, prepare for the test”), a short beep sounded, accompanied by a visual signal from the hands of the evaluator to mark the beginning of the test. To facilitate the identification of the correct rhythm for every lap swum, a longer beep of a different type was heard by the swimmer to indicate that he should be near the middle of the pool. After completing each lap, another short beep was given to indicate the start of the next lap.

The swimmer was instructed to try to keep to the rhythm sounded by the MP3 v2 so as to be always within the beginning/ending 5-meter zone when the short beep was heard.

As identifiers for both the swimmers and the evaluators, cones were placed along the edge of the testing lane, and 4 kg rings inside the pool, to mark the initial and final 5 meters and 12.5 meters (half-distance) (Figure 1). During the test, while the main evaluator counted the number of laps (Table 1), the evaluator’s assistant monitored the swimmer, giving verbal and visual signals of the progressive test pace.

The test ended when the swimmer didn’t reach the 5-meter zone preceding the edge of the pool two laps in a row. At this point, the main evaluator immediately measured the heart rate of the swimmer, while the assistant evaluator presented a rate of perceived exertion (RPE). Then, the two evaluators registered the number of laps performed by the swimmer. If any swimmer, for any reason, were to stop during the test, their test would be aborted and not counted in the results.

### Variables for analysis

For result analysis, it was decided to measure the resting heart rate 30 minutes before the start of the test (before heart rate [BHR]) and after the end of the test (after heart rate [AHR]) with a heart rate monitor (Polar Model FT1; Polar Electro Oy, Kempele, Finland). These measurements were taken inside the pool at the swimmer’s chest level. At the end of the test, a Borg’s RPE score (scale from 1 to 10) was presented to the swimmer, but the number of laps performed (NLP) was registered by the evaluators. The swimmers were evaluated on different days, in the same order, on their same respective day each week and at their same respective time, with a main and assistant evaluator present at each test. Before each test the water temperature was measured in order to identify any difference in swimmers’ environmental conditions (Bestway® Floating Pool Thermometer, accurate to 1°C; Bestway Inflatables North America Inc, Mississauga, ON, Canada); the temperature in the two days of testing was 28°C.

### Ethical aspects

All the procedures adhered to the guidelines of the Declaration of Helsinki (www.wma.net/e/policy/b3.htm). This study was approved by the Ethics Committee of the Rio Grande do Norte University, Brazil – CEP/UERN, protocol 070/2011. All participants, including swimmers, teachers, evaluators, and assistants, signed a Free and Clarified Consent.

---

**Table 1** Laps performed during the Progressive Swim Test for nonexpert swimmers

<table>
<thead>
<tr>
<th>Lap</th>
<th>Time</th>
<th>m/s</th>
<th>k/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28″30</td>
<td>0.88</td>
<td>3.18</td>
</tr>
<tr>
<td>2</td>
<td>27″30</td>
<td>0.92</td>
<td>3.30</td>
</tr>
<tr>
<td>3</td>
<td>26″30</td>
<td>0.50</td>
<td>3.42</td>
</tr>
<tr>
<td>4</td>
<td>25″30</td>
<td>0.99</td>
<td>3.56</td>
</tr>
<tr>
<td>5</td>
<td>24″30</td>
<td>1.03</td>
<td>3.70</td>
</tr>
<tr>
<td>6</td>
<td>23″30</td>
<td>1.07</td>
<td>3.86</td>
</tr>
<tr>
<td>7</td>
<td>22″30</td>
<td>1.12</td>
<td>4.04</td>
</tr>
<tr>
<td>8</td>
<td>21″30</td>
<td>1.17</td>
<td>4.23</td>
</tr>
<tr>
<td>9</td>
<td>20″30</td>
<td>1.23</td>
<td>4.43</td>
</tr>
<tr>
<td>10</td>
<td>19″30</td>
<td>1.30</td>
<td>4.66</td>
</tr>
<tr>
<td>11</td>
<td>18″30</td>
<td>1.37</td>
<td>4.92</td>
</tr>
<tr>
<td>12</td>
<td>17″30</td>
<td>1.45</td>
<td>5.20</td>
</tr>
<tr>
<td>13</td>
<td>16″30</td>
<td>1.53</td>
<td>5.52</td>
</tr>
<tr>
<td>14</td>
<td>15″30</td>
<td>1.63</td>
<td>5.88</td>
</tr>
<tr>
<td>15</td>
<td>14″30</td>
<td>1.75</td>
<td>6.29</td>
</tr>
<tr>
<td>16</td>
<td>13″30</td>
<td>1.88</td>
<td>6.77</td>
</tr>
</tbody>
</table>

**Note:** This table was used as a monitoring aid by evaluators during the execution of the Progressive Swim Test.

**Abbreviations:** m/s, meters per second; k/h, kilometers per hour.

---
Statistical analysis

The sample size was estimated by the software G*Power 3.0.10 (Franz Faul, University of Kiel, Kiel, Germany), considering the effect size 0.75, the minimum power 0.80, and \( \alpha = 0.05 \), resulting in 13 swimmers.\(^{29}\) However, it was possible to evaluate 24 swimmers. Atkinson and Nevill\(^{31}\) believed that statistical methods are necessary in order to quantify the results substantially. From this perspective, the descriptive values were expressed as mean and standard deviation. After testing the normality of the data through the Shapiro–Wilk test and Levene tests, a paired \( t \)-test was performed for comparison between the days. To analyze the reliability, the \( r \)-test was used (Pearson’s linear correlation), classified according to da Silva et al\(^{30}\) as very low (<0.20), low (0.20–0.39), moderate (0.40–0.59), high (0.60–0.79), and very high (0.80–1). In analyzing the consistency among the variables, intraclass correlation coefficient (ICC) was used, as categorized by Santos et al\(^{31}\) and Sánchez and Echeverry:\(^{32}\) not acceptable (<0.70), acceptable (0.70–0.79), good agreement (0.80–0.89), and excellent agreement (>0.90). The coefficient of variation (CV) was used to ascertain absolute reproducibility as:

\[
CV = \frac{\text{Standard deviation}}{\text{mean}} \times 100.
\]

First the equation was applied in each subject between the days of testing to calculate the general CV average. Limits of agreement and bias of the absolute and relative values between the days were evaluated with Bland–Altman analysis:

\[
\text{Limit of agreement (LC)} = (1.96 \times \text{Standard deviation (SD)}) \pm \text{Mean difference (Mdif)}.
\]

All statistics were considered significant at \( P < 0.05 \). Data were analyzed using the statistical packages SPSS for Windows (v 20; IBM Corporation, Armonk, NY) and MedCalc for Windows (v 12.3.0; MedCalc Software bvba, Mariakerke, Belgium).

Table 2 Descriptive values of variables applied to nonexpert swimmers

<table>
<thead>
<tr>
<th>Variables</th>
<th>Day</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHR (bpm)</td>
<td>1</td>
<td>80.09</td>
<td>7.32</td>
</tr>
<tr>
<td>AHR (bpm)</td>
<td>1</td>
<td>79.26</td>
<td>6.45</td>
</tr>
<tr>
<td>AHR (bpm)</td>
<td>2</td>
<td>173.82</td>
<td>9.79</td>
</tr>
<tr>
<td>RPE</td>
<td>1</td>
<td>8.40</td>
<td>0.78</td>
</tr>
<tr>
<td>RPE</td>
<td>2</td>
<td>8.25</td>
<td>0.72</td>
</tr>
<tr>
<td>NLP</td>
<td>1</td>
<td>6.86</td>
<td>1.36</td>
</tr>
<tr>
<td>NLP</td>
<td>2</td>
<td>7.50</td>
<td>1.79</td>
</tr>
</tbody>
</table>

Abbreviations: BHR, before (test) heart rate; AHR, after (test) heart rate; RPE, rate of perceived exertion; NLP, number of laps performed; bpm, beats per minute.

Table 3 Paired differences between days of the variables applied to nonexpert swimmers

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>SD</th>
<th>SEM</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>t-test</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHR (bpm)</td>
<td>2.56</td>
<td>6.31</td>
<td>1.49</td>
<td>-0.58</td>
<td>5.69</td>
<td>1.72</td>
<td>0.10</td>
</tr>
<tr>
<td>AHR (bpm)</td>
<td>-4.50</td>
<td>7.64</td>
<td>1.91</td>
<td>-8.57</td>
<td>-0.43</td>
<td>-2.36</td>
<td>0.03*</td>
</tr>
<tr>
<td>RPE</td>
<td>0.05</td>
<td>0.78</td>
<td>0.18</td>
<td>-0.32</td>
<td>0.43</td>
<td>0.29</td>
<td>0.77</td>
</tr>
<tr>
<td>NLP</td>
<td>-0.48</td>
<td>0.68</td>
<td>0.15</td>
<td>-0.77</td>
<td>-0.17</td>
<td>-3.21</td>
<td>&lt;0.01*</td>
</tr>
</tbody>
</table>

Note: *Significant difference (\( P \leq 0.05 \)).

Abbreviations: SD, standard deviation; SEM, standard error measurement; 95% CI, confidence interval; BHR, before (test) heart rate; AHR, after (test) heart rate; RPE, rate of perceived exertion; NLP, number of laps performed; bpm, beats per minute.

Results

Table 2 shows the descriptive values of the BHR, AHR, RPE, and NLP variables between the days.

Table 3 shows the mean and standard deviation of the differences between test days, the standard errors of measurement, and the significant differences found between the AHR and NLP.

Table 4 shows the reproducibility on the Pearson’s linear correlation values of \( r > 0.50 \), the consistency between variables at ICC > 0.66, and an absolute CV < 6%.

Figure 2 shows the absolute values of bias and limits of agreement (LC 95%) for BHR, AHR, RPE, and NLP between the days.

Discussion

This study aimed to verify the reproducibility of an aerobic test for nonexpert swimmers and presented favorable conditions for the implementation of the Progressive Swim Test to assess the fitness level by increasing the load beep frequency with the number of laps performed. The results shown for BHR, AHR, RPE, and NLP by \( r \), ICC, SEM, and CV demonstrated that this new test was stable across the interday and intrarater variations found. According to Barbosa et al,\(^{23}\) the consistency of a test becomes greater as...
training adaptation occurs. As the Progressive Swim Test was directed at nonexpert swimmers, the stability was analyzed with caution due to the few studies testing nonexperts, so as not to overstate their results given the amount of research conducted with expert swimmers.

By comparing the results between the days and different evaluators, a decrease in BHR and RPE ($P > 0.05$) and an increase in AHR and NPP ($P < 0.05$) were observed, which showed improvements in the physical conditioning that did not interfere with the variable results. The use of heart rate and RPE parameters gave the possibility of controlling the intensity of aerobic exercises both in research tests and in training. The results obtained in this study of nonexpert swimmers were similar to those observed in water running tests in deep-water pools. The calculation of maximum heart rate, estimated by Karvonen (ie, $220 - \text{age}$), may also be useful in predicting if the swimmer, at the end of the Progressive Swim Test, was within the zone of maximum effort training; AHR in the first and second day was 88% and 89% maximum, respectively, showing that such a test required a performance of intense form by the swimmer seeking to keep up with the speed controlled by progressively more frequent beeps in each lap.

Relative reproducibility was considered to be moderate for the BHR, AHR, and RPE ($r > 0.50, P < 0.04$), and very high for the NLP ($r > 0.90, P < 0.01$). The consistency of the results of AHR and RPE were considered not acceptable (ICC $< 0.70, P < 0.02$), BHR acceptable (ICC $> 0.70, P < 0.01$), and NLP excellent (ICC $> 0.90, P < 0.01$). The absolute reproducibility in all variables was regarded as acceptable (CV $< 6\%$), being less than 10%, and the SEM, (95% CI) showed ±2% actual change in the values of variables. The absolute values of the mean difference

---

**Figure 1** Schematic visual of the Progressive Swim Test for nonexpert swimmers.

**Figure 2** Representation of the differences between paired measurements of before (test) heart rate (BHR), after (test) heart rate (AHR), rate of perceived exertion (RPE), and number of laps performed (NLP) against their mean values.
and limits of agreement (95% CI) were shown graphically in Bland–Altman plots, confirming that there were no heteroscedastic errors. However, one swimmer in the RPE and two in the NLP data were outside the limits of acceptance; in other words, the RPE confirmed a low reproducibility and the NLP reliability was not lost, since the values of Pearson’s $r$ and ICC were presented as excellent. The variables in this study require attention for the use of the Progressive Swim Test, because of the low relative reproducibility values. This attention does not apply to NLP, which showed robustness of the results for its applicability.

The variability of these results can be found in swimming studies that deal with reproducibility. Barbosa et al,23 in a strength test through tied swimming, found high scores on the values of Pearson’s $r$ and ICC and low scores for most of the results of SEM and CV. The same was observed by Albert et al13 in the execution of three types of swimming tests, and by Martin and Thompson,24 in the reproducibility of diurnal variation in submaximal swimming. It is likely that the results of these studies have shown significant values because the tests were conducted with athletes and were moderated by the same appraiser.15,23,24 However, Barbosa et al3 and Zamparo12 confirmed that NLP for nonexpert swimmers are lower than for elite swimmers and that high anthropometric and biomechanical factors are highly associated with performance.

In this way, NLP was a predictor of physical conditioning of the nonexpert swimmers due to the distances of a predetermined time, since studies with continuous testing, intervals testing, and progressive testing use time and distance as criteria of exercise intensity.9,15–18 This new test may also be applied to swimming lessons, due to the progressive requirements of each lap, and the variables of HR and RPE may be useful in the prediction of the physiological conditions of nonexpert swimmers.

There were a number of limitations to the reproducibility of this study: (1) the evaluation of response rate and/or psychophysiological factors only: in the future, selection of other physiological variables (eg, VO$_2$, energy cost, lactate) might give a more detailed answer; (2) it was unable to evaluate the biomechanical responses of the nonexpert swimmers; and (3) the study included only nonexpert swimmers, who were not compared with elite swimmers.

In conclusion, the Progressive Swim Test for nonexpert swimmers suggests a favorable level of reproducibility, with possible application in the indirect assessment of aerobic fitness for the prescription of noncompetitive swimming for fitness.

Disclosure

The authors report no conflicts of interest in this work.

References