Modeling the Links Between Young Swimmers’ Performance: Energetic and Biomechanic Profiles

Tiago M. Barbosa and Mário Costa
Polytechnic Institute Bragança

Daniel A. Marinho
University of Beira Interior

Joel Coelho
Polytechnic Institute Bragança

Marc Moreira and António J. Silva
University of Beira Interior

The aim was to develop a path-flow analysis model for young swimmers’ performance based on biomechanical and energetic parameters, using structural equation modeling. Thirty-eight male young swimmers served as subjects. Performance was assessed by the 200-m freestyle event. For biomechanical assessment the stroke length, the stroke frequency and the swimming velocity were analyzed. Energetics assessment included the critical velocity, the stroke index and the propulsive efficiency. The confirmatory model explained 79% of swimming performance after deleting the stroke index-performance path, which was nonsignificant (SRMR = 0.06). As a conclusion, the model is appropriate to explain performance in young swimmers.

The goal of competitive swimming is to travel the race distance as fast as possible. The identification of the variables that predict swimming performance is one of the main aims of the swimming “science” community. Anthropometrics, hydrodynamics, psychology, pedagogy, medicine and traumatology are some of the main scientific approaches used to understand swimming performance (9). Even so, it is consensual that biomechanics and physiology/energetics are the most determinant areas to enhance performance and achieve high-standard levels in competitive swimming (5).
Some research groups dedicate their attention to understanding the relationships established between energetics and biomechanical variables on adult swimmers, with special emphasis on elite adult swimmers (e.g., 2,3,29). Several research projects explored those relationships and identified the links between them. It was reported that, on one hand, performance is strongly related to several energetic variables, such as oxygen up-take (e.g., 13), energy cost (e.g., 7) or time limit and minimum velocity to achieve maximal oxygen up-take (e.g., 16). On the other hand, energetics variables appear to depend on biomechanical ones, such as speed fluctuation (e.g., 1), swimming velocity (e.g., 2) or stroke mechanics (e.g., 3). In this sense, there are relationships that might be described as a “flow chart” type. These types of exploratory and confirmatory models are often used in other scientific domains, such as economy, genetics or epidemiology. However, to our knowledge, the development of “flow chart” models was never attempted in competitive swimming.

Furthermore, the research dedicated to competitive swimming in young people is reduced in comparison with the one about adult/elite swimmers. Investigation on the relationships between performance, energetics and biomechanics in young swimmers has to change some of the variables analyzed. Not only economical but also ethical issues are raised when children are evaluated. Several parameters commonly assessed in adult swimmers cannot be used in young swimmers due to several reasons. Assessments in young swimmers are less expensive, invasive, complex or time consuming in comparison with the ones assessed in adult/elite swimmers. Nevertheless, on a regular basis, young swimmers’ coaches also perform biomechanical and energetic assessments. Besides, there are assessment batteries suggesting procedures for data collection, its analysis and interpretation for young swimmers in particular. They include both energetic and biomechanical assessment procedures among others (e.g., 6,13,34).

While children are swimming, the relationships established between performance, energetics and biomechanics domains, seem not to be fully understood. Some papers examined the influence of the energy cost of swimming, body composition, drag coefficient and technical parameters on swimming performance (e.g., 22–25,30). There, it was reported that stroke index ($R^2 = .898$), in-water $VO_2$ peak ($R^2 = .358$), and arm span ($R^2 = .454$) were the best predictors of swimming performance in young swimmers (22). Higher levels of leg-kick force, peak $VO_2$, stroke efficiency and muscularity were the most potent variables contributing to the classification of the faster swimmers (23). When scaling for body surface area and body mass, children were less economical than adults; when scaling for body length, children were equally economical and; when considering energy cost per meter and absolute $VO_2$, children were more economical than the adults computing two-sample student’s t test (24). For young swimmers, anthropometric characteristics, stroke frequency and stroke length did not present large association with energy cost in front-crawl swimming ($R < .57$; 30). Stepwise regression showed that biomechanical factors ($R^2 > .32$) best characterized the 400-m event in young females, followed by energetic ($R^2 > .31$) and physical ones ($R^2 > .20$; 25). These researches contributed to a higher level in the state of the art about performance in children swimming. Nevertheless, several evaluated parameters are not often at the disposal of young swimmers’ coaches, such as, energy cost, blood lactate or $VO_2$. So, there is a chance to develop a model to explain young swimmers’ performance based on variables assessed on a regular basis by field practitioners (e.g., 34).
The aim of this research was to develop a path-flow analysis model for young swimmers’ performance based on biomechanical and energetic parameters using structural equation modeling. The theoretical model adopted is presented in Figure 1. This model was developed according to the main review papers about these relationships in competitive swimming (e.g., 5, 26). It was reported that swimming performance is determined by energetic profile and the latter by biomechanics behavior (5). So, the sequence of the path-flow domains was respectively biomechanics-energetic-performance. For biomechanical assessment the most cited variables are the stroke length, stroke frequency and swimming velocity (14). In the energetic domain, aerobic capacity is often estimated through the critical velocity concept in young swimmers (20). Estimation of swimming efficiency has been accomplished in the past by the stroke index (12) and more recently by the propulsive efficiency estimation concept (40) based on the theoretical model of

**Figure 1** — Theoretical path-flow model. SL—stroke length; SF—stroke frequency; CV—critical velocity; SI—stroke index; ηp—propulsive efficiency; βx,y—beta value for regression model between exogenous (x) and endogenous (y) variables; εx—disturbance term for a given endogenous variable; rxy—correlation coefficient between two variables; x → y—variable y depends from variable(s) x; x ← y—variable y is associated to variable x.
Martin et al. (27). All these parameters are easy to record and have an ecological validity for young swimmers’ coaches.

**Methods**

**Subjects**

Thirty-eight male young swimmers with several competitive levels volunteered as subjects (12.53 ± 0.58 years old; 1.56 ± 0.10 m of height; 46.58 ± 9.25 kg of body mass; 14.98 ± 4.95% fat mass by bioimpedance; Tanner stages 1–2 assessed by self-evaluation). Parents and coaches gave their consent for the swimmers’ participation in this study. All procedures were in accordance to the Declaration of Helsinki in respect to Human research. The Institutional Review Board of the Polytechnic Institute Bragança approved the study design.

**Biomechanical Data Collection**

For biomechanical assessment swimming velocity, stroke frequency and stroke length were measured. Each swimmer made a maximal 25-m swim with an underwater start. Subject performed the bout alone with no other swimmer in the same swim lane to reduce the drafting or pacing effects. The swimmers were advised to reduce gliding during the start. Swimming velocity was measured in the middle 15-m as:

\[
\bar{v} = \frac{d}{t}
\]

Where \( \bar{v} \) is the mean swimming velocity, \( d \) the distance covered by the swimmer, \( t \) the time spent to cover such distance and measured with a chronometer by an expert evaluator on the swimming pool deck during the bout. The stroke frequency (SF) was measured with a crono-frequency meter from 3 consecutive stroke cycles, in the middle of the 15-m distance by another expert evaluator on the swimming pool deck during the bout. Stroke length was estimated as (14):

\[
SL = \frac{\bar{v}}{SF} \quad \bar{v} = \frac{d}{t}
\]

**Energetics Data Collection**

The assessment of energetics included the analysis of the stroke index and the propulsive efficiency as swim efficiency estimators as well as critical velocity as an aerobic capacity estimator. Stroke index (SI), considered as one of the swim stroke efficiency indexes was computed as (12):

\[
SI = SL \cdot \bar{v}
\]

Propulsive efficiency (hp) was estimated as being (40):

\[
\eta_p = \left( \frac{\bar{v} \cdot 0.9}{2\pi \cdot SF \cdot l} \right) \cdot \frac{2}{\pi}
\]

Downloaded by Purdue Univ on 09/19/16, Volume 22, Article Number 3
Where \( v \) is the swim velocity, \( SF \) is the stroke frequency and \( l \) is the arm’s length.

The \( l \) was computed trigonometrically measuring the arm’s length and taking into account the average elbow angles during the insweep of the arm pull reported by Zamparo (41) for subjects of similar chronological age and gender. Equation 4 is properly speaking the Froude efficiency. The difference between Froude and propelling efficiency is that the first one does not take into account the effect of the internal mechanical work to total mechanical work production. As reported by Zamparo et al. (40), at the range of swim velocities verified in these swimmers, internal mechanical work is rather low and can be neglected. So, propulsive efficiency is very similar to Froude efficiency.

Critical velocity (CV) was computed based on the swimmers curriculum (10). The CV was assessed based on the 200 m and 800 m freestyle short course events. The mean (\( \pm SD \)) time gap between the swimmers personal records (the difference between the oldest and the most recent personal best) was 2.3 \( \pm 1.2 \) months. The \( CV \) was computed using the slope of the simple linear regression model, plotting the swimming performances over the time and determined by (39):

\[
d = at + b
\]

Equation 5

Where \( d \) is the distance of the swim event, \( a \) is the slope of the fit line, \( t \) is the time spent to cover the distance and \( b \) the y-interception in the origin of the \( xx \) axis.

**Performance Data Collection**

Swimming performance was assessed by time lists of the 200 m freestyle event of official short course competitions of local, regional or national level. The time gap between biomechanical plus energetics assessments and swimming performance was made in less than two weeks.

**Statistical Procedures**

The normality and homocedasticity assumptions were checked respectively with the Shapiro-Wilk and the Levene tests. Descriptive statistics (minimum, maximum, mean and one standard deviation) from all variables were calculated. Pearson correlation coefficients were computed between swimming performance and all remains variables. The statistical significance was set at \( p \leq .05 \).

Structural equation modeling is a mathematical approach for testing and estimating causal relationships using a combination of statistical data and qualitative causal assumptions previously defined by the researcher that will be (or would not be) confirmed. Path analysis is one special case of structural equation modeling. This kind of modeling aims at confirming the existence of relationships between variables and not to identify them. Moreover, this approach, rather than to identify variables, also suggests the kind of interplay (direct, indirect and spurious effects) which exists.

Path-flow analysis was performed with the estimation of linear regression standardized coefficients between the exogenous and endogenous variables. All assumptions to perform the path-flow analysis were taken into account. When appropriate, according to the theoretical model, simple or multiple linear regression models were computed. Standardized regression coefficients (\( \beta \)) were considered. Significance of each \( \beta \) was assessed with the Student’s \( t \) test (\( p \leq .05 \)). The effect
size of the disturbance term, reflecting unmeasured variables, for a given endogenous variable, was $1 - R^2$.

To measure the quality of the model fit, the standardized root mean square residuals (SRMR) was computed:

$$SRMR = \sqrt{\frac{\sum_{i=1}^{p} \sum_{j=1}^{q} (r_{ij} - p_{ij})^2}{p + q}}$$  \hspace{1cm} \text{Equation 6}

Where $r$ is the Pearson correlation coefficients and $p$ the correlation predicted by the model (based on total effect, i.e., the addition of the direct and indirect effects plus spurious effects). SRMR measures the standardized difference between the observed covariance and predicted covariance. It is considered qualitatively if: (i) $SRMR < 0.1$ that the model adjust to the theory; (ii) $SRMR < 0.05$ that the model adjusts very well to the theory and; (iii) $SRMR \sim 0$ that the model is perfect (adapted from 21).

**Results**

Table 1 presents descriptive statistics from all variables studied. Data dispersion, expressed as $1SD$, was moderate-high for almost every variable. This same idea can be supported analyzing the range values. This was especially clear for the swim performance ($130.27 \leq 200 \text{ m Freestyle} \leq 206.27 \text{ s}$), one biomechanical variable ($1.25 \leq SL \leq 2.14 \text{ m}$) and one energetics variable ($17.25 \leq \eta p \leq 34.93\%$).

Table 2 presents the correlation between swim performance and remaining variables analyzed. All variables presented significant association with swim performance, except for $SL$. For the significant ones, correlation coefficients ranged between moderate (e.g., $r_{\text{performance, SF}} = -0.35$) and high (e.g., $r_{\text{performance, CV}} = -0.76$) associations.

Figure 2 presents the confirmatory path-flows for young swimmer’s performance. Almost every partial relationship confirmed the hypothesis. The only exception was the relationship between $SI$ and performance ($\beta = -0.003, p > .05$).

| Table 1  Descriptive Statistics of Biomechanics, Energetics and Swim Performance Variables |
|---------------------------------|-----------------|------------|-----------------|
| SL (m)                          | 1.64            | 0.20       | 1.25            | 2.14            |
| SF (Hz)                         | 0.89            | 0.08       | 0.69            | 1.03            |
| $v$ (m.s$^{-1}$)                | 1.46            | 0.13       | 1.15            | 1.69            |
| SI (m$^2$.c$^{-3}$.s$^{-1}$)    | 2.42            | 0.44       | 1.55            | 3.21            |
| $hp$ (%)                        | 23.31           | 4.53       | 17.25           | 34.93           |
| CV (m.s$^{-1}$)                 | 1.07            | 0.10       | 0.87            | 1.26            |
| 200 m Freestyle (s)             | 156.80          | 17.30      | 130.27          | 206.27          |
Age-Group Swim Performance

The confirmatory model explained 79% of swimming performance for both path-flows including (Figure 2A) and deleting (Figure 2B) SI-performance path with subsequent recomputation of the remaining data. So, it can be stated from a qualitative viewpoint that a high prediction of swim performance was verified. SRMR, quantifying the fit of the model purposed, was 0.10 for the first confirmatory model, including the nonsignificant path (Figure 2A) and 0.06 without that path (Figure 2B). In this sense, the confirmatory path-flow model can be considered suitable of the theory presented.

**Discussion**

The aim of this research was to develop a path-flow analysis model for young swimmers’ performance based on biomechanical and energetics parameters. The main results were that the confirmatory models explained 79% of swimming performance including (SRMR = 0.10) or deleting (SRMR = 0.06) the stroke index-performance path, which was nonsignificant.

It was considered that swimming performance is strongly related to the energetics profile and, in turn, this one to technical level (5). So, the path model designed adopted the same flow sequence: biomechanics-energetic-performance. The analysis of the relationships between the three domains is often made with adult/elite swimmers. However, in young swimmers these interplays seem not to be fully explored. Moreover, “typical” variables to assess biomechanical and energetics profile in young swimmers are not the same as for adult/elite swimmers. With the younger swimmers, apparatus and procedures have to be less expensive, invasive, complex and time consuming. Field practitioners do: (i) biomechanical assessments based on the stroke mechanics ($SF$, $SL$ and $v$); (ii) energetics assessments estimating the swim efficiency ($SI$ and $hp$) and the aerobic capacity ($CV$) and; (iii) swimming performance assessment according to final time in real competition context. All the variables are easy to collect and have ecological meaning to design, control and evaluate the training process. Swimming performance was evaluated based on the 200 m freestyle event. Most of the races in the competitive swimming have distances up to 200 m. Indeed the 200 m event is one of the most challenging ones for young swimmers not only because the technical level (i.e., biomechanical behavior) is relevant but also because of the energetics profile.

**Table 2** Correlation Matrix Between Swim Performance and Remains Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>R</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL (m)</td>
<td>-0.23</td>
<td>NS</td>
</tr>
<tr>
<td>SF (Hz)</td>
<td>-0.35</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>v (m.s$^{-1}$)</td>
<td>-0.52</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>SI (m$^2$.c$^{-1}$.s$^{-1}$)</td>
<td>-0.37</td>
<td>&lt; 0.05</td>
</tr>
<tr>
<td>hp (%)</td>
<td>0.48</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>CV (m.s$^{-1}$)</td>
<td>-0.76</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
Mean data values are somewhat within the range of values reported in the literature for swimmers with similar gender, chronological and biological ages (e.g., 18, 31, 32). In this sense, it can be speculated that our data are in accordance to main literature. Thus, it is possible to give a step further in the “state of the art” about competitive swimming, exploring the links established between performance, energetic and biomechanical variables in young swimmers. Data dispersion was moderate-high for almost every variable. Moderate-high data dispersion will allow analyzing and confirming significant determinants of swim performance for the pooled data of several swimmers with different competitive levels.

All variables assessed presented significant association with swim performance, except for \( SL \). Higher performances were associated with higher \( SF, v, SI \) and \( CV \). Several papers reported similar associations for young (e.g., 20, 22) and adult swimmers (e.g., 8). However, performance was also positive and significantly associated
with ηp. It can be hypothesized that efficiency is related to better performances. As described in equation 4, ηp is directly related to v and inversely to SF. Probably young swimmers achieve higher v increasing SL and SF although with more emphasis on the last one. As a consequence, for a reduced range of swim paces, v increases especially due to SF, and therefore ηp decreases.

Another redundant explanation is that propulsive efficiency can also be described as:

$$\eta_p = \frac{P_d}{P_d + P_{kin}}$$  \hspace{1cm} Equation 7

Where Pd is the mechanical power needed to overcome drag force and Pkin the mechanical power of the kinetic energy transferred to the water from the swimmer.
$P_d$ is computed as:

$$P_d = w_d \cdot SF$$  \hspace{1cm} (Equation 8)

Where $w_d$ is the mechanical work per stroke when submitted to the drag force. So,

$$P_d = K \cdot v^2 \cdot SL \cdot SF = K \cdot v^3$$  \hspace{1cm} (Equation 9)

On the other hand, $P_{kin}$ is computed as:

$$P_{kin} = SF \cdot \frac{1}{2} \sum_{i=1}^{n} m_i (\Delta v_i)^2$$  \hspace{1cm} (Equation 10)

Where $m_i$ is the mass of water displaced by the swimmer trying to propel himself and $v_i$ is the velocity of the added mass of water. So, taking into account equations 7, 9 and 10, the increase of the $SF$ will lead to a decrease of the $\eta_p$.

One of the main aims of swimming research is to identify determinant parameters related to competitive level and to understand how they interplay to enhance performance. This kind of research can be developed (5): (i) comparing different competitive level swimmers; (ii) employing neural network; (iii) computing cluster analysis or; (iv) developing statistical models to predict performance from the swimmer’s profile. In this last case of statistical modeling, several mathematical procedures can be employed, such as computing: (i) associations between performance and other variables (i.e., correlation coefficients); (ii) prediction equations for performance based on independent variables (i.e., simple or multiple regression models) or; (iii) prediction of performance based on several independent variables and how they interplay between themselves (i.e., structural equation modeling). Indeed, path analysis has been used in several knowledge areas, such as economy, genetics or epidemiology. However, its use in sport sciences, mainly in the research related to sports performance is almost nonexistent, despite its relevance. Some of the few exceptions are the works by Fogarty et al. (17) and Feltz et al. (15). In this manuscript path analysis was employed mostly to confirm the type of effects suggested on regular basis in the literature for the relationship between swim performance, biomechanics and energetic. In the present state of the art of competitive swimming, understanding such relationships is more interesting rather than trying to predict and model swimming performance. The performance modeling can be as much challenging as frustrating if based on today’s knowledge.

Almost every path previously hypothesized was confirmed. The only nonsignificant path was the $SI$-performance one ($\beta = -0.003$, $p > .05$), meaning that the direct effect of the energetic variable in the performance is reduced. However, $SI$ has an indirect effect in the swim performance mediated by the $hp$ ($r = .492$, $p < .01$). This means that the $\eta_p$ has a higher capability to predict performance than $SI$, resulting in the exclusion of this one from the model. However, from an ecological point of view, field practitioners can continue using $SI$ as evaluation criteria on regular basis as it has a relationship with swimming performance.

As a measure of the models fit the $SRMR$ was used. Numerically the model is as fit as $SRMR$ is closer to zero. Qualitatively it is possible to define a cutoff value when $SRMR$ is equal or lower than 0.10, meaning that the model adjusts to the theory. Deleting the nonsignificant path $SI$-performance and recomputing all data, $SRMR$ was 0.06. In this sense, the theoretical model after a smooth change was adjusted.
Only variables that are easy to be assessed by field practitioners were selected to be included in the theoretical model. As a result, it seems that such variables and the interplay reported in Figure 2 can be useful for age-group coaches’ performing training control and evaluation.

The determination coefficient of swim performance was 0.79. It can be stated that qualitatively a high prediction of swim performance was verified. Even so, 21% of swim performance cannot be explained by the variables analyzed. A higher performance predictability and even a decrease of SRMR might be achieved including other biomechanical variables such as speed fluctuation (2), segmental kinematics (4); or energetics ones, e.g., VO₂max and minimum velocity to achieve VO₂max (16,22), energy cost of swimming (3), blood lactate (25) and anaerobic parameters (19,35,38) according to the literature. However, the inclusion of such variables could increase the gap between the theoretical model proposed and practice since the procedures and apparatus to assess these variables are not easily at the disposal of young swimmers’ coaches. Moreover, swimming performance is a multifactorial phenomenon. Predictability could also increase if other scientific domains reported in the literature as having strong relationships with swim performance were included in the model, e.g., motor control (33), anthropometry (34,37), hydrodynamics (13,24), genetics (11,28) or psychology (36).

As a matter of fact, some limitations in the study can explain at least part of the prediction model. It might be supposed to increase the model quality if: (i) other biomechanical variables are included in it, such as speed fluctuation or segmental kinematics; (ii) other energetic variables are also included in it, such as oxygen up-take, energy cost, time limit, minimum velocity to achieve VO₂max, oxygen deficit or blood lactate; (iii) instead of estimating propulsive efficiency, assess it using another evaluation method; (iv) include in the model parameters from other knowledge areas, e.g., genetics, motor control, anthropometry, hydrodynamics, psychology, etc. Moreover, it should be stressed that the model is only suitable for young male swimmers performing the 200-m freestyle event in short course.

As a conclusion, the model based on biomechanical and energetic variables, according to the relationships suggested, is appropriate to explain performance in young swimmers. Nevertheless, the model should be expanded, including other variables to increase the prediction level and show a deeper understanding of the swimming performance.

Acknowledgments

The authors would like to acknowledge Ana M. Cruz and Mário C. Marques for their valuable help during data collection.

References


