Journal of Science and Medicine in Sport 16 (2013) 281–285

Contents lists available at SciVerse ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams

Original research

# Data modelling reveals inter-individual variability of front crawl swimming

Gautier Bideault[a], Romain Herault[a], Ludovic Seifert[b,*]

[a] Laboratoire d'Informatique, de Traitement de l'Information et des Systèmes (LITIS), National Institute of Applied Sciences (INSA), France
[b] Centre d'Etude des Transformations des Activités Physiques et Sportives (CETAPS), University of Rouen, France

ARTICLE INFO

*Article history:*
Received 25 January 2012
Received in revised form 22 June 2012
Accepted 3 August 2012

*Keywords:*
Variability
Cluster analysis
Biomechanics
Motor control
Modelling

ABSTRACT

*Objectives:* In accordance with dynamical systems theory, which assumes that motor behaviour emerges from interacting constraints (task, organismic, environmental), this study explored the functional role of inter-individual variability in inter-limb coordination.
*Design:* 63 front crawl swimmers with a range of characteristics (gender, performance level, specialty) performed seven intermittent graded speed bouts of 25 m in front crawl.
*Methods:* Each bout was video-taped with a side-view camera from which speed, stroke rate, stroke length and index of arm coordination (IdC) were analysed for three cycles. Cluster analysis was used to classify the swimmers through speed and IdC values.
*Results:* Cluster analysis and validation showed four profiles of IdC management expressing the swimmers' characteristics as cluster 1: mainly national distance male swimmers, cluster 2: mainly international male sprinters, cluster 3: distinguished by female characteristics, and cluster 4: swimmers with the lowest level of performance.
*Conclusions:* These profiles generated different IdC-speed regression models, which (i) showed how the swimmers adapted their motor behaviour to overcome task constraints and (ii) supported the key idea that there is not a single ideal expert model to be imitated, but rather adapted behaviour emerging from individually encountered constraints.


## 1. Introduction

Variability in human behaviour has long been considered a dysfunctional aspect of motor control,[1] indicative of the amount of noise to be reduced.[2] From this perspective, high skill can be defined as the capability to automatically reproduce the exact same movement. In practice, however, variability in human behaviour occurs at many levels in the training process, suggesting that the achievement of skilled behaviour does not depend on a specific profile. Dynamical systems theory emphasises that variability has a functional role,[2] because it allows for the flexibility required to adapt to a variety of constraints.[3] In this context, variability is a mechanism by which individuals adapt their movements to the interaction of organismic, environmental and task constraints.[4,5] Variability allows performers to explore different motor solutions, facilitating the discovery and adoption of individualised optimal patterns of coordination.

In swimming, several constraints can be manipulated to make a desired behaviour emerge and partially explain inter-individual variability in arm coordination.[6,7] For example, water density and temperature, the direction of water flow, underwater visibility, and waves on the water surface act as environmental constraints that cause more or less aquatic resistance to the forward displacement of the body. Active drag is also related to the individual's properties, and thus individual anthropometric characteristics, locomotor disabilities, passive drag and floatation parameters, strength and muscular fibres, and energetic capacities can be considered as organismic constraints that contribute to inter-individual variability. Last, traditional task constraints like speed, stroke rate, instructions from the coach, and equipment or devices (e.g., paddles, fins) are often manipulated to pace the swimmer or to bring about change in the swimmer's technique. An understanding of how these constraints interact would further elucidate the functional role of inter-individual variability and ultimately contribute to the definition of coordination profiles.[8]

In front crawl, Chollet et al.[9] presented three modes of arm coordination (catch-up, opposition, superposition), which can be used differently, notably as a function of speed (i.e., task constraint). It was also shown that the changes of arm coordination from a catch-up to a superposition mode could relate to the increase of aquatic resistance (i.e., environmental constraint).[7] Beyond the effect of task and environmental constraints on arm coordination changes, recent studies highlighted the effect of organismic constraints. In particular, these studies showed inter-individual

* Corresponding author.
*E-mail address:* ludovic.seifert@univ-rouen.fr (L. Seifert).

1440-2440/$ – see front matter 
http://dx.doi.org/10.1016/j.jsams.2012.08.001

variability regarding gender,[10,11] skill[6] and swim specialty,[13,14] suggesting that the range of arm coordination modes was individual. In other words, all swimmers might not use the three modes of arm coordination in front crawl supporting the theory that there are several ways to adapt arm coordination to increase speed. The main aim of the present study was to revisit the pioneering study of Chollet et al.[9] and show that the arm coordination changes usually observed when speed increases should be reconsidered according to the swimmer's characteristics (especially gender, skill level and swim specialty).

Cluster analysis, by unsupervised machine learning,[15] is an interesting technique to detect patterns within high-dimensional datasets, in our case, the swimmer's characteristics without human intervention and the inherent bias. One significant advantage of movement pattern clustering is that no a priori assumptions about the structure of the dataset are required to identify similar patterns.[8,16,17] Rein et al.[18] explored inter-individual differences in the hook shot of basketball players with various skill levels. Cluster analysis provided several profiles, showing that coordination modes differed between participants as they threw from various distances from the basket.[18] Cluster analysis also showed the functional role of variability in recreational breaststroke swimmers, as some swimmers used an in-phase mode of arms-legs coordination to keep their body close to the water surface, while others used an out-of phase mode to overcome aquatic resistance.[19]

The second aim was to provide a regression model for each cluster of swimmers determined by cluster analysis. Indeed, most biomechanical modelling in swimming concerns performance modelling[20] and the associated parameters (e.g. stroke rate and stroke length).[21,22] However, few studies have tried to model the underlying motor organisation of this cyclic task. On one hand, Seifert et al.[6] modelled the change in arm coordination with speed increases by piecewise linear regression, showing a bifurcation in arm coupling between the slow and fast race paces. On another hand, Hay[22] modelled the stroke frequency-speed and stroke length-speed relationships by quadratic regression for cyclic activities such as swimming, canoeing, kayaking. In that way, Seifert and Chollet[23] showed that the arm coordination-speed relationships of elite swimmers could also be modelled by quadratic regression. Thus, as both the index of coordination and active drag increase with speed square, these authors suggested that arm coordination changes mostly related to environmental constraint (i.e., aquatic resistance). However, these studies only considered the effects of gender and skill level by assigning each swimmer to a category, which prevented the analysis of inter-individual variability.

## 2. Methods

Sixty-three French front crawl swimmers with various characteristics volunteered for this study (Table 1). The protocol was approved by the Rouen university ethics committee and followed the declaration of Helsinki. The protocol was explained to the swimmers who then gave their written informed consent to participate.

The swimmers performed seven intermittent graded speed bouts of 25 m in front crawl. To avoid fatigue effects, each swimmer simulated the seven individual 25 m bouts at paces corresponding to specific race distances: 1500 m, 800 m, 400 m, 200 m, 100 m, 50 m, and maximal speed, with 4 min of rest before the next bout was swum. They started in the water without diving and each bout was self-paced to avoid the speed variations that can arise when swimmers follow a target.

The swimmers were video-taped by two underwater video cameras (Sony compact FCB-EX10L, $f = 50$ Hz), with one camera placed to obtain a frontal view and the other to obtain a side view. The frontal underwater camera was fixed on the edge of the pool, 0.4 m below the water. The side underwater camera was fixed on a trolley and an operator followed the swimmer's head to avoid parallax. Both cameras were connected to a timer, a video recorder and a screen to mix and genlock the frontal and side views on the same screen. A third camera mixed with the side view for time synchronisation video-taped all trials with a profile view from above the water.

The lateral aerial view allowed the calculation of the average speed ($S$ in m s$^{-1}$) over a 10 m distance (from 10 m to 20 m) using the swimmer's head as the marker. Over this distance, a mean period (defined as the time that separates two consecutive entries of the same hand in the water) was determined with the timer on three consecutive arm strokes ($T_{cycle}$) taken in the 10 m central part of the pool. An average stroke rate value ($SR = 1/T_{cycle}$ in Hz) was calculated. The stroke length (SL in m) was calculated from $S$ and SR ($SL = S/SR$).

In line with Chollet et al.,[9] four arm phases per stroke (entry and catch of the hand in the water, pull, push, aerial recovery) were determined from the two underwater views every 0.02 s by three independent operators measuring with a blind technique. Three strokes were analysed. The index of arm coordination (IdC) calculated the time gap between the propulsions of the two arms.[9] The duration of each phase and the IdC were expressed as a percentage of the arm stroke duration. Catch-up mode corresponded to lag time between two propulsive actions (IdC < 0%), opposition mode corresponded to continuity between two propulsive actions (IdC = 0%), and superposition mode corresponded to overlap between two propulsive actions (IdC > 0%).

Statistical analysis was conducted through clustering method. Machine learning method has been implemented using unsupervised training wherein categories of examples were ignored (performance level, gender, and specialty) and only features from each recorded example were known. From these features, the algorithm looked at similar records, classified them into clusters, and labelled each cluster. Then, a regression model was built on each cluster. Sixteen features were used to classify the 63 swimmers: seven bouts of coordination, seven bouts of speed, variation in speed ($S_{max} - S_{min}$), and variation in coordination ($IdC_{max} - IdC_{min}$). The stochastic neighbour embedding (SNE) algorithm[24] was applied to optimise data visualisation. This algorithm projected the data in lower dimensional space while trying to preserve the neighbourhood. It is based on the asymmetric probability $p_{ij}$ that $i$ would pick $j$ as a neighbour:

$$p_{ij} = \frac{\exp(-d_{ij}^2)}{\sum_{k \neq i} \exp(-d_{ik}^2)}$$

The dissimilarities were the scaled squared Euclidean distance:

$$d_{ij}^2 = \frac{\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2}{2\sigma_i^2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are two high-dimensional points and $\sigma_i$ is a normalised value for the entropy of the distribution set by hand. Then, thanks to a dendrogram, cluster hierarchical analysis (CHA) determined several profiles within this cluster of swimmers (as done before[16,17,19]). Robustness was achieved through the strong shape method: the CHA was run with 20 different initialisations and each swimmer was then labelled according to majority rule.

From there, the clustering analysis was validated by two methods. First in line with Breiman,[25] Duda et al.,[15] Rein et al.,[17] clusters and labels were validated by bagging. Bagging consists of repeating this operation several times while excluding a different participant each time and then determining whether the obtained clusters are stable. Stability was assessed by the number of turnover, i.e.,