Paper:
http://dx.doi.org/10.1007/s00421-014-2890-1
Longitudinal Investigation of Training Status & Cardiopulmonary Responses in Pre- and Early-Pubertal Children

M. A. McNarry¹, K. A. Mackintosh¹, K. Stoedefalke²,³

¹College of Engineering, Swansea University, Wales, UK.
²Exercise and Sport Sciences, Colby-Sawyer College, New London, USA.
³School of Sport and Health Sciences, University of Exeter, UK

Dr. M.A. McNarry
College of Engineering, Swansea University
Singleton Park
Swansea, SA2 8PP
Tel 01792 513069
Fax 01792 295676
Email: m.mcnarry@swansea.ac.uk

There are no conflicts of interest or funding to disclose.

Key Words: Pre-pubertal, peak $\dot{V}O_2$, multilevel modelling, stroke volume

Peak $\dot{V}O_2$ - peak oxygen uptake
Abstract

Purpose
The presence of a maturational threshold that modulates children’s physiological responses to exercise training continues to be debated, not least due to a lack of longitudinal evidence to address this question. The purpose of this study was to investigate the interaction between swim-training status and maturity in nineteen trained (T, 10±1 yrs, -2.4±1.9 yrs pre-peak height velocity, 8 boys) and fifteen untrained (UT, 10±1 yrs, -2.3±0.9 yrs pre-peak height velocity, 5 boys) children, at three annual measurements.

Methods
In addition to pulmonary gas exchange measurements, stroke volume (SV) and cardiac output (Q̇) were estimated by thoracic bioelectrical impedance during incremental ramp exercise.

Results
At baseline and both subsequent measurement points, trained children had significantly (P<0.05) higher peak oxygen uptake (Yr.1: T 1.75±0.34 vs. UT 1.49±0.22; Yr.2: T 2.01±0.31 vs. UT 1.65±0.08; Yr. 3: T 2.07±0.30 vs. UT 1.77±0.16 l·min⁻¹) and Q̇ (Yr.1: T 15.0±2.9 vs. UT 13.2±2.2; Yr.2: T 16.1±2.8 vs. UT 13.8±2.9; Yr.3: T 19.3±4.4 vs. UT 16.0±2.7 l·min⁻¹).
Furthermore, the SV response pattern differed significantly with training status, demonstrating the conventional plateau in UT but a progressive increase in T. Multilevel modelling revealed that none of the measured pulmonary or cardiovascular parameters interacted with maturational status, and the magnitude of the difference between T and UT was similar, irrespective of maturational status.

Conclusion
The results of this novel longitudinal study challenge the notion that differences in training status in young people are only evident once a maturational threshold has been exceeded.
Introduction

The age at which young athletes are being encouraged to train increasingly intensely for sporting competitions is ever decreasing, leading to considerable interest in the physiological trainability of children. Whilst it is now widely accepted that children demonstrate significant physiological adaptations to training (Baquet et al. 2003), questions continue to surround those factors that may modulate the magnitude of these adaptations.

For decades, it has been postulated that there may be a maturational threshold below which significant physiological adaptations to training cannot occur and/or a “golden” period during which training has an especially pronounced effect (Katch 1983). The physiological mechanisms for this phenomenon, which are comprehensively reviewed elsewhere (Rowland 1997), are suggested to be changes in the hormonal environment, specifically androgens and growth hormones, which are associated with the “initiation of puberty and that influence functional development and subsequent organic adaptations” (8). Although some studies appear to support this contention, suggesting that the influence of training increases with sexual maturity (Mirwald et al. 1981; Ostojic et al. 2009), others report no relationship between the two variables (Danis et al. 2003; McNarry et al. 2010; McNarry et al. 2011), whilst Weber et al. (Weber et al. 1976) found training to be less influential during the pubertal years. The interpretation of these studies is partly confounded, however, by their predominantly cross-sectional nature, which precludes the appropriate partitioning of the influences of training status from the concomitant influences of growth and maturation. Longitudinal investigations and analyses are therefore warranted to address such issues (Armstrong and Welsman 1994; Mirwald et al. 1981).
Few studies have investigated the longitudinal influence of training in children during the maturational period, with those that have reporting contradictory results. Specifically, using a mixed-longitudinal design, Baxter-Jones et al. (Baxter-Jones et al. 1993) found similar influences of training status irrespective of maturity, whilst Kobayashi et al. (1978) reported significant influences of training status only to be manifest after peak height velocity (PHV). This discrepancy may be attributable to differences in the populations studied (maturation status, aerobic fitness) as well as in their training programmes (type/sport, frequency, intensity, duration). Indeed, Baxter-Jones et al. (1993) showed substantial differences between sports, with aerobic sports demonstrating the most pronounced influence of training status. However, several methodological issues limit the interpretation of these earlier studies: i) fundamental to the analysis of the longitudinal influence of training is appropriately accounting for the concomitant influences of growth and maturation; ii) the hierarchical nature of non-independent, nested data in longitudinal studies, whereby an individual’s responses over time will be correlated with each other, needs to be controlled for by taking into account the dependency of observations. These issues, which may be resolved through the application of allometric scaling and multilevel modelling techniques, remain to be addressed.

The purpose of the present study was to investigate the influence of, and interaction between, training and maturity status on the cardiopulmonary responses to ramp incremental exercise in swim-trained and untrained children over a period of three years. Pivotal, the aim of the present study was to apply novel statistical techniques to appropriately account for growth and maturation and the co-dependence of measurement points. It was hypothesized that significantly higher peak values would be evident in the trained children’s pulmonary gas exchange and cardiovascular responses at each measurement point, and that the magnitude of
difference between trained and untrained children would be similar irrespective of maturity, in accord with recent cross-sectional findings in comparable participant populations (McNarry et al. 2010).

**Materials and Methods**

*Participants and anthropometrics*

In total, 19 trained (T, 10.4 ± 1.1 yrs, 8 boys) and 15 untrained (UT, 9.8 ± 0.9 yrs, 5 boys) children initially participated in this study. One trained boy, and three untrained and two trained girls did not complete the study. The swimmers were all competing at regional level and reported a mean training volume of 6 ± 3, 8 ± 3 and 12 ± 4 hours/week in years 1, 2 and 3, respectively. The age-matched UT children were volunteers from local schools.

Participants were asked to arrive at the laboratory in a rested and fully hydrated state, at least 3 hours postprandial and to refrain from consuming caffeinated drinks in the 6 hours prior to testing. The methods employed during this study were approved by the School of Sport and Health Sciences research ethics committee at the University of Exeter, UK and all participants and their parents/guardians gave written informed consent and assent, respectively.

On each visit, prior to the incremental ramp test, participant anthropometrics were assessed. Standing and seated height were measured to 0.1 cm using a Holtain stadiometer (Holtain, Crymych, Dyfed, UK) and body mass was determined using Avery beam balance scales to 0.05 kg (Avery, Birmingham, UK). Skinfold thickness was assessed three times at five sites around the body (bicep, triceps, subscapular, supra-iliac crest and thigh) by the same researcher for all participants using Harpenden callipers (Baty International, Burgess Hill,
UK), accurate to the nearest 0.2 mm. The mean of the three measurements was taken. Percentage body fat was subsequently estimated based on the tricep and subscapular equations of Slaughter et al. (1988). Age was calculated from the date of the first test, and maturity was estimated using the methods described by Mirwald et al. (2002) and expressed as the estimated time in years from the age at PHV (APHV).

**Experimental procedures**

Participants visited the laboratory three times at approximately yearly intervals. On each of these occasions, the participant completed a ramp incremental test to volitional exhaustion on a cycle ergometer (Lode Excalibur Sport, Netherlands). Prior to each test, the handle bar height, seat height and crank length (cycle ergometer) were adjusted to suit each child.

After a 3 minute warm up, the resistance increased at a pre-determined rate to achieve a test duration of 8-12 minutes. The rate was 10 W·min⁻¹, 12 W·min⁻¹ or 15 W·min⁻¹ for years 1, 2 and 3, respectively. Throughout the tests, the children were instructed to maintain a cadence of 70 rpm. Peak efforts were considered to have been given if, in addition to subjective indications such as sweating, hyperpnoea and facial flushing, there was a consistent reduction in cadence despite strong verbal encouragement. The achievement of conventional secondary criteria were noted but not relied upon given the limitations associated with these criteria (Barker et al. 2009). The present results are therefore accepted as reflecting “peak” \( \dot{V}O_2 \) achieved during the test.

**Experimental measures**

Throughout each test, gas exchange and ventilation (Metalyser 3B Cortex, Biophysik, Leipzig, Germany) and heart rate (Polar S610, Polar Electro Oy, Kempele, Finland) were
measured on a breath-by-breath basis and displayed online. Prior to each test, the gas analyzers were calibrated using gases of known concentration and the turbine volume transducer was calibrated using a 3-litre syringe (Hans Rudolph, Kansas City, MO). The delay in the capillary gas transit and analyzer rise time were accounted for relative to the volume signal, thereby time aligning the concentration and volume signals.

Cardiac output (\(\dot{Q}\)) and stroke volume (SV) were determined non-invasively throughout the exercise test using a thoracic bioelectrical impedance device (PhysioFlow, PF-05 Lab1, Manatec Biomedical, France), previously shown to provide reliable results in paediatric populations (Welsman et al. 2005). The electrodes were positioned on the forehead, neck, xiphoid process and on the left hand side lower ribs, avoiding the abdominal muscles, as suggested to be appropriate for children (Welsman et al. 2005). Prior to testing, blood pressure was measured by the same researcher in triplicate using a manual sphygmomanometer with the participant seated at rest. The mean systolic/diastolic blood pressure was entered into the Physioflow following auto-calibration, which was conducted with the participant seated at rest on the ergometer.

**Data Analysis**

The gas exchange data were interpolated to 1-s intervals and peak \(\dot{V}O_2\) was taken as the highest 10-s stationary average during the test. The gas exchange threshold (GET) was determined by the V-slope method (Beaver et al. 1986) as the point at which carbon dioxide production began to increase disproportionately to \(\dot{V}O_2\) as identified using purpose designed software developed using LabVIEW (National Instruments, Newbury, UK).
Peak stroke volume and cardiac output were taken as the highest 15-s mean value measured during the test. To provide an insight into the balance between oxygen delivery and utilisation, the peak arterial-venous oxygen (\(a - \bar{v}O_2\)) difference was estimated by rearrangement of the Fick equation:

\[
\text{Peak } a - \bar{v}O_2 \text{ difference} = \frac{\text{peak } \dot{V}O_2}{\text{peak } \dot{Q}}
\]

To identify the most appropriate model fit describing the stroke volume response as a function of percentage peak \(\dot{V}O_2\), the linear relationship (\(Y = a + bX\)) was compared to the general quadratic relationship (\(Y = a + bX + cX^2\)) using sum of least-squares and maximum likelihood estimation for linear and non-linear regression, respectively (Graphpad Prism, Graphpad Software, San Diego, CA). The best fitting model was determined on the basis of the \(R^2\) values, the residual sum of squares and the F-value.

The influence of body size was accounted for using analysis of covariance (ANCOVA) on log transformed data to determine the allometric relationship between body mass and peak \(\dot{V}O_2\) and between body surface area and peak SV and \(\dot{Q}\) (Welsman and Armstrong 2000). Common allometric exponents were confirmed for all groups and power function ratios \((Y/X^b)\) were computed. Body surface area was calculated according to the equations of Haycock et al. (1978).

Descriptive statistics (means and standard deviations) for anthropometric and exercise variables were computed for the participants on each test occasion. Differences in these variables between genders and trained and untrained children were determined using two-
way analysis of variance (ANOVA); all analyses were conducted using SPSS 19 (IBM, Chicago, IL) and significance was accepted at $P < 0.05$.

Factors associated with the longitudinal development of peak $\dot{V}O_2$ were investigated using multilevel modelling, which is an extension of multiple regression analysis. Multilevel modelling is preferable to traditional analytical approaches for longitudinal data (e.g. repeated-measures analysis of variance) because, in addition to describing the population mean response, this method recognises and describes variation around the mean at both levels. Specifically, multilevel models can analyse the hierarchical nature of non-independent, nested data by taking into account the dependency of observations (Twisk 2006a). To account for the outcome measures from different time points being nested in participants, a 2-level data structure was used. Timing of the year 2 and year 3 measurements was defined as the first level unit of analysis, and participant was the second level unit of analysis. Association models were used to assess the average effect of training on the outcome variables over the year 2 and year 3 time points, after being adjusted for potentially confounding variables.

The two follow-up measurements (i.e., year 2 and year 3) of each outcome measure were the dependent variables. To estimate the average effect of training on outcome measures, potential confounding variables based on previous research were added to the models as they may influence the change in magnitude of the training effect (Twisk 2006b). Regression coefficients for the group variables (where ‘0’ indicated UT and ‘1’ indicated T), reflected average differences in the outcome measures over time (adjusted for baseline values and covariates). Separate ‘crude’ interaction analyses adjusted for each interaction term, group, time, and baseline value of the outcome measure were initially performed. ‘Adjusted’
interaction analyses (i.e., interaction term included in adjusted multilevel models) were consequently conducted for each effect modifier (Twisk 2006a). Regression coefficients in the main and interaction models were assessed for significance using the Wald statistic. Absolute values were utilised for this analysis rather than allometrically scaled variables to avoid co-linearity in the variables assessed. Analyses were performed using MLwiN 2.26 software (Centre for Multilevel Modelling, University of Bristol, UK).

**Results**

*Anthropometrics*

The anthropometric characteristics of the children changed significantly over time, but there was no significant difference between groups (trained and untrained; Table 1). Participants were classified as pre-pubertal and pubertal according to previous studies cross-referencing age at PHV with invasive or intrusive measures of maturation (Al Alwan et al. 2010; Pantsiotou 2007).

*Influence of training status*

As shown in Table 2, peak $\dot{V}O_2$ was significantly higher in the trained children across all three years ($F_{(1,32)} = 675.91, P < 0.001$). Subsequent tests revealed significant differences between trained and untrained children at year 1 ($t_{(32)} = 2.04, P < 0.02$), 2 ($t_{(28)} = 2.42, P < 0.02$) and 3 ($t_{(26)} = 1.83, P < 0.02$). These differences persisted following allometric scaling ($F_{(1,32)} = 798.96, P < 0.001$; yr 1, $t_{(32)} = 2.49, P < 0.02$; yr 2, $t_{(28)} = 3.20, P < 0.01$; yr 3, $t_{(26)} = 1.99, P < 0.02$) and therefore allometrically scaled peak $\dot{V}O_2$ will be used hereafter. In contrast, there was no influence of training status on the GET regardless of the method of expression.
As shown in Table 2, absolute peak $\dot{Q}$ was higher in the trained children at each year (yr 1, $t(32) = 2.04, P < 0.02$; yr 2, $t(28) = 2.16, P < 0.02$; yr 3, $t(26) = 2.10, P < 0.02$). This difference persisted when scaled to account for body size (yr 1, $t(32) = 2.28, P < 0.02$; yr 2, $t(28) = 2.76, P < 0.01$; yr 3, $t(26) = 2.00, P < 0.02$). Peak SV was similarly significantly higher in the trained children at year 1 and 3 (absolute: yr 1, $t(32) = 1.86, P < 0.02$; yr 3, $t(26) = 1.84, P < 0.02$; scaled: yr 1, $t(32) = 2.30, P < 0.02$; yr 3 $t(26) = 1.74, P < 0.02$). Peak HR achieved during the test did not differ by training status whilst the estimated peak $\Delta$-VO$_2$ difference was significantly higher in the trained children at year two only ($t(26) = 4.71, P < 0.01$).

Influence of time

Irrespective of training status, peak VO$_2$ increased significantly from year 1 to year 2 (T, $t(18) = 8.01, P < 0.01$; UT, $t(14) = 2.76, P < 0.02$). This increase was ameliorated when expressed relative to absolute body mass whilst a significant difference was evident from year 2 to year 3 when the allomaterically scaled peak VO$_2$ was considered (T, $t(18) = 17.53, P < 0.01$; UT, $t(14) = 17.70, P < 0.01$). Peak heart rate (HR) was not influenced by year irrespective of training status. The influence of time on peak SV and $\dot{Q}$ depended on training status: in the trained children, both increased significantly from year 2 to year 3 (SV, $t(18) = 2.41, P < 0.02$; $\dot{Q}$, $t(18) = 7.59, P < 0.01$), whereas in the untrained group neither increased between years with the exception of peak $Q$ from year 1 to year 2 ($t(18) = 6.24, P < 0.01$).

Multilevel modelling

In adjusted analyses, significant between-group intervention effects (i.e., the effect of training status) were observed for peak VO$_2$ ($\beta$ for average intervention effects over time = 19.06 ml·kg$^{-1}$·min$^{-1}, P = 0.011$; see Table 3 for main intervention effects). Children, on average,
increased peak \( \dot{V}O_2 \) by 0.74 and 67.92 ml·kg\(^{-b} \)· min\(^{-1} \) at year 2 and year 3, in comparison to year 1 and year 2, respectively (\( P<0.001 \)). No intervention effects were observed for SV or \( \dot{Q} \).

Peak HR achieved during the test did not differ by time or training status whilst the estimated peak \( \dot{V}O_2 \) difference was influenced by both training status and time. Specifically, maturation status was not associated with any outcome variables (\( P>0.05 \)).

**SV Response Pattern**

The SV response pattern was influenced by training status across all three measurement points. Specifically, as illustrated in Figure 1, SV evidenced a linear response in trained children (proportion of linear responses, year 1, 100\%, year 2, 94\%, year 3, 86\%) compared to a curvilinear response in untrained children (proportion of curvilinear responses, year 1, 79\%, year 2, 83\%, year 3, 67\%). The response pattern was not influenced by the method of expressing SV, therefore scaled values are presented.

**Discussion**

The main findings of this study were that when body size is appropriately accounted for, the significant increase in scaled peak \( \dot{V}O_2 \) observed with training was not related to the presence of a maturational threshold. Specifically, the trained children demonstrated a higher peak \( \dot{V}O_2 \) at all three measurement points. However, while the magnitude of the difference between the trained and untrained children significantly increased over the three year period, multilevel modelling revealed that this increase was not associated with changes in maturity status. This therefore refutes the notion that the onset of puberty represents a trigger point beyond which the influences of training are manifest more strongly.
The higher peak $\dot{V}O_2$ in the trained children agrees with previous studies (McManus et al. 1997; Obert et al. 1996; Rowland et al. 2009a). An increased peak $\dot{V}O_2$ in trained participants is generally associated with an increased peak SV and consequently an increased peak $\dot{Q}$ (Nottin et al. 2002; Obert et al. 2003; Rowland et al. 2002), congruent with the trends found in the current study. Also, in agreement with some (Nottin et al. 2002; Obert et al. 2003; Rowland et al. 2009b), but not all previous studies (Regaieg et al. 2013), peak HR was not influenced by training status, indicating that the increased peak $\dot{V}O_2$ was predominantly related to an elevated SV and $\dot{Q}$. The exception to this was the peak $a - \dot{V}O_2$ difference at year two, which was significantly higher in the trained participants, with this difference subsequently ameliorated at year 3. It is presently unclear why this pattern may have been observed; the multilevel modelling indicates that maturity status was not a significant explanatory variable. Given that the $a - \dot{V}O_2$ difference reflects the balance between $O_2$ delivery and extraction, this increase may be indicative of an improved ratio of perfusion to utilisation and/or of an increased mitochondrial density or content, previously reported in trained children (Eriksson et al. 1973). However, such suggestions remain speculative on the basis of the present data.

The higher peak SV reported in trained children is widely attributed to morphological adaptations of the myocardium, which include an increased left ventricular dimension and mass, and intra-ventricular and posterior wall thickness (e.g. Bianchi et al. 1998; Nottin et al. 2002; Rowland et al. 2009a). However, in agreement with other recent findings (McNarry et al. 2010; Rowland et al. 1998; Unnithan et al. 1997), the current results indicate both quantitative and qualitative differences in SV, as shown by the divergent SV response patterns in the trained and untrained children at all three measurement points. Specifically, while the conventionally accepted SV response pattern of an initial increase until 40-50%
peak $\bar{VO}_2$ followed by a plateau until exhaustion was exhibited by the untrained children, the trained children demonstrated a progressive, linear increase in SV until exhaustion. These divergent patterns thereby suggest an additional or alternative functional basis for the higher SV in trained children. The mechanistic basis of these functional differences with training status is unclear; whilst an enhanced diastolic filling has been suggested as the most likely mechanism in adults, the relative contribution of upstream and downstream factors to this effect remain to be resolved (Rowland 2009), and the applicability of this explanation to younger populations is unclear.

In agreement with our hypotheses and others (Danis et al. 2003; McNarry et al. 2010; Weber 1976), there was no interaction between the influence of training status and maturity in the present study. These findings, which contradict some other studies (Kobayashi et al. 1978; Mirwald et al. 1981; Ostojic et al. 2009), suggest that the onset of puberty does not represent a trigger point that must be surpassed for significant influences of training to be manifest or after which the influences are more apparent. Comparisons to earlier studies are largely confounded, however, by their cross sectional nature or failure to adequately account for the simultaneous effects of growth and maturation on the variables of interest. Indeed, this is also the first study to apply the principles of multilevel modelling to account for the co-dependency of individual’s data within a longitudinal training study design. The importance of accounting for the potential confounding influence of maturation is exemplified in the present study in which it is pertinent to note the earlier maturing of the trained participants. (i.e. closer to age at PHV at the comparable chronological ages). A failure to account for this when making comparisons between the present populations could have resulted in erroneous conclusions being drawn.
It is possible that previous reports of the presence of a maturational threshold (Kobayashi et al. 1978; Mirwald et al. 1981; Ostojic et al. 2009) are actually a reflection of an insufficient training volume in the younger participants, therefore artificially indicating an influence of maturation. Indeed, such a contention is supported by the present study whereby a significant influence of training on scaled peak $\dot{V}O_2$ was demonstrated over the three year period, which also saw training volume to increase two-fold. The dose response relationship between training volume and physiological adaptations largely remains to be elucidated. The mechanistic basis for this increase in peak $\dot{V}O_2$ remains unclear; a central influence of training status is not supported since neither peak SV, HR nor, consequently, $\dot{Q}$ were effected by training status. The increased scaled peak $\dot{V}O_2$ may therefore be a reflection of peripheral adaptations, such as an improved ratio of perfusion to utilisation and/or of an increased mitochondrial density or content, as has been reported elsewhere in trained children (Eriksson et al. 1973). Whilst such adaptations may be supported by the elevated peak $a - \dot{V}O_2$ difference observed in the current trained participants at year 2, the amelioration of this increase at year 3 raises significant questions. Finally, while the present study suggests that the influence of training status is not modulated by the onset of puberty, it is important to highlight that this finding may be sport specific: it is possible that the influence of the hormones associated with the onset of puberty may be more important in more “anaerobic” sports.

Whilst the present study is novel in its longitudinal design and statistical analyses, thereby addressing a pivotal question remaining in the literature, it did nonetheless have several limitations. Firstly, the cross-sectional nature at year one precludes the attribution of the differences observed to training per se; the observed differences in their responses may reflect genetic traits which predisposed the trained children to higher values. Also, although a
range of ages were assessed throughout the study it is pertinent to note that even by year three all the participants were classified as pre-PHV. Thus a limitation of the present study is the absence of a post-pubertal group. While the onset of puberty is reported to precede the age at PHV (Al Alwan et al. 2010; Pantsiotou 2007), future research should seek to extend the present study design to incorporate measurements that span the maturational period. However, it is noted that the present data cannot eliminate the possibility that a maturational threshold occurs later in puberty, although such a suggestion is not even supported by those studies that have purportedly observed a maturational threshold. With regard to the assessment of maturity, additional, more direct, measures would have been informative. Furthermore, whilst the training volume of the swimmers is reported, no assessment of physical activity levels within the untrained controls was completed. The test modality also lacked a degree of commonality to the training modality of the swimmers; an upper body ergometer would have demonstrated a greater similarity. Finally, the present study was not able to address the potentially interesting question regarding the presence of a sex difference in occurrence of a maturational threshold due to the limited sample numbers.

In conclusion, this study provides the novel longitudinal evidence that the influence of training on peak $\dot{V}O_2$ is not altered by the onset of puberty, thereby refuting the notion of a trigger point beyond which the influence of training is manifest to a greater extent. The present results provide additional support for the notion that the superior peak $\dot{V}O_2$ observed in trained children is related to a higher Q due to both morphological and functional adaptations that result in an altered SV response pattern that culminates in a higher peak SV.
References
Mirwald RL, Bailey DA, Cameron N, Rasmussen RL (1981) Longitudinal comparison of aerobic power in active and inactive boys aged 7.0 to 17.0 years. Ann Hum Biol 8: 405-414
Table 1. Participant anthropometrics.

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trained</td>
<td>Untrained</td>
<td>Trained</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>10.4 ± 1.1</td>
<td>9.8 ± 0.9</td>
<td>11.3 ± 1.2</td>
</tr>
<tr>
<td>Stature (m)</td>
<td>1.40 ± 0.09</td>
<td>1.40 ± 0.06</td>
<td>1.47 ± 0.09*</td>
</tr>
<tr>
<td>Mass (kg)</td>
<td>35.1 ± 5.7</td>
<td>32.3 ± 5.1</td>
<td>38.6 ± 5.7*</td>
</tr>
<tr>
<td>BMI</td>
<td>17.8 ± 2.2</td>
<td>16.6 ± 2.2</td>
<td>17.8 ± 1.8</td>
</tr>
<tr>
<td>Lean Body Mass (kg)</td>
<td>29.1 ± 4.3</td>
<td>29.5 ± 6.4</td>
<td>31.8 ± 3.6</td>
</tr>
<tr>
<td>Maturity offset (yrs)</td>
<td>-2.4 ± 1.9</td>
<td>-2.3 ± 0.9</td>
<td>-1.4 ± 0.9*</td>
</tr>
</tbody>
</table>

Mean ± SD. BMI, body mass index. * Significant influence of time within training status group.
Table 2. Peak oxygen uptake and cardiovascular responses in trained and untrained children across the three year study period.

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th></th>
<th>Year 2</th>
<th></th>
<th>Year 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trained</td>
<td>Untrained</td>
<td>Trained</td>
<td>Untrained</td>
<td>Trained</td>
<td>Untrained</td>
</tr>
<tr>
<td>Peak $\dot{V}O_2$ (l·min$^{-1}$)</td>
<td>1.75 ± 0.34</td>
<td>1.49 ± 0.22$^*$</td>
<td>2.01 ± 0.31$^T$</td>
<td>1.65 ± 0.08$^{TT}$</td>
<td>2.07 ± 0.30$^*$</td>
<td>1.77 ± 0.16$^{TT}$</td>
</tr>
<tr>
<td>Peak $\dot{V}O_2$ (ml·kg$^{-1}$·min$^{-1}$)</td>
<td>48.8 ± 7.6</td>
<td>46.8 ± 7.1</td>
<td>51.2 ± 8.1</td>
<td>48.4 ± 6.4</td>
<td>53.5 ± 9.2</td>
<td>46.6 ± 8.1</td>
</tr>
<tr>
<td>Peak $\dot{V}O_2$ (ml·kg$^{-b}$·min$^{-1}$)</td>
<td>160.3 ± 31.0</td>
<td>137.3 ± 20.0$^*$</td>
<td>183.2 ± 24.3$^T$</td>
<td>158.1 ± 17.6$^{TT}$</td>
<td>241.6 ± 25.5$^T$</td>
<td>216.1 ± 21.6$^{TT}$</td>
</tr>
<tr>
<td>GET (l·min$^{-1}$)</td>
<td>0.91 ± 0.21</td>
<td>0.85 ± 0.27</td>
<td>1.05 ± 0.20</td>
<td>0.9+ ± 0.20</td>
<td>1.05 ± 0.21</td>
<td>1.08 ± 0.13</td>
</tr>
<tr>
<td>GET (% peak $\dot{V}O_2$)</td>
<td>55 ± 6</td>
<td>54 ± 8</td>
<td>55 ± 5</td>
<td>53 ± 4</td>
<td>54 ± 8</td>
<td>61 ± 7</td>
</tr>
<tr>
<td>Peak HR (b·min$^{-1}$)</td>
<td>195 ± 6</td>
<td>200 ± 12</td>
<td>193 ± 10</td>
<td>196 ± 10</td>
<td>193 ± 8</td>
<td>199 ± 8</td>
</tr>
<tr>
<td>Peak SV (ml·m$^{-b}$)</td>
<td>73 ± 13</td>
<td>64 ± 8$^*$</td>
<td>66 ± 5$^T$</td>
<td>64 ± 12</td>
<td>82 ± 16$^T$</td>
<td>71 ± 13$^*$</td>
</tr>
<tr>
<td>Peak $\dot{Q}$ (l·m$^{-b}$·min$^{-1}$)</td>
<td>14.0 ± 2.4</td>
<td>12.3 ± 1.9$^*$</td>
<td>13.8 ± 1.3$^T$</td>
<td>12.0 ± 1.9$^{TT}$</td>
<td>15.8 ± 3.1$^T$</td>
<td>13.5 ± 2.3$^*$</td>
</tr>
<tr>
<td>Peak $a - \bar{VO}_2$ diff (ml·dl)</td>
<td>11.8 ± 2.9</td>
<td>11.4 ± 2.5</td>
<td>14.0 ± 2.4$^T$</td>
<td>10.5 ± 1.0$^*$</td>
<td>10.5 ± 1.5$^T$</td>
<td>11.5 ± 1.3</td>
</tr>
</tbody>
</table>

Mean ± SD. $\dot{V}O_2$, oxygen uptake; GET, gas exchange threshold; HR, heart rate; SV, stroke volume; Q, cardiac output, $a - \bar{VO}_2$ diff, arterial-venous oxygen difference. * Significant difference between trained and untrained children within a year; † significant difference relative to previous year within trained or untrained children.
Table 3. Multilevel model analyses of outcome variables

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Constant</th>
<th>Baseline</th>
<th>Training†</th>
<th>Time††</th>
<th>Maturity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak (\bar{\dot{V}O_2} ) (ml·kg(^{-b})·min(^{-1}))</td>
<td>40.70 ± 95% CI 43.80 18.02</td>
<td>0.74*** ± 95% CI 0.29 0.42</td>
<td>19.06** ± 95% CI 14.69 11.94</td>
<td>67.92*** ± 95% CI 20.30 17.36</td>
<td>-5.45 ± 95% CI 6.31 -2.64</td>
</tr>
<tr>
<td>Peak SV (ml·m(^{-b}))</td>
<td>31.96 ± 95% CI 10.25** ± 95% CI 7.35</td>
<td>0.48 ± 95% CI 0.06 0.51</td>
<td>19.17 ± 95% CI 1.10 2.88</td>
<td>17.82 ± 95% CI 1.99 2.99</td>
<td>3.81 ± 95% CI -0.18 0.76</td>
</tr>
<tr>
<td>Peak Q (l·m(^{-b})·min(^{-1}))</td>
<td>18.02 ± 95% CI 10.25** ± 95% CI 7.35</td>
<td>0.06 ± 95% CI 0.51 -0.13</td>
<td>11.94 ± 95% CI 1.10 2.88</td>
<td>17.36 ± 95% CI 1.99 2.99</td>
<td>-0.18 ± 95% CI 0.76 -0.51</td>
</tr>
<tr>
<td>Peak (\bar{\dot{\bar{\bar{V}O_2}}} \text{diff} )</td>
<td>10.98 *** ± 95% CI 4.98</td>
<td>-0.13 ± 95% CI 0.35</td>
<td>3.56 ** ± 95% CI 2.85</td>
<td>-5.48 *** ± 95% CI 3.85</td>
<td>1.08 ± 95% CI</td>
</tr>
</tbody>
</table>

†Reference category = Control group; †† Reference category = Year 2; The Intervention \(\beta\) values represent the estimated difference in outcome variables for trained participants’ against the untrained participants’ when covariates are included in the final model. A positive \(\beta\) value for outcome variables indicates a positive training effect in the trained children compared with the untrained children. * \(p<0.05\), ** \(p\leq0.01\), *** \(p\leq0.001\)

Figure Captions
Figure 1. The stroke volume (SV) response pattern for a representative trained and untrained girl during cycle ergometry (see text for details). $\dot{VO}_2$, oxygen uptake.