REVIEW ARTICLE

Sports Med 2003; 33 (4): 285-300
0112-1642/03/0004-0285/$30.00/0

Secular Trends in the Performance of Children and Adolescents (1980–2000)

An Analysis of 55 Studies of the 20m Shuttle Run Test in 11 Countries

Grant R. Tomkinson,[1] *Luc A. Léger*,[2] *Tim S. Olds*[1] and *Georges Cazorla*[3]

1 Computer Simulation Laboratory, School of Health Sciences, University of South Australia, Underdale, South Australia, Australia
2 Department of Kinesiology, University of Montreal, Centre-Ville Branch, Montreal, Quebec, Canada
3 Association for Research and Evaluation of Physical Activity and Sports, Faculty of Sport Sciences, University of Bordeaux II, Talence, France

Contents

Abstract

It is widely believed that the performance of children and adolescents on aerobic fitness tests is declining. To test this hypothesis, this meta-analysis compared the results of 55 reports of the performance of children and adolescents aged 6–19 years who have used the 20m shuttle run test (20mSRT). All data were collected in the period 1981–2000.

Following corrections for methodological variation, the results of all studies were expressed using the common metric of running speed (km/h) at the last completed stage. Raw data were combined with pseudodata generated from reported means and standard deviations using Monte Carlo simulation. Where data were available on children and adolescents from the same country of the same age and sex, but tested at different times, linear regression was used to calculate rates of change. This was possible for 11 (mainly developed) countries, representing a total of 129 882 children and adolescents in 151 age × sex × country slices.

There has been a significant decline in performance in the 11 countries where data were available, and in most age × sex groups, with a sample-weighted mean decline of 0.43% of mean values per year. The decline was most marked in older age groups and the rate of decline was similar for boys and girls.

There has been a very rapid secular decline in the 20mSRT performance of children and adolescents over the last 20 years, at least in developed countries. The rate of decline is not related to the change in the country's relative wealth, as quantified by per capita gross domestic product (GDP).

It is widely believed that the aerobic test performance of children and adolescents has declined over the last few decades.[1] Sedentary technologies, the easy availability of energy-rich food, and declines in community-based physical activity have been implicated.[2] However, despite a great deal of anecdotal and lay speculation, there have been very few studies of secular trends in the performance of children and adolescents on aerobic fitness tests. This is partly due to the wide variety of tests used (and methodological differences even when the same test has been used), and partly due to the lack of conduits and incentives for data sharing.

Since its description in 1984 by Léger et al.,[3] the 20m shuttle run test (20mSRT) with 1-minute stages has been widely used to assess the aerobic fitness of children and adults. This test consists of a number of 'stages' (also called 'levels'), each lasting about 1 minute and comprising a number of 20m 'laps' (also called 'shuttles'), paced by beeps on a cassette or compact disk (CD). At each stage, the required running speed increases. Each stage includes seven or more laps, depending on the required running speed and exact protocol used. The test has been shown to be a reliable and valid field test to estimate maximal oxygen uptake ($\dot{V}O_{2max}$).[4-6] The 20mSRT has probably been the most widely used test to assess the aerobic fitness of children and adolescents.

Until now, there have been no attempts to cumulate and analyse the results of studies which have used the 20mSRT. This is partly because of variations which have developed in the way the test is administered. There are at least three major variants of the test extant today:

1. Léger et al.'s,[3] original 1-minute protocol, which starts at a speed of 8.5 km/h, and increases in speed by 0.5 km/h each minute.
2. The protocol used by the Eurofit,[7] the Australian Coaching Council,[8] the British National Coaching Foundation,[9] and the American Progressive Aerobic Cardiovascular Endurance Run (PACER)[10] system, among others. In this protocol, participants start at a speed of 8.0 km/h, the second stage is at 9.0 km/h, and thereafter increases in speed by 0.5 km/h each minute.
3. The Queen's University of Belfast protocol,[11] which starts at 8.0 km/h, and increases in speed by 0.5 km/h each minute.

In addition, there are often several different cassettes or CDs, often produced in house, used for the same protocol. Methodological variations on these cassettes (e.g. calling the stage number at the start versus the finish of each stage; using only full minutes versus both full minutes and half minutes to indicate completed stages) will mean that identical performances are reported in different ways. Authors very often appear to have been unaware of the protocol variants, frequently referring inappropriately to the paper of Léger and Lambert[12] in their methods sections (Léger and Lambert actually describe a 20mSRT with 2-minute stages), regardless of which protocol they have actually used. In some reports, the description of the test used and the reference do not match, and in others, there is no indication as to which protocol has been used. Table I shows the structure of each of the three main protocols.

In addition to variation in protocols, there has been variation in how results have been reported.

Table I. Running speed (km/h) and number of laps for each stage of the 20m shuttle run test according to the three major protocol variants

Stage no.[a]	Protocol 1 (Léger et al.[3])		Protocol 2 (Eurofit et al.[7-10])		Protocol 3 (QUB[11])	
	speed (km/h)	no. of laps	speed (km/h)	no. of laps[b]	speed (km/h)	no. of laps
1	8.5	7	8.0	7	8.0	7
2	9.0	8	9.0	8	8.5	7
3	9.5	8	9.5	8	9.0	8
4	10.0	8	10.0	8	9.5	8
5	10.5	9	10.5	9	10.0	8
6	11.0	9	11.0	9	10.5	9
7	11.5	10	11.5	10	11.0	9

a The first stage is sometimes called 'Stage 1' (if 'Stage 1' is called at the start of the stage), and sometimes 'Stage 0' (if 'Stage 1' is called at the end of the first stage).

b In some versions of Protocol 2 (e.g. the Eurofit,[7] Australian Coaching Council[8]) the number of laps per stage varies in places.

QUB = Queens University of Belfast.

Individual results have been reported as the number of completed stages, the running speed at the last completed stage, the number of completed laps (or stages plus laps), the number of minutes the test lasted, or as an estimated maximum oxygen uptake ($\dot{V}O_{2max}$) based on regression equations.[4,13]

There have also been inconsistencies when group results have been presented. Consider the case where, using the Léger et al.[3] protocol (protocol 1 in table I), individual A has completed 5 stages and 4 laps, a total of 44 laps, and individual B has completed 6 stages and 7 laps, a total of 56 laps. If we express our results as the average number of completed laps, the average will be 50. If we express our results as the number of stages completed (5 for A and 6 for B), we will report an average of 5.5 stages, which may be translated as 45 laps.

These methodological issues complicate attempts to meta-analyse or cumulate the results of the many 20mSRT studies which have been conducted, often on very large samples of children and adolescents. In addition, in many cases, these studies have been commissioned by government agencies or sports organisations, and the results have not been published or even analysed. This paper attempts to collate all available studies which have used the 20mSRT with children and adolescents, and expresses the results using the common metric of running speed at the last completed stage. Our intention was to chart the evolution of performance where tests have been conducted on children and adolescents of the same age and sex in the same country across different years. Data on secular trends in performance may complement the surprisingly few published studies[14-16] and many anecdotal reports suggesting precipitous declines in the performance of children and adolescents. Comparison of secular trends between countries, sexes and age-groups may also help to elucidate mechanisms and correlates of changes in test performance.

1. Methods

1.1 Data Sources

Studies were located by searching online databases (Sports Discus, Medline, AustROM, CINAHL, Digital Dissertations, Current Contents) using the following keywords: shuttle, shuttle run, MSFT, 20MST, 20mSRT, beep test, multi-stage. These keywords were use in combination with the following modifiers: child, children, pre-adolescent, adolescent, adolescence, pubescent, pubescence boy, girl, young, youth and infant. When published reports were obtained, all relevant references contained in the studies were followed up. Finally, attempts were made to personally contact the authors of each report to ask if they knew of further studies, and to clarify details of their own study.

A total of 103 candidate studies were located. Of these, 21 were excluded for one or more of the following reasons:

- the authors reported their results in terms of $\dot{V}O_{2max}$ only, which could not be transformed to the common metric we adopted, namely running speed at the last completed stage (see section 1.2);

- results were reported for large, undifferentiated age ranges (e.g. 12–17 years) or combined boys and girls into a single group;
- the test protocol used was unknown;
- more complete data were provided by the study authors;
- the study reported data, or subsets of the data, which had been reported in other located studies;
- permission to use unpublished data was withheld.

Of the remaining 82 studies, 55 were used to determine changes in performance. They included data on 129 882 children and adolescents from 11 countries in 151 age × sex × country slices. The breakdown of sample numbers by country and sex is shown in table II. Of these 55 studies, 75% were papers published in scientific journals and books, 5% were commissioned technical reports, 7% were unpublished postgraduate theses, and 13% were unpublished data.

Age was recorded as the estimated mean age of all children and adolescents in the age × sex slice for each study. In most studies, age was reported as age at last birthday, so we assumed the mean age was 0.5 years after the last completed year (e.g. unless otherwise reported, if a study reported testing 12 year olds, we assumed the mean age was 12.5 years). Three studies covered a range of ages, with a maximum span of 3 years. In these cases, the midpoint of the age range was taken as the sample age. For example, for a study which reported measuring 10–11-year-old children (i.e. anywhere between just 10 and almost 12 years), the sample age was taken as 11.0 years.

Table II. Sample sizes for the cumulated studies by country and sex

Country	Boys	Girls
Australia	20 220	18 160
Belgium	8 958	8 748
Canada	7 228	6 246
France	5 014	4 866
Greece	480	452
Italy	2 567	3 217
Netherlands	164	174
Northern Ireland	2 062	2 334
Poland	11 015	10 145
Spain	8 849	8 527
USA	246	210
Total	**66 803**	**63 079**

Measurement year was recorded as the middle of the year of measurement. For example, if a study reported measuring children in 1996, the measurement year was taken to be 1996.5. If measurements took place over 2 years (e.g. 1996–1997, or anywhere between the very beginning of 1996 and the very end of 1997), the midpoint year (1996.0) was used. In four studies (1.9% of all datapoints), it was impossible to determine the year of measurement. In these cases the measurement year was taken to be 1 year before the publication date of the study. Table III summarises the demographic characteristics of children and adolescents in the studies analysed.

1.2 Data Treatment

To compare studies, it was important to express the results using a common metric, to know which protocol was used, and to standardise the method of analysis. We chose running speed (km/h) at the last completed stage as the common metric across studies. While this allowed us to cumulate studies, it incurred a small cost in terms of comparability. This is because a running speed at the last completed stage of say 9.0 km/h translates into different test histories according to the protocol used. In protocol 1, the participant who finishes at 9.0 km/h will have run one previous minute at 8.5 km/h. In protocol 2, they will have run one previous minute at 8.0 km/h. In protocol 3, they will have run one previous minute at 8.0 km/h, and one minute at 8.5 km/h. It also means that the minimum possible speed will vary between protocols – 8.5 km/h with protocol 1, and 8.0 km/h with protocols 2 and 3. However, because very few participants – particularly among the older age groups – complete only one or two stages, the differences in the early stages would serve largely as different types of warm-up, and the effect would have 'washed out' by later stages.

Table III. A summary of the studies used in this analysis. Shown are the year when measurements were taken, the 20m shuttle run test protocol used, the sex and age range of the children and adolescents tested, and the range of sample sizes for each age × sex slice in the study

Country	Year	Protocol	Sex	Age range (y)	Sample size	Reference
Australia	1994	2	M, F	9–18	39–104	17
Australia	1993	2	M, F	12–17	301–542	18
Australia	1997	2	M, F	9,11,13,15	399–634	19
Australia	1991	2	M, F	11–14	57–256	20
Australia	1998	2	M, F	11–16	339–636	21
Australia	1999	2	M, F	7–12	5–20	22
Australia	1988	1	M, F	12 ± 0.4	527–565	23
Australia	ND	2	M, F	11–12	35–53	24
Australia	1996–98	2	F	8–10	45–79	Mulkearns (personal communication)
Australia	1996	2	M	14	51	25
Australia	1995–2000	2	M, F	12–15	1018–4451	15
Australia	1990	1	M, F	10–12	485–486	26
Belgium	1997	1	M, F	11–18	6–88	27
Belgium	1990	2	M, F	6–18	84–423	28
Belgium	1993,1997	2	M, F	12–18	40–288	29
Belgium	1991–92	2	M, F	10–18	90–247	30
Belgium	1984–85	3	M, F	6–19	9–23	31
Canada	1981	1	M, F	6–17	112–404	3
Canada	1989–90	1	M, F	6–17	75–402	32
France	1992	1	M, F	6–11	14–49	33
France	1997	1	M, F	10–15	8–88	34
France	2000	1	M, F	8–10	3–12	35
France	ND	1	M, F	11–16	20–124	36
France	1987	2	M, F	7–11	342–641	37
France	1996–97	2	M, F	7–11	47–110	38
France	1999	2	M, F	11–18	85–234	Unpublished observations
France	1986	1	M	7,12	15–18	39
Greece	1990–91	1	M, F	6–18	103–297	40
Greece	1992–93	1	M, F	6	264–305	41
Italy	1994	3	M, F	12–14	10–23	42
Italy	1992	3	M, F	12–14	189–259	43
Italy	1995	3	M, F	12–15	23–120	44
Italy	1997	3	M, F	12–19	73–404	45
Italy	1997	3	M, F	14–19	16–68	46
Italy	1985–86	2	F	14–18	44–79	47
Netherlands	1983	3	M, F	12	41	6
Netherlands	1987	3	M, F	12–16	101–285	48
Northern Ireland	1988	3	M, F	15	18–23	49
Northern Ireland	1986	3	M, F	13,15	13–32	50
Northern Ireland	1989–90	3	M, F	12,15	251–258	51
Northern Ireland	ND	3	M, F	7,9,11	91–109	52
Northern Ireland	1990–91	3	M, F	12	50–53	53

Continued next page

Table III. Contd

Country	Year	Protocol	Sex	Age range (y)	Sample size	Reference
Northern Ireland	1986	3	M, F	12–15	32–163	54
Northern Ireland	1988–89	3	M, F	11–18	87–237	11
Northern Ireland	1992–93	3	M, F	15	229–230	55
Poland	1991–92	2	M, F	15–19	150–303	56
Poland	1999	2	M, F	7–19	623–3470	57
Spain	ND	2	M, F	10–19	35–50	58
Spain	1999	2	M, F	13–18	10–70	59
Spain	1984–85	2	M, F	10–18	140–267	60
Spain	1987	2	M, F	9–18	6–303	Rivas (personal communication)
Spain	1986–89	2	M, F	10–15	37–180	61
Spain	1990–94	2	M, F	9–17	54–671	62
US	1995	2	M, F	10	113–144	63
US	1997	2	M, F	10.5 ± 0.5	97–102	64

F = female; **M** = male; **ND** = not determined.

Where the protocol used was uncertain (i.e. when the protocol described did not match that of the reference cited, when the same author reported using different protocols in different studies, or when the protocol was not mentioned) we contacted the study authors for clarification, and also obtained a copy of the actual cassette or CD used in the study. It is important to understand that testers are usually instructed to write down the last stage (or last minute) value they hear called on the cassette. Therefore the way they report performances will differ according to whether the stage number is called at the beginning or end of the stage (that is, whether 'Stage 1' is called at the beginning or the end of the first stage), and whether elapsed time is called on the full minute, or each half minute. There are a number of commercially available cassettes, and many in-house versions. Having the actual cassette used allowed us to correct for these methodological differences, and also to be sure of the protocol used. Personal communication with the study authors also allowed us to check suspected typographical errors, obtain more detailed information (e.g. finer resolution within age groupings), clarify analytical methods, and verify which cassette was used. Several researchers also provided us with the original raw data.

Where raw data were available, the running speed at the last completed stage for each participant was calculated according to the protocol used. If the study reported results as average running speed at the last completed stage, we first determined whether the cassette used called 'Stage 1' at the beginning of the cassette, or at the end of the first stage. If 'Stage 1' was called at the end of the first stage, we retained the reported value. If 'Stage 1' was called at the beginning of the first stage, we subtracted 0.5 km/h (the speed increment from stage to stage) from the reported mean speed.

If the study reported results as the average number of stages completed (or as average minutes), we first clarified whether the cassette called the stage number at the start or end of the stage, and whether the cassette they used reported half minutes as well as full minutes. These factors will affect the conversion of completed stages or minutes to running speeds. Testers are asked to record the last stage (or minute) value called on the cassette. Consider a participant who completes 6 stages and 8 laps. If the cassette uses only full stages called at the end of each stage, this performance will be reported as 6 stages. If the cassette also uses half stages, the performance will be recorded as 6.5 stages. If the cassette uses full stages called at the start of each stage, the performance will be reported as 7 stages. Therefore, if the testers used cassettes with stages called at the start of each stage, we subtracted 1 from the reported stage value. On average, testers using cassettes with half-minute calls will record 0.25 stages longer than those using full stages, so in this

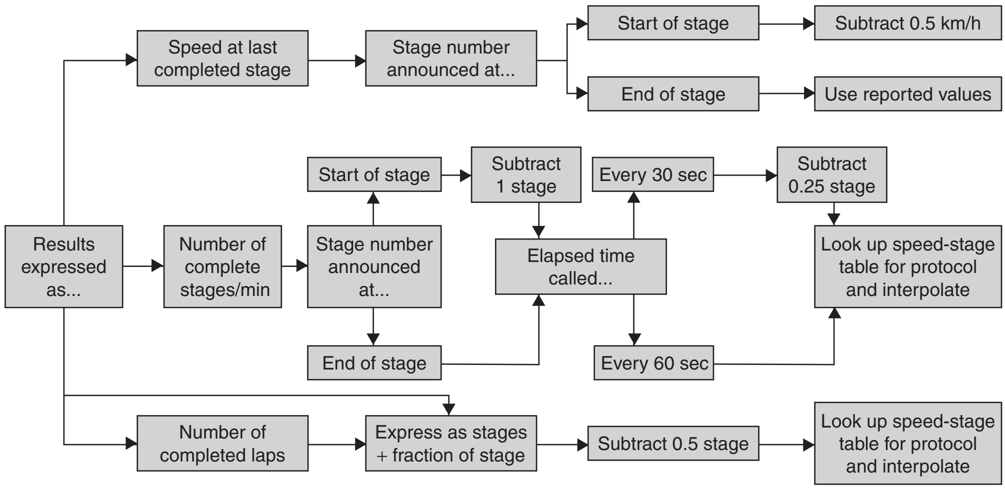

Fig. 1. Steps taken in primary data treatment in order to express the results of all studies using a common metric.

case we subtracted 0.25 from the reported stage value. In all cases where cassettes with half-minute calls were used, we verified procedures with the authors or associates. Having calculated the adjusted number of stages (or minutes), we converted this to a running speed according to the protocol used, as shown in table I. For example, a reported average of 5.5 stages would equate to 10.75 km/h for protocols 1 and 2, and 10.25 km/h for protocol 3. Finally, stages are not always precisely 1-minute long. Some studies called elapsed time on the minute, while others called stages at the end of stages (which might differ by a second or two from the exact minute).[1]

If results were reported as the average number of laps, the average number of completed stages was calculated. For example, if a study reported an average of 43 laps using protocol 1, this was equivalent to 5 + 3/9 = 5.33 stages (since stage 6 has 9 laps). However, this will overestimate the performance compared with a method that reports just the number of completed 'stages'. If we assume that the number of completed laps within each stage is normally distributed about the middle, we can subtract 0.5 from the converted mean completed number of stages. We therefore correct the converted mean to 5.33 − 0.5 = 4.83 stages. The running speed at stage 4 (protocol 1) is 10.0 km/h. The converted mean value of 43 laps was therefore interpolated as 10.0 + 83/100 × 0.5 = 10.42 km/h. Checking this procedure using a large Australian data set (the South Australian Sports Institute's Talent Search data [n = 18 631]),[15] the difference between the mean calculated by using the conversion procedure, and direct calculation from individual final running speeds, was <0.5% for all age × sex slices. The overall data treatment procedures are illustrated in figure 1.

1.3 Statistical Analysis

It was necessary to combine data from different sources to calculate rates of change. To combine data sets, we used Monte Carlo simulation to generate pseudodata. This technique attempts to 'recreate' the unavailable raw data by using a random normal generator to produce datapoints based on reported means and standard deviations (SDs). It assumes that distributions are approximately normal, which was true of the raw data sets that were available.

One study,[20] which constituted 0.7% of all datapoints, reported median rather than mean val-

1 It should be noted that a reported value of, for example, 3.2 stages (using protocol 1) may mean either 3 stages and 2 laps (i.e. 3 stages + 2/8 of a stage = 3.25 stages) or 3.2 stages. This was checked with the study authors where it was unclear.

ues. In this case, we substituted the median values for means. To check the validity of this procedure, we compared the means and medians for all those studies where both were reported. The mean-median difference was on average 0.3% of the mean value.

Two studies did not report SDs, and in another[47] the reported SDs were improbably large. In these cases, constituting 1.3% of all datapoints, we estimated SDs based on the sample-weighted coefficients of variation (CVs) from all other studies. CVs varied with mean running speed and with sex, so separate regressions were calculated relating CV to mean running speed for both sexes. These were used to estimate the CV, and hence the SD, for those studies where values were missing. The mean coefficient of variation was 8.5 ± 2.1%.

Raw data were available from seven studies, comprising 21 817 or 17% of all datapoints. For those studies where raw data were not available, pseudodata were repeatedly generated until the calculated mean differed from the reported mean by less than 0.5%, and the calculated SD differed from the reported SD by less than 2.5%. Pseudodata were then merged with raw data where necessary before analysing the combined data set.

Where two or more data sets were available from the same country for children and adolescents of the same age and sex but measured at different times, linear regression (with year of test as the predictor variable, and running speed at the last completed stage as the response variable) were used to determine the rate of change of performance. This was then expressed as a percentage of the mean value for all datapoints in the regression. For each of these 11 countries, an overall rate of decline for boys and girls was calculated using the sample-weighted mean percentage change across all age × sex slices.

An unpaired t-test was used to compare rates of change in boys and girls. A one-group t-test was used to determine whether the mean of calculated rates of change differed from zero.

2. Results

Changes in performance were calculated in 11 countries, of which ten (Australia, Belgium, Canada, France, Greece, Italy, the Netherlands, Northern Ireland, Spain and the US) were developed economies, and one (Poland) a transitional former socialist economy. In these countries, changes were calculated in a total of 151 age × sex × country slices, covering an average span of 7.5 years (range 1–14.5 years). Each calculation involved as few as two to as many as 12 separate studies. Of the 151 changes, 106 were negative (declines in performance), and 45 positive (improvements in performance). The sample-weighted mean change was −0.046 km/h per year, or −0.43% of the mean running speed per year. This was significantly different from zero ($p \leq 0.0001$). The 95% confidence interval for the overall sample-weighted mean was −0.46 to −0.40%. Considering only the statistically significant changes ($p < 0.05$; 96 of the 151 calculated changes), the sample-weighted mean was −0.062 km/h per year, or −0.57% per year.

A funnel plot of the calculated changes is shown in figure 2. This figure shows that where sample sizes were small, there was a great deal of lability in the calculated rates of change. As sample size increased, the rates of change stabilised near the overall mean of −0.43% per year. For example, of the 70 statistically significant changes where the sample size was above 500, only two were positive, and 68 were negative. There was a similar pattern when the rates of change were plotted against the span of years over which measurements were taken.

When sample-weighted mean rates of change were calculated across all age groups for each of the 11 countries, there was a great deal of variability between countries in the magnitude of secular changes, ranging from a maximum of +0.48% per year for girls from Greece to −1.89% per year for US boys. The mean rates of change for boys and girls from different countries are shown in figure 3.

When sample-weighted mean rates of change were calculated for age-groups rather than for countries, a pattern emerged whereby the rate of decline in performance was reasonably consistent in children (−0.5 to −0.3% per year), but was much greater in adolescents (to about −1.0% per year; figure 4).

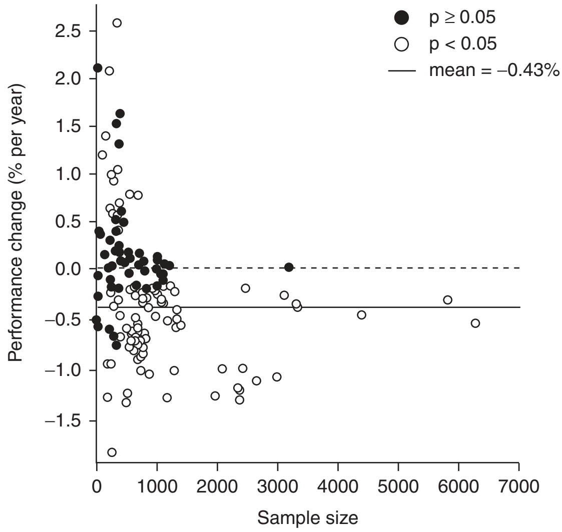

Fig. 2. Funnel plot showing the number in each age × sex × country sample (on the x-axis), and the calculated rate of change in performance (% per year; y-axis). Statistically significant results (i.e. where the slope is significantly different from zero) are shown as open circles, and nonsignificant results as filled circles. The sample-weighted mean rate of change (–0.43% per year) is shown as a thin line.

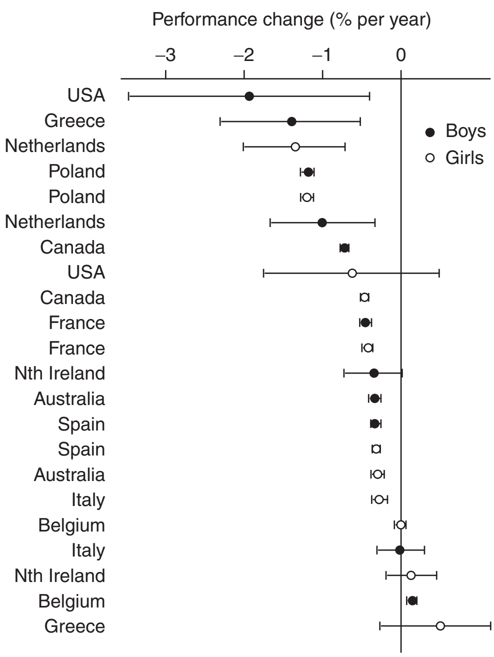

Fig. 3. Sample-weighted mean rates of change in performance (% per year) for boys and girls from the 11 countries where rates of change were calculated. The error bars show the 95% confidence intervals.

3. Discussion

3.1 Methodological Issues

Ideally, a study of this sort would use large raw data sets from random samples taken in different

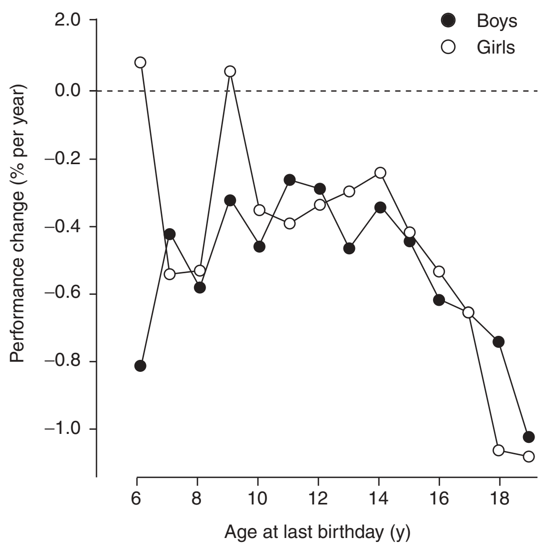

Fig. 4. Sample-weighted mean rates of change in performance (% per year) for boys and girls across the age groups 6–19 years.

countries over regular time intervals. This was clearly not possible in the present case. The studies used a variety of sampling procedures. Some used large stratified random samples on a regional or national basis,[19] while others used convenience samples,[33] occasionally from a single school.[42] Others used proportional sampling based on education systems or geographical areas.[29] In many studies, there was inevitably a degree of self-selection.[15] The participants may have come from different ethnic groups, and from groups differing in their exposure to physical activity. These different sampling procedures obviously raise issues of representativeness – a problem common to almost all cumulation studies. Some of these factors may produce unrepresentatively high values, others low values. It is likely that the sheer number of datapoints in this meta-analysis will dampen irregularities arising from sampling inconsistencies. In any case, the present study offers the most comprehensive picture to date of recent secular trends in the performance of children and adolescents on aerobic tests in the developed world.

The pseudodata method will produce samples similar to the original data sets, providing the original data sets are normally distributed. If the original data sets were skewed, the reported means will overestimate (positive skew) or underestimate (negative skew) the actual medians. However, unless there are systematic changes in skewness over time, the trends in performance scores will not be biased. The few raw data from the studies analysed here where comparisons could be made did not show any consistent change in skewness. Increasingly positive skews in 1600m run times in 10–12-year-old Australian children between 1985 and 1997 have been found,[14] and such changes in skewness could produce artefactual declines in performance using the methods we have employed, or indeed whenever means are compared. However, when the Dollman et al.[14] 1600m values were expressed as speeds rather than times, the distributions were much closer to normal, and there was very little distributional shift between 1985 and 1997. It is hard to quantify how much potential shifts in skewness would affect the results in this study. In the Dollman et al.[14] study, the decline in performance calculated using mean values was at most 10% greater than when it was calculated using the medians.

A review of 13 studies[4-6,11,13,49,52,53,65-69] of the validity of the 20mSRT in normal, untrained children (relative to peak $\dot{V}O_2$) showed a sample-weighted average coefficient of determination (r^2) of 0.5. This shows that a large part of the variability in 20mSRT performance can be explained by the variability in peak $\dot{V}O_2$. In addition, the coefficient of determination varied with age, with r^2 values of 0.2 and 0.7 in 7- and 17-year-old children, respectively. However, it is important to remember that factors other than peak $\dot{V}O_2$ also contribute to 20mSRT performance, particularly in younger children where motor skills and cognitive ability are likely to play a major role. Running efficiency, anaerobic capacity, motivation and social dynamics are all likely to be important.[70] Validation studies of the 20mSRT typically involve children and adolescents across a range of ages tested cross-sectionally. Factors contributing to longitudinal variability within the same age group may be quite different. However, we really have no evidence of systematic changes over time in these other factors. It is the changes in performance, rather than underlying mechanisms, which are the primary focus in this paper.

It is known that many other factors can affect 20mSRT performance. These include environmental differences, clothing and running surfaces, test familiarisation and instructions, and the purpose and context of testing. These were not always reported in the studies analysed here. However, we have no evidence of systematic changes in these factors over time, so while they certainly affect variability in test results, they are unlikely to affect the overall trends.

3.2 Comparisons with Other Studies

There have been few studies which have directly and systematically addressed the issue of secular declines in aerobic test performance, although it is widely believed that children and adolescents are becoming fatter and less fit. Using data from published studies, government reports and personal communications (see next paragraph), we have been able to compare aerobic fitness tests covering a 20-year span from 1980–2000 using tests other than the 20mSRT. They have been conducted in seven developed countries on children and adolescents aged 7–19 years, and have involved over one million study participants. They present a very consistent picture: all have found declines in aerobic performance across all age × sex slices, averaging 0.2–1.1% per year.

Dollman et al.[14] compared a national survey of the performance of Australian children in 1985 to matched measurements taken on 1463 10–11-year-old South Australian children in 1997. Over the 12-year period, performance on the 1600m run test declined by 0.5–0.8% per year. A similar study of 2450 7–10-year-old Tasmanian children found a decline in 1600m run performance between 1985 and 1995 of 0.02–0.4% per year.[71] A recent New Zealand study[72] examined changes in 550m run performance in 5579 children from a single intermediate school (ages 10–14 years) over a 9-year period from 1991–2000. The average decline in performance was 0.2% per year for both boys and girls.

Analysis of data from mass testing programs run by the Japanese Ministry of Education, Science and Culture[73] shows that distance run performance (1000 and 1500m) decreased by 0.4% per year (range 0.3–0.6% per year) in 12–17-year-old Japanese adolescents between 1985 and 1998. In South Korea, tests on over 260 000 10–17-year-old children and adolescents conducted from 1983–1985 and again in 1998 show an average decline of 1.1% (range 0.8–1.8%) per year in 600–1000m run performance.[74] A second South Korean survey[75] using the 1200m run test, found a decline of 1.1% per year (range 0.8–1.4%) in 11 636 12–16-year-old adolescents tested between 1988 and 1998.

Using one-quarter to one mile (402–1609m) runs, Updyke and Willett[16] reported an average decline of 1.1% (range 0.7–1.9%) per year in the aerobic performance of 6–17-year-old American adolescents between 1980 and 1989. In Italy, 11–14-year-old children were tested in 1981[76] and again between 1995 and 2000 (Buonaccorsi, personal communication) using a 1200m run test. A total of 2499 children were tested. On average, performance deteriorated at the rate of 0.9% per year (range 0.4–1.5%). Distance run tests conducted in Poland in 1989[77] were repeated in 1999.[57] A total of 274 014 children and adolescents aged from 7–19 years were tested using 600–1000m runs. The average annual decline in performance was 0.7% (range 0.3–1.1%). These results are consonant with those found in the present study. The rate of decline of performance has, in biological terms, been very rapid.

3.3 What is Causing the Observed Decline in Running Performance?

Performance fitness in running can be reduced by lower aerobic fitness, or increased fatness, or both. Children and adolescents are certainly getting fatter. Cumulated results from 28 reports of changes in body mass index (BMI) in 5–16-year-old children and adolescents since 1980 from the US and Australia show that BMI has increased at a median rate of 0.6% per year[78] – comparable to the rate of decline in aerobic performance found in the present study. There have been similar trends in skinfold thicknesses and body mass.[79] Increased fatness may be the result of increases in energy intake, decreases in energy expenditure, or both. Decreases in energy expenditure are also likely to be associated with reduced vigorous activity, and hence a lower level of aerobic conditioning.

In 1995, Prentice and Jebb[80] asked the question in relation to adults: is the increased incidence of obesity due to gluttony or sloth? We may ask the same question here in relation to declines in performance. There are several lines of evidence suggesting that in children and adolescents the major factor is sloth, i.e. decreased physical activity. This is because:

- energy intake appears to be relatively stable;
- longitudinal studies suggest that activity levels of children and adolescents are declining;
- there appears to have been an increase in inactivity (e.g. television viewing).

Data on secular trends in energy intake are sparse and inconsistent, largely due to sampling and methodological variation. Few data are available on specific trends in the energy intake of children and adolescents, and fewer still on changes in the distribution of intakes, and on changes in saturated fat consumption as part of total intake. However, some 'snapshots' are available. The mean energy intake of 14–15-year-old British children declined by about 20–30% between the 1930s and the 1980s,[81] while there was a 20% reduction in the average intake of UK residents between 1970 and 1990.[80] Based on food balance equations, the Food and Agriculture Organisation reported that per capita energy availability in developed countries did not increase between 1980 and 1998.[82] Other studies, however, have shown opposite trends. Using apparent consumption data, Harnack et al.,[83] for example, estimated that in the US, per capita food availability had increased by 15% between 1974 and 1990.

Not all studies have found a strong link between fitness and physical activity in childhood.[84,85] However, inter-subject variability in fitness is associated with many factors aside from physical activity, including nutritional status and genetic constitution. Furthermore, the quantification of physical activity is inevitably limited in time and perhaps type, and fitness changes may be the result of long-term physical activity patterns. There are also few reliable data on secular trends in physical activity patterns in children. One area where quantifiable data are available is in children's use of transport. Between 1975–1976 and 1989–1994, the percentage of 5–10-year-old children in Britain walking to school fell from 71– 62%, while the percentage travelling by car rose from 15–28%.[86] Between 1985 and 1993, the average yearly distance walked by all British children aged under 15 years fell from 395–317km, and the average yearly distance ridden on bicycles fell from 61–45km.[87] In the US, Department of Transportation data show that between 1977 and 1995 there was a 37% decline in the number of trips made by children on foot or by bicycle.[88] Comparison of data from surveys in 1985 and 1997 showed a decline in the number of organised sports 10–11-year-old Australian children reported playing (from a median of 2 to a median of 1 for boys, and from 1–0 for girls) [Hill AM and Olds TS, unpublished observations]. In the US, enrolment in high school physical education has fallen from 42% in 1991 to 27% in 1997.[89] Heath et al.[90] estimated that between 1984 and 1990, the percentage of US high school students participating in 20 minutes or more of vigorous physical activity three or more times a week declined from 62% to 37%. Given the consistency of declines in the performance and physical activity in children, it appears that both school-based physical education and government intervention strategies are failing, highlighting the pressing need for new approaches.

There is stronger evidence in relation to increasing inactivity both in children and in adults. Some studies on the association between television watching and levels of obesity have found that as the hours of television watched increases, physical activity levels decrease and obesity increases, both in adults and children.[91] However, other studies have been unable to find such links.[92] With increasing television ownership, viewing time has increased: television viewing has doubled in Britain compared with the 1960s.[80] In the US, large use-of-time surveys have been conducted amongst American adults every 10 years since 1965. Reported time spent televiewing among adults has increased from 1.5 hours per day in 1965, to 2.1 hours per day in 1975 and 1985, and 2.3 hours per day in 1995.[93]

Both increases in energy intake and decreases in energy expenditure have been characterised as products of increasing affluence. Reduced active commuting, increased access to sedentary and labour-saving technologies, a trend towards mediatisation and vicarious experience of vigorous activity, and the disintegration of community-based organised sport with increased household and job mobility have all been implicated.[2] Given this, then beyond a certain threshold – associated with freedom from hunger and disease – we would expect to see aerobic performance declining with increasing affluence. To test this hypothesis, we obtained historical data on national gross domestic product (GDP)[94] for the 11 countries where changes in performance levels were calculated. The relationship between the per annum percentage change in GDP (in $US) over the measurement period, and the per annum percentage change in 20mSRT speed was not significant. There was a weak positive relationship ($p = 0.01$) between the absolute annual rate of change of GDP (in $US) and changes in performance – higher rates of increase of GDP in absolute terms were associated with lower rates of performance decline. It should be noted, however, that there were fairly consistent increases in GDP in all the countries in this study in the period from 1980–2000 (2.5 ± 1.1% per annum), so the resolution is quite low.

The decline in performance was not different between boys (sample-weighted mean change = 0.46% per year) and girls (0.41% per year). Nor was the rate of change related to the mid-year of the measurement period over which the change was calculated, suggesting that the secular decline has been relatively constant over the last 20 years. However, there were differences in the rate of decline across age groups. Declines were quite consistent in children, whereas declines in adolescents became increasingly larger. It is hard to know how to interpret this pattern. It may signal the cumulative effect of exposure to environments that are 'toxic for exercise', or the presence of environmental factors which have a greater effect on older rather than younger children, or more optimistically – but less probably given evidence of the constancy of the trend – activity-reducing factors which were stronger in the past than they are now.

3.4 Recommendations

Over the years, many different tests have been used to assess the aerobic fitness of children and adolescents. This has made data pooling impossible. The 20mSRT has many advantages as a fitness test. A large number of children and adolescents can be tested at the same time, it is reliable and valid, and is part of the Eurofit test battery. To facilitate data pooling in the future, we make the following recommendations:

- Methodological drift has led to the results of 20mSRTs being largely incommensurable. A single test protocol should be used (Léger et al.'s[3] protocol 1), or at least the protocol used should be accurately reported.
- More care should be taken in the standardisation and reporting of factors such as environmental conditions, running surfaces, clothing, pre-test instructions, and test familiarisation. Studies should be conducted to assess the effect of these factors on performance.
- A standard multilingual test package should be made available, including a CD, cassette, instruction booklet, reporting forms, and data summaries.
- Results should ideally be expressed as running speed at the last completed one-minute stage, and certainly not as estimated $\dot{V}O_{2max}$ values only.
- When summary statistics are provided, they should be broken down into 1-year age and sex slices based on age at last birthday.
- The year of measurement should be reported.
- Both mean and median values should be reported.
- Researchers should be encouraged to make their raw data available, and an Internet-based data repository should be established.
- Finally, mixed longitudinal studies with standardised sampling procedures should be initiated in as many countries as possible.

Acknowledgements

The authors would like to thank the following people for allowing them access to their data: Scott Baker and the South Australian Sports Institute, Georges Baquet, Serge Berthoin, Alberto Buonaccorsi, Sara Mulkearns and Jeff Walkley and the Australian Council for Health, Physical Education and Recreation. Thank you also to the following people for their kindness in providing extra information about their studies: Natalie Balagué, Mario Bellucci, Michael Booth, Colin Boreham, Jan Borms, Valerie Burke, Dean Cooley, William Duquet, Juan García, Giorgos Georgiadis, Beth Hands, Deborah Hoare, Johan Lefèvre, Matt Mahar, Craig Mahoney, Denis Massicotte, Lars McNaughton, Tony Okely, Ryszard Przeweda, Chris Riddoch, Javier Rivas, Willem van Mechelen and Emmanuel Van Praagh. No sources of funding were used to assist in the preparation of this manuscript. The

authors have no conflicts of interest that are directly relevant to the content of this manuscript.

References

1. Savage MP, Scott LB. Physical activity and rural middle school adolescents. J Youth Adolesc 1998; 27: 245-53
2. French SA, Story M, Jeffery RW. Environmental influences on eating and physical activity. Annu Rev Public Health 2001; 22: 309-35
3. Léger LA, Lambert J, Goulet A, et al. Capacité aérobie des Québécois de 6 à 17 ans -test navette de 20 mètres avec paliers de 1 minute. Can J Appl Sport Sci 1984; 9 (2): 64-9
4. Léger LA, Mercier D, Gadoury C, et al. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci 1988; 6 (2): 93-101
5. Liu NY-S, Plowman SA, Looney MA. The reliability and validity of the 20-meter shuttle test in American students 12 to 15 years old. Res Q Exerc Sport 1992; 63 (4): 360-5
6. van Mechelen W, Hlobil H, Kemper HCG. Validation of two running tests as estimates of maximal aerobic power in children. Eur J Appl Physiol Occup Physiol 1986; 55 (5): 503-6
7. Council of Europe. Eurofit: handbook for the Eurofit tests of physical fitness. Rome: Council of Europe, 1988
8. Australian Sports Commission. 20m shuttle run test: a progressive shuttle run test for measuring aerobic fitness. Belconnen (ACT): Australian Coaching Council, 1999
9. Brewer J, Ramsbottom R, Williams C. Multistage fitness test: a progressive shuttle-run test for the prediction of maximum oxygen uptake. Leeds: National Coaching Foundation, 1988
10. Cooper Institute for Aerobics Research. The Prudential FITNESSGRAM test administration manual. Dallas (TX): Cooper Institute for Aerobics Research, 1992
11. Riddoch CJ. The Northern Ireland health and fitness survey -1989: the fitness, physical activity, attitudes and lifestyles of Northern Ireland post-primary schoolchildren. Belfast: The Queen's University of Belfast, 1990
12. Léger LA, Lambert J. A maximal multistage 20m shuttle run test to predict $\dot{V}O2max$. Eur J Appl Physiol 1982; 49 (1): 1-12
13. Barnett A, Chan LYS, Bruce IC. A preliminary study of the 20m multistage shuttle run as a predictor of peak $\dot{V}O2$ in Hong Kong Chinese students. Pediatr Exerc Sci 1993; 5 (1): 42-50
14. Dollman J, Olds T, Norton K, et al. The evolution of fitness and fatness in 10–11-year-old Australian schoolchildren: changes in distributional characteristics between 1985 and 1997. Pediatr Exerc Sci 1999; 11 (2): 108-21
15. Tomkinson GR, Olds TS, Gulbin J. Secular trends in physical performance of Australian children: evidence from the Talent Search program. J Sports Med Phys Fitness. In press
16. Updyke WF, Willett MS. Physical fitness trends in American youth 1980-1989 [press release]. Bloomington (IL): Chrysler Fund-AAU Physical Fitness Program, 1989
17. Australian Council for Health, Physical Education and Recreation. Australian fitness education award: user's manual and curriculum ideas. Adelaide (SA): ACHPER, 1996
18. Australian Sports Commission. Sport search: norms for sport related fitness tests in Australian students aged 12-17 years. Belconnen (ACT): Australian Sports Commission, 1994
19. Booth M, Macaskill P, McLellan L, et al. NSW schools fitness and physical activity survey 1997. Sydney (NSW): NSW Department of Education and Training, 1997
20. Brewer J, Ramsbottom R, Williams C. Multistage fitness test: a progressive shuttle-run test for the prediction of maximum oxygen uptake. Belconnen (ACT): Australian Coaching Council, 1988
21. Cooley D, McNaughton L. Aerobic fitness of Tasmanian secondary school children using the 20m shuttle run test. Percept Mot Skills 1999; 88 (1): 188-98
22. Hands B. Fitness and motor skill levels of Western Australian primary school children. Perth (WA): University of Western Australia, 2000
23. Jenner DA, Vandongen R, Beilin L. Relationships between blood pressure and measures of dietary energy intake, physical fitness, and physical activity in Australian children aged 11-12 years. J Epidemiol Community Health 1992; 46 (2): 108-13
24. Lloyd KC, Antonas KN. Nutritional habits and fitness levels of schoolchildren. Proceedings of the Nutrition Society of Australia twenty-fourth annual scientific meeting; 2000 Dec 3-6; Fremantle (WA). Adelaide (SA): Nutrition Society of Australia, 2000: 138
25. Okely AD, Gray T, Cotton WG. Effect of an extended stay outdoor education program on aerobic fitness. In: Gray T, Hayllar B, editors. Catalysts for change. Proceedings from the 10th National Outdoor Education Conference; 1997 Jan 20-24; Collaroy Beach (NSW). Sydney (NSW): The Outdoor Education Council, 1997: 206-10
26. Vandongen R, Jenner DA, Thompson C, et al. A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10- to 12-year-old children. Prev Med 1995; 24 (1): 9-22
27. Baquet G, Berthoin S, Padovano C, et al. Effets d'un cycle de course de duree de type intermittent (court-court) sur la condition physique des adolescents. Rev Educ Phys 2000; 40 (2): 51-60
28. Beunen G, Borms J, Vrijens J, et al. Fysieke fitheid en sportbeoefening van de Vlaamse jeugd. Volumen 1: fysieke fitheid van de jeugd van 6 tot 18 jaar. Brussels: Bloso, 1991
29. Lefèvre J, Bouckaert J, Duquet W. De barometer van de fysieke fitheid van de Vlaamse jeugd 1997: de resultaten. Sport (Bloso Brussel) 1998; 4: 16-22
30. Pirnay F. Le baromètre de la condition physique. Sport 1995, 61
31. Poortmans J, Vlaeminck M, Collin M, et al. Estimation indirecte de la puissance aérobie maximale d'une population Bruxelloise masculine et féminine âgée de 6 à 23 ans. Comparaison avec une technique directe de la mesure de la consommation maximale d'oxygène. J Physiol 1986; 81: 195-201
32. Massicotte D. Partial curl-ups, push ups and multistage 20 meter shuttle run, national norms for 6 to 17 year-olds. Montreal: University of Quebec, 1990
33. Baquet G, Berthoin S, Gerbeaux M, et al. Assessment of the maximal aerobic speed with the incremental running field tests in children. Biol Sport 1999; 16 (1): 23-30
34. Baquet G, Berthoin S, Gerbeaux M, et al. High-intensity aerobic training during a 10 week one-hour physical education cycle: effect on physical fitness of adolescents aged 11 to 16. Int J Sports Med 2001; 22 (4): 295-300
35. Baquet G, Berthoin S, Dupont G, et al. Effects d'un entrainement intermittent a haute intensite sur le pic de $\dot{V}O2$ de garcons et de filles prepuberes. Sci Sports 2001; 16 (1 Suppl.): 19S-21S
36. Blonc S, Falgairette G, Fayet J-C, et al. Performance aux tests de terrain d'enfants de 11 à 16 ans: influence de l'âge, du sexe et de l'activité physique. Sci Motricité 1992; 17: 11-7
37. Cazorla G. Batterie France-eval: Mesures, épreuves et barêmes: evaluation des qualités physiques des jeunes Français d'âge scolaire: 7-11 ans. Rapport pour le Secrétariat d'Etat auprès du

Premier Ministre Chargé de la Jeunesse et de Sports. Paris: Ministère de la Jeunne et de Sports, 1987
38. Cazorla G, Portes A, James F. Opération Martinique-eval. Centre d'Evaluation Sport Santé, Fort de France (Martinique). Rapport pour l'Inspection d'Académie de la Martinique. Fort de France: Centre d'Evaluation Sport Santé, 1997
39. Van Praagh E, Bedu M, Falgairette G, et al. Comparaison entre V̇O2 direct et indirect chez l'enfant de 7 et 12 ans. Validation d'une épreuve de terrain. Sci Sports 1988; 3 (4): 327-32
40. Georgiadis G. Evaluation of physical fitness of Greek youth aged 6-18 years [dissertation]. Athens: University of Athens, 1993
41. Manios Y, Kafatos A, Codrington C. Gender differences in physical activity and physical fitness in young children in Crete. J Sports Med Phys Fitness 1999; 39 (1): 24-30
42. Bellucci M. I test Eurofit nella scuola media Mameli di Roma. Alcmeone 1997; 1: 22-7
43. Cilia G, Bellucci M. Eurofit: tests Europei di attitudine fisica. Roma: Istituto Superiore Statale di Educazione Fisica, 1993
44. Cilia G, Bellucci M, Riva M, et al. Eurofit 1995. Roma: Istituto Superiore Statale di Educazione Fisica, 1996
45. Cilia G, Bellucci M, Bazzano C, et al. Eurofit 1997: banche dati per la scuola. Alcmeone 1997; 3: 13-32
46. Cilia G, Bazzano C, Bellucci M, et al. I risultati dei test Eurofit nella scuola Matteuccii di Roma. Alcmeone 1998; 2: 16-20
47. Council of Europe. Évaluation de l'aptitude physique: Eurofit batterie expérimentale. Rome: Council of Europe, 1986
48. van Mechelen W, van Lier WH, Hlobil H, et al. Eurofit: Handleiding met referentieschalen voor 12- tot en met 16-jarige jongens en meisjes in Nederland. Haarlem: Uitgeverij de Vrieseborch, 1991
49. Boreham CAG, Paliczka VJ, Nichols AK. A comparison of the PWC170 and 20-MST tests of aerobic fitness in adolescent schoolchildren. J Sports Med Phys Fitness 1990; 30 (1): 19-23
50. Boreham CAG, Paliczka VJ, Nichols AK. Fitness testing of Belfast schoolchildren. 5th European research seminar on testing physical fitness; 1986 May 12-17; Formia, Italy. Strasbourg: Council of Europe, 1987: 52-7
51. Boreham C, Twisk J, Murray L, et al. Fitness, fatness, and coronary heart disease risk in adolescents: the Northern Ireland Young Hearts Project. Med Sci Sports Exerc 2001; 33 (2): 270-4
52. Mahoney CA, Boreham CAG. Validity and reliability of fitness testing in primary school children. In: Williams T, Almond L, Sparkes A, editors. Sport and physical activity: moving towards excellence. London: E & FN Spon, 1992: 429-37
53. Mahoney C. 20-MST and PWC170 validity in non-Caucasian children in the UK. Br J Sports Med 1992; 26 (1): 45-7
54. Nichols AK, Riddoch CJ. The development of fitness test batteries for use in higher education. In: Trends and developments in physical education. Proceedings of the VIII Commonwealth and International Conference on Sport, Physical Education, Dance, Recreation and Health; 1986 Jul 18-23; Glasgow. London: E & FN Spon, 1986: 378-84
55. Twisk JWR, Boreham C, Cran G, et al. Clustering of biological risk factors for cardiovascular disease and the longitudinal relationship with lifestyle of an adolescent population: the Northern Ireland Young Hearts Project. J Cardiovasc Risk 1999; 6: 355-62
56. Mleczko E, Ozimek M. Rozwój somatyczny i motoryczny mlodziezy Krakowskiej miedzy 15 a 19 rokiem zycia z uwzglednieniem czynników srodowiskowych. Kraków: Akademia Wychowania Fizycznego, 2000
57. Przeweda R. KBN Research Project No. 002-15. Warsaw: Akademia Wychowania Fizycznego, 1999
58. Brito EM, Navarro M, García D, et al. La condición física en la población escolar de gran Canaria (10-19 años). Las Palmas de Gran Canaria, Spain: Excmo. Cabildo Insular de Gran Canaria, 1995
59. García J. La condición fisica en la educación secundaria. Trabajo de investigación [dissertation]. Madrid: Universidad Nacional de Educación a Distancia, 1999
60. Prat JA, Casamort J, Balagué N, et al. Eurofit: la batería Eurofit en Catalunya. Barcelona: Secretaria General de l'Esport, 1998
61. Sainz RM. Aptitudes psiquicas y fisicas: estudio ed la aptitud fisica de los adolescentes de la provincia de Vizcaya y su relacion con la personalidad [dissertation]. Bilbao, Spain: Universidad de Deusto, 1992
62. Sainz RM. La batería Eurofit en Euskadi. Vitoria-Gasteiz, Spain: Instituto Vasco de Educación Fisica, 1996
63. Mahar MT, Rowe DA, Parker CR, et al. Criterion-referenced and norm-referenced agreement between the mile run/walk and PACER. Meas Phys Educ Exerc Sci 1997; 1 (4): 245-58
64. Wolford N. The difference in physical fitness levels of fifth graders according to socioeconomic groups and genders [dissertation]. Lawrence (KS): University of Kansas, 1998
65. Anderson GS. The 1600m run and multistage 20m shuttle run as predictive tests of aerobic capacity in children. Pediatr Exerc Sci 1992; 4 (4): 312-8
66. Armstrong N, Williams J, Ringham D. Peak oxygen uptake and progressive shuttle run performance in boys aged 11-14 years. Res Suppl 1988; 4: 11-2
67. McVeigh SK, Payne AC, Scott S. The reliability and validity of the 20-meter shuttle test as a predictor of peak oxygen uptake in Edinburgh school children, age 13 to 14 years. Pediatr Exerc Sci 1995; 7 (1): 69-79
68. Van Praagh E, Bedu M, Falgairette G, et al. Comparaison entre V̇O2max direct et indirect chez l'enfant de 7 et 12 ans. Validation d'une épreuve de terrain. Sci Sports 1988; 3 (4): 327-32
69. Van Praagh E, Falgairette G, Bedu M, et al. Laboratory and field tests in 7-year-old boys. In: Oseid S, Carlsen K-H, editors. Children and exercise XIII. Champaign (IL): Human Kinetics, 1989: 11-7
70. Armstrong N, Welsman JR. Assessment and interpretation of aerobic fitness in children and adolescents. Exerc Sport Sci Rev 1994; 22: 435-76
71. McNaughton L, Morgan R, Smith P, et al. An investigation into the fitness levels of Tasmanian primary schoolchildren. ACHPER Healthy Lifestyles J 1996; 43 (1): 4-10
72. Dawson K, Hamlin M, Ross J. Trends in the health-related physical fitness of 10-14 year old New Zealand children. J Phys Educ N Z 2001; 34 (1): 26-39
73. Ministry of Education, Science and Culture. Statistical abstract of education, science and culture. Tokyo: Ministry of Education, Science and Culture, 1999
74. Ministry of Education. Statistical yearbook of education. Seoul: Ministry of Education, 1999
75. Ministry of Culture and Tourism. National survey of physical fitness. Seoul: Korean Sport Science Institute, 1998
76. Merni F, Carbonaro G. Test motori. Rome: Comitato Olimpico Nazionale Italiano, 1981
77. Przeweda R, Trzesniowski R. Sprawnosc fizyczna Polskiej mlodziezy w swietle badan z roku 1989. Warsaw: Akademia Wychowania Fizycznego, 1996

78. Harten NR. The evolution of body size and shape in Australian children [dissertation]. Adelaide (SA): University of South Australia, 1999
79. Olds TS, Harten NR. One hundred years of growth: the evolution of height, mass, and body composition in Australian children, 1899-1999. Hum Biol 2001; 73: 727-38
80. Prentice AM, Jebb SA. Obesity in Britain: gluttony or sloth? BMJ 1995; 311 (7002): 437-9
81. Durnin JVGA. Physical activity levels: past and present. In: Norgan NG, editor. Physical activity and health. 34th symposium volume of the Society for the Study of Human Biology. New York: Cambridge University Press, 1992: 20-7
82. Food and Agriculture Organisation [online]. Available from URL: http://apps.fao.org [Accessed 2001 Jan 20]
83. Harnack LJ, Jeffery RW, Boutelle KN. Temporal trends in energy intake in the United States: an ecologic perspective. Am J Clin Nutr 2000; 71 (6): 1478-84
84. Armstrong N, Balding J, Gentle P, et al. Peak oxygen uptake and physical activity in 11- to 16-year-olds. Pediatr Exerc Sci 1990; 2 (4): 349-58
85. Cunningham DA, Stapleton JJ, MacDonald IC, et al. Daily energy expenditure of young boys as related to maximal aerobic power. Can J Appl Sport Sci 1981; 6 (4): 207-11
86. Black C, Collins A, Snell M. Encouraging walking: the case of journey-to-school trips in compact urban areas. Urban Stud 2001; 38: 1121-41
87. Roberts I. Safely to school [letter]? Lancet 1996; 347 (9016): 1642
88. Tudor-Locke C, Ainsworth BE, Popkin BM. Active commuting to school: an overlooked source of children's physical activity? Sports Med 2001; 31 (5): 309-13
89. Pratt M, Macera CA, Blanton C. Levels of physical activity and inactivity in children and adults in the United States: current evidence and research issues. Med Sci Sports Exerc 1999; 31 (11 Suppl.): S526S-S33
90. Heath GW, Pratt M, Warren CM, et al. Physical activity patterns in American high school students: results from the 1990 Youth Behavior Risk Survey. Arch Pediatr Adolesc Med 1994; 148 (11): 1131-6
91. Gortmaker SL, Must A, Sobol AM, et al. Television viewing as a cause of increasing obesity among children in the United States, 1986-1990. Arch Pediatr Adolesc Med 1996; 150: 356-62
92. Robinson TN, Hammer LD, Killen JD, et al. Does television viewing increase obesity and decrease physical activity?: cross-sectional and longitudinal analyses among adolescent girls. Pediatrics 1993; 91 (2): 273-80
93. Robinson J, Godbey G. Time for life: the surprising ways Americans use their time. Pennsylvania (PA): Pennsylvania State University Press, 1999
94. Energy Information Association. [online]. Available from URL: http//www.eia.doe.gov/emeu/international/other.html# IntlGDP [Accessed 2002 Feb 6]

Correspondence and offprints: *Grant R. Tomkinson*, Computer Simulation Laboratory, School of Health Sciences, University of South Australia, Holbrooks Road, Underdale, SA 5032, Australia.
E-mail: grant_tomkinson@yahoo.com.au