Schools With Fitter Children Achieve Better Literacy and Numeracy Results: Evidence of a School Cultural Effect

Richard D. Telford and Ross B. Cunningham
Australian National University

Rohan M. Telford
University of Canberra

Walter P. Abhayaratna
Australian National University

Relationships of academic achievement (government tests) with physical fitness (multistage run), physical activity (pedometers) and percent body fat (dual emission X-ray absorptiometry) were examined at both the aggregate school level and the individual child level using data collected from 757 children in 29 elementary schools. Statistical adjustments included gender, grade and socioeconomic status. Between-school relationships of the academic scores with fitness and physical activity were strong and positive, with some evidence of (negative) relationships with percent body fat. The between-child relationships were weaker, and nonexistent with percent body fat. Stronger between-school than between-child relationships favor the argument that variation in school cultures, characterized by concurrent attention to fitness and academic achievement, might play a more dominant role in explaining these relationships than any direct effect of fitness on academic achievement.

As in other parts of the world, academic standards of schools in Australia are under public scrutiny through recent introduction of national literacy and numeracy testing. At the same time teachers are aware that many children are overweight and insufficiently active. Teachers are therefore faced with making decisions as to how much curriculum time should be committed to physical activity pursuits, given these take time away from classroom teaching of literacy and numeracy.

Consequently there is considerable interest in the relationships between academic achievement, physical activity (PA) and cardio-respiratory fitness (CRF). A recent review considered the weight of cross-sectional evidence to conclude that

Ri. Telford and Abhayaratna are with the Medical School, Australian National University, Acton, ACT, Australia and Clinical Trials Unit, the Canberra Hospital, Garran, ACT, Australia. Cunningham is with the Fenner School for Environment and Society, Australian National University, Acton, ACT, Australia. Ro. Telford is with the Faculty of Health, University of Canberra, Canberra, ACT, Australia.
physically fitter children perform cognitive tasks more rapidly, and that relatively short and specific aerobic exercise training interventions improve executive function, a form of mental processing involving strategically-based decision making (24). Cross-sectional studies also indicate positive associations between academic achievement and PA or CRF, but longitudinal investigations of PA programs in schools are less decisive, reviews concluding they exert neither a positive nor negative impact (16,19,20,24,25).

One of the more recent reports (4) involved a cross-sectional investigation of 259 grade 3 and grade 5 students in four schools. This study supported previous findings of significant relationships between academic achievement and CRF (7,15,18), and made reference to a comment in a previous report (17) that the relationship between academic achievement and PA might be explained, at least in part, by positive student and teacher attitudes to both academic achievement and PA. These workers went on to suggest that further research should target the mechanisms behind the relationships, recommending that socioeconomic status (SES) should be included in any such investigation.

To throw some light on mechanisms, these relationships were explored not only at the between-child level but also at the between-school (cluster) level. Little attention has been given to investigations of the relationships between academic achievement and PA and CRF at the school level. One study that did address relationships in this way was carried out in Hong Kong, where schools are “banded” 1–5 according to the academic standard of its pupils; band 1 accepting the highest achievers and band 5 the lowest (14). Schools achieving better academic results tended to have higher levels of PA, but interpretation of these findings is clouded in that higher banded schools generally had better facilities and greater access to sports programs as well as higher fees, so relationships may have reflected a bias in family and/or school district socioeconomic status (SES).

Our primary objective was to determine the relationships between academic achievement and CRF, PA and percent body fat (%BF) at the school (or cluster) level as well as the child level, and to make use of the relative strength of these relationships to provide insights into underlying mechanisms. In accordance with the above-mentioned recommendations (4) an appropriate index of SES was introduced. Also incorporated was a more appropriate method of measuring %BF than use of its surrogate BMI, and statistical power was increased by involving both more participants and schools.

**Methods**

**Participants**

This study was part of the Lifestyle of our Kids (LOOK) longitudinal study involving 29 elementary (primary) schools in an Australian education jurisdiction as previously described (21) and the characteristics of the sample of 757 participants are set out in Table 1. Approximately 86% of the children had one or both parents of Caucasian descent, 8% of Asian descent, 3% Australian Aboriginal or Torres Strait Islander or 1% Polynesian, and we had no data on 2% of the families.

Schools involved in this study were from the outer suburbs of the territory where the socioeconomic characteristics were similar as indicated by the relatively
Table 1  Descriptive Statistics (Raw, Unadjusted Measurements) for Boys and Girls in Grade 2 and Grade 4. Numbers in Cells Are Means and Standard Deviation, in Brackets.

<table>
<thead>
<tr>
<th></th>
<th>Boys Grade 2</th>
<th>Boys grade 4</th>
<th>Girls grade 2</th>
<th>Girls Grade 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=389</td>
<td>N = 352</td>
<td>N = 368</td>
<td>N = 333</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>8.5 (0.3)</td>
<td>10.5</td>
<td>8.5 (0.3)</td>
<td>10.5</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>130.2 (5.5)</td>
<td>141.6 (6.3)</td>
<td>128.7 (5.4)</td>
<td>140.6 (6.1)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>27.7 (5.3)</td>
<td>36.5 (7.2)</td>
<td>28.6 (5.7)</td>
<td>36.8 (8.3)</td>
</tr>
<tr>
<td>% body fat</td>
<td>22.7 (5.9)</td>
<td>24.5 (6.6)</td>
<td>27.9 (6.3)</td>
<td>29.0 (6.6)</td>
</tr>
<tr>
<td>PA Index** (steps)</td>
<td>108 (11)</td>
<td>102(11)</td>
<td>97(10)</td>
<td>94(9.5)</td>
</tr>
<tr>
<td>Fitness*** (stages)</td>
<td>2.1(1.2)</td>
<td>2.3(1.3)</td>
<td>1.9 (1.1)</td>
<td>2.1(1.2)</td>
</tr>
<tr>
<td>Writing</td>
<td>403(70)</td>
<td>471 (75)</td>
<td>429 (70)</td>
<td>497 (66)</td>
</tr>
<tr>
<td>Reading</td>
<td>410 (89)</td>
<td>490 (74)</td>
<td>428(87)</td>
<td>510 (80)</td>
</tr>
<tr>
<td>Numeracy</td>
<td>417 (78)</td>
<td>555 (75)</td>
<td>410 (74)</td>
<td>477 (68)</td>
</tr>
</tbody>
</table>

*N varies slightly among the variables due to students being absent for some assessments;**the PA index is approximately the square root of the steps per day;***square root of the number of stages reached in the multistage run

confined range of 178 units on the Australian Bureau of Statistics advantage-disadvantage scale (2006), compared with the range of Australian suburbs overall of 653. The relative homogeneity of the schools in our study was also indicated in that all schools were part of a local jurisdiction public education system, receiving similar funding with teachers and principals all drawn from the same overall pool.

Measurements

**Anthropometry.** Height was measured by a portable stadiometer to the nearest 0.001 m and body mass by portable electronic scales to the nearest 0.05 kg. Body composition was measured using dual energy x-ray absorptiometry (DXA, Hologic Discovery QDR Series, Hologic Inc., Bedford, MA, USA) and QDR Hologic Software Version 12.4:7 was used to generate fat mass from which %BF was calculated. For the DXA measurements, the same technician was involved for every measurement in both grades 3 and 5.

**Fitness and Physical Activity.** The 20-m multistage run was used to estimate CRF, being well-established as a field test with children (23). To measure PA, children wore pedometers on their hip for seven consecutive days and a PA index calculated as previously described (21). The same scientist was involved in every PA and fitness measurement, supervising a small team of technicians.

**Literacy and Numeracy.** The literacy and numeracy tests were administered in grade 3 and grade 5 (in 2006 and 2008) by the classroom teachers. In grade 3 the tests were designed and assessed by the local government education authority and in
grade 5 all the tests, including the scaling of results, became a national responsibility (3). It is recognized that while literacy and numeracy are important components of academic achievement, they do not represent a child’s academic achievement in its entirety. Reference to academic achievement in the context of this study is limited to that represented by the government literacy and numeracy test scores.

**Socioeconomic Status (SES).** The Australian Bureau of Statistics Index of Relative Socioeconomic Advantage and Disadvantage (SEIFA) was employed (2). This value is a continuum of advantage (high values) and disadvantage (low values) derived from government Census variables such as income, educational attainment, and employment. The territory in which the schools were situated is relatively homogeneous in this regard, and the average SES index of the suburbs in our study (1085 ± SD 40 and range 982–1160) was higher than the average index of all towns and cities throughout Australia (980 ± 84, 598–1251).

**Procedure**

Government designed literacy and numeracy tests were administered on two occasions, first in grade 3 of elementary school when the average ages of the boys and girls were both 8.5 (SD = 0.3) years and again in grade 5 at the same time of the year. All other assessments were performed in the same order and in the same months before the literacy and numeracy tests administered in grades 3 and 5. The DXA %BF tests, the 7-day pedometer tests, and the CRF tests were completed two months before each of the two literacy and numeracy assessments.

This study was approved by the ACT Health and Community Care Human Research Ethics Committee and the Ethics Committee at The Australian Institute of Sport. Parental and child consent was obtained for all measures in this study.

**Design**

The study design features each school as the semirandomized unit of measurement, facilitating exploration of relationships between academic achievement, PA, CRF and %BF at the level of the school, as well as investigating relationships at the between-child level. Given potential relationships of each of these variables with each other as well as with SES it was important to apply statistical adjustment and an appropriate statistical model.

**Statistical Methods**

The following description is of a statistical model designed to determine relationships at the between-child and between-school levels as well as at the within-child (longitudinal) level. The current study reports on the between-child and between-school relationships, but because another aspect of the LOOK study involved a physical education intervention over successive years (22), it was important to account for any potential effect of this on the between-child and between-school analyses. Consequently, details of the model are presented in its full form.

The data are multilevel and the response variables representing academic achievement vary at three levels; between-school, between-child (within a school),
and within-child. The same applies to the candidate explanatory variables representing PA, CRF and %BF. Other candidate explanatory variables, such as the measure of SES, vary only at the school level.

It is recognized that children within the same school share common experiences, which may make their academic results more homogeneous than those of a random sample of children drawn across schools. Similarly, at the next level, repeat observations on the same child share the same genetic, family and environmental influences, so repeat observations within a child are likely to be more homogeneous than observations between children. It is important that the statistical model reflects the sampling design and so it is necessary to account for the above dependence structure. This suggests that variables School and Child (within a school) should each be regarded as random effects in our model.

A feature of this study is the multilevel data, which allow inferences to be segregated pertaining to regression relationships at different levels. The difference in academic achievement can be estimated across schools within suburbs which differ by one unit in SES score or in the mean of any given explanatory variable (PA, CRF or %BF), and these estimates can be distinguished from effects of explanatory variables on academic achievement for individuals.

**The Statistical Model**

To distinguish effects at the different levels and to take account of the dependence structure, the statistical model has the form set out in the following example using the Writing Score as an example of a literacy and numeracy response variable and PA as an example of the explanatory variable.

The design structure involves three levels, the School, the Child within school, and the repeat observations (i.e., longitudinal observations) on a child.

Writing Score = constant + Group effect + SES effect + PA + school random effect + Sex effect + Sex.Group interaction + PA + child random effect + Year effect + Sex.Year effect + Group.Year effect + PA + possible interactions between fixed effects + within-child random error

where PA denotes the vector of PA means for each school which varies only at the school level; PA is the vector of differences between the PA mean of each child and the school mean, which varies only at the child level; and PA is the vector of differences between repeat observations and the relevant child PA mean, which varies only at the observation level.

PA + PA + PA represents the totality of components of the original vector of PA.

Similar models apply when PA is replaced by the other explanatory variables CRF and %BF. This mean or fixed effect model can be readily extended and interpreted within this framework. Using PA as the example variable again, the current paper is concerned directly with PA and PA, the relationships at the child and the school level.

The above model fits within the general framework of general linear mixed models (9). Restricted maximum likelihood is used to estimate variance components and weighted least squares for estimating fixed effects. Statistical significance of effects was assessed by calculating adjusted Wald statistics (12). Explanatory
variables PA and CRF were scaled by square roots to better meet linearity assumptions. General model checking procedures were routinely used to identify aberrant data and to check the model assumptions.

Results

Table 1 presents summary data for each relevant variable for the boys and girls in grades 3 and 5. The number of observations for different characteristics varies slightly, as some children were absent from school on some of the assessment days or did not return parental consent forms on time. Statistical analyses are not included in Table 1, as the values included are raw, not adjusted values.

The statistical relationships between the response variables, Numeracy, Reading and Writing with each of the explanatory variables, both at the between-school and between-child levels are shown in Table 2. The relationships between each of the response variables and explanatory variables PA, the square root of CRF, and %BF were adjusted for each other, as well as for SES and gender. While it had little impact on our results, exploration of these data revealed evidence of a random school by year interaction effect (variation between schools in the changes in literacy and numeracy scores) so this was incorporated as an additional term in the statistical model.(Figures 1, 2, and 3)

Between-School Relationships

School reading scores were significantly and positively associated with both the school CRF ($p < .001$) and PA ($p = .01$) and there was evidence ($p = .05$) that reading score was negatively related to %BF at the school level. School numeracy scores were also significantly and positively associated with school CRF levels ($p < .001$) and PA ($p = .02$) and associated negatively with %BF ($p = .02$).

School writing scores were associated with the school CRF ($p < .001$) with little evidence of any association with PA ($p = .17$) or %BF ($p = .23$).

Between-Child Relationships

Child reading scores were not associated with CRF, PA or %BF ($p = .82, 0.80, 0.62$ respectively).

Child numeracy scores were associated with CRF ($p = .03$) with weak evidence of an association with PA ($p = .09$) and no evidence of any relationship with %BF ($p = .61$).

Child writing scores were related to their CRF ($p = .03$) with strong evidence of a relationship with PA ($p = .007$), but there was no evidence of any relationship with %BF ($p = .19$).

Consideration of SES was important, but when we excluded SES in the model there was no impact of any consequence on the inferential status (statistical significance or nonsignificance) of any relationship.
Table 2  Table of Effects ($\beta$, (SE), and p Values) for Between-School and Between-Child Relationships of the Literacy and Numeracy Scores With the Key Explanatory Variables Physical Activity Index (PA), Cardio-Respiratory Fitness (Square Root of CRF) and Percent Body Fat (%BF). Relationships Were Adjusted for These and Design Variables as Well as SES

<table>
<thead>
<tr>
<th>Effect</th>
<th>Writing Score</th>
<th>SE</th>
<th>p</th>
<th>Reading Score</th>
<th>SE</th>
<th>p</th>
<th>Numeracy Score</th>
<th>SE</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PA_s$</td>
<td>2.1</td>
<td>1.5</td>
<td>0.17</td>
<td>-0.07</td>
<td>0.29</td>
<td>0.80</td>
<td>4.11</td>
<td>1.7</td>
<td>0.02</td>
</tr>
<tr>
<td>$PA_c$</td>
<td>0.62</td>
<td>0.23</td>
<td>0.007</td>
<td>0.07</td>
<td>0.29</td>
<td>0.80</td>
<td>0.43</td>
<td>0.25</td>
<td>0.09</td>
</tr>
<tr>
<td>$\sqrt{CRF}_s$</td>
<td>145</td>
<td>29</td>
<td>&lt; 0.001</td>
<td>143</td>
<td>35</td>
<td>&lt; 0.001</td>
<td>142</td>
<td>37</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>$\sqrt{CRF}_c$</td>
<td>14.9</td>
<td>7.0</td>
<td>0.03</td>
<td>2.0</td>
<td>8.7</td>
<td>0.82</td>
<td>17.3</td>
<td>7.7</td>
<td>0.03</td>
</tr>
<tr>
<td>$%BF_s$</td>
<td>-3.0</td>
<td>2.4</td>
<td>0.23</td>
<td>-5.5</td>
<td>2.7</td>
<td>0.05</td>
<td>-6.5</td>
<td>2.6</td>
<td>0.02</td>
</tr>
<tr>
<td>$%BF_c$</td>
<td>-0.47</td>
<td>0.36</td>
<td>0.19</td>
<td>0.23</td>
<td>0.47</td>
<td>0.62</td>
<td>-0.21</td>
<td>0.41</td>
<td>0.61</td>
</tr>
</tbody>
</table>

The subscript 's' refers to the effect at the school level. The subscript 'c' refers to the effect at the child level.
Figure 1 — Numeracy test scores and OCRF (the square root of the number of run stages completed). Open circles represent the adjusted mean values for each school. The thick black line is line of best fit at the school level. The lines through the open circles depict the relationship between children within each school. The dots are values for individual children.

Discussion

The most important and novel finding of this study was the strong evidence of positive relationships at the school level between the literacy and numeracy scores and cardio-respiratory fitness. Further evidence of the between-school association of academic achievement with the physical measures was provided by the positive relationships between PA and the reading and numeracy scores together with their negative relationships with %BF.

On considering why schools with fitter children achieved better literacy and numeracy scores, it is acknowledged that neither between-school (nor between-child) associations per se permit unequivocal claims of causality. However support for such a premise does arise from previous literature linking CRF and PA with cognitive processes and brain function in children. For example, aerobically fitter children perform better in tasks associated with cognitive function (6) and neuro-electrical studies provide supporting evidence of the influence of fitness on cognition (10). A review of the literature (24) prompted these authors to write that
“exercise training programs may prove to be simple, yet important, methods of enhancing aspects of children’s mental functioning”.

The current findings would seem to provide support for the premise that increasing a child’s fitness may improve his or her academic performance. On the other hand, it is possible that schools better able to develop children physically are coincidentally and independently better at developing them academically. Extending the suggestion that child and/or teacher attitudes may explain, at least in part, the relationships between academic and physical measures (17), the suggestion is made that schools may vary in what might be described as “school culture”, a set of attitudes and behavior transcending both classroom and playground.

Should the relationships between academic achievement and CRF (or PA) be stronger between schools than between children within a school, this would lend more support for the school cultural explanation than for any direct effect of exercise or fitness on brain function. As shown in Table 2, associations were much stronger between schools than between children; these stronger associations existing despite less variability between schools than between children. Therefore, more support is
provided for a school cultural influence than for a direct causal relationship between academic performance and fitness or PA.

However, it does remain that the true explanation may lie somewhere in between. As alluded to in the review (24), animal studies indicate that enriched and challenging environments promote neurological development and cognitive function. Furthermore, a report from the current group found that a two-year intervention of physical education improved literacy and numeracy development without any detectable increases in CRF or PA (22). This prompted the suggestion that the academic improvements associated with the intervention may have been the result of enhanced communication and movement challenges provided for the children through activities requiring concentration and creativity. Referring again to the current study, similar mechanisms may apply at the school administration and teaching staff levels; processes adopted within a school to promote increased fitness and PA may help to create the kind of school environment more conducive to learning in young children. Further exploration of this line of reasoning might

**Figure 3** — Writing test score and ÖCRF (the square root of the number of run stages completed). Open circles represent the adjusted mean values for each school. The thick black line is line of best at the school level. The lines through the open circles depict the relationship between children within each school. The dots are values for individual children.
benefit from an interdisciplinary approach, incorporating elements of pedagogy, exercise science and neuropsychology.

The question arises as to how variation in school cultural effects might arise. A consistent influence from year to year on teaching philosophies, practices and attitudes to physical activity is likely to have arisen through the school principal and the senior executive teachers, together with the School Board of which they form a part. Well before any data analysis, the research team visiting the schools regularly over a four year period reported an awareness of varying school “atmospheres”. For example, distinct differences were reported in the manner in which schools displayed the pupils’ work, and in the level of interest and support for research into PA and academic performance from the principal, teachers and office staff. It is recalled that the SES of the schools has been both controlled in our design (with selection of schools), as well as adjusted for statistically, supporting our premise of a cultural influence arising from within the school staff rather than home or other local environmental influences.

Literacy and numeracy measures were more strongly related to CRF than to PA, both at the between-school and between-child levels (with and without adjusting for %BF and other potential confounders). This is consistent with the previous literature (24) where it was concluded that fitness, rather than PA, was the variable most likely to exert an impact on learning and executive cognitive function. On the other hand the authors of a meta-analysis (8) concluded that CRF was unlikely to be related to academic achievement, although most of the participants in the studies analyzed were of greater than school-age, which may have contributed to their different conclusion. Considering the findings of these two summary papers together with our current data, it may be that CRF is more closely related to brain function during the earlier stages of child development where plasticity associated with neuromuscular interaction is likely to be greater.

The absence of any significant relationships between literacy and numeracy scores and %BF at the between-child level the current report is consistent with three large studies, each of which involved statistical adjustment for potential confounders. Researchers making use of data from the Early Childhood Longitudinal Study in kindergarten and first grade children reported that relationships between academic achievement and overweight status represented by BMI disappeared following adjustment for PA and SES (5). A subsequent study using the same database (11), this time from the third grade, found that associations between overweight and mathematics and reading scores disappeared when they adjusted for SES and maternal education. A third large population study, using data from the National Health and Nutrition Examination Survey (13) reported an absence of evidence relating academic performance and BMI, following adjustment for SES. Of interest in the latter study was that separate tests of cognition were significantly and negatively associated with BMI. However, interpretation of this study is clouded due to the lack of experimental control or statistical adjustments for fitness; fitness also being negatively related to BMI in children (1) as well as being associated with cognition (24). Equally difficult to interpret is the review (26) which reports overall evidence of a negative correlation between intelligence (as measured by IQ) and obesity in children in the absence of any reference to adjustments or control for CRF or PA. On the other hand, in the current study, evidence of (negative)
relationships between numeracy and reading scores with %BF did emerge at the school level. This suggests that if indeed a school cultural influence does explain part of the fitness-academic score relationship, then it may apply to any relationship between body composition and academic achievement as well.

Strengths of our study were the design, including the numbers of participating children and schools, the rigor of the statistical analyses, and the objectivity of the measures of %BF, fitness and PA. On the other hand, the pedometer measures were a relative weakness, because they were unable to discriminate the intensity of PA, and do not measure activities such as swimming or cycling. However, while relationships between PA and academic performance may have been affected to some degree, there is no reason to suspect any impact on inferences or conclusions. Replication of this project would be of interest not just for the scientific verification of findings but to determine whether the effects we showed in Australian schools apply to other educational jurisdictions in and outside Australia.

In conclusion, we found that elementary schools with fitter children achieved better literacy and numeracy results; and to a lesser degree, better academic results also occurred in the more physically active and leaner children. These relationships at the school level were much stronger than between children within the schools, lending more weight to a school cultural explanation of the relationships than to a direct influence of fitness on literacy and numeracy achievement.

Acknowledgments

We thank the local government jurisdiction and Australian government for their cooperation with this project; the teachers, parents and children for their willing participation over the three years; Ms Karen Gravenmaker for carrying out the DXA measurements, together with her supervisory staff at The Canberra Hospital, Drs Ahmad Javaid and Marc Budge; Dr Robert Fitzgerald, University of Canberra for his review of the manuscript, and the Commonwealth Education Trust for their financial contribution to this study.

References


