Over ten years on from “The horse is dead...”: surely it must be time to “dismount”?!
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
Title
Over ten years on from “The horse is dead…”: surely it must be time to “dismount”?!

Authors
Dr Lorraine Cale*, Dr Jo Harris (Loughborough University, United Kingdom) and Ming-Hung Chen (National Pingtung University of Science and Technology, Taiwan)

*Correspondence/galleys to Lorraine Cale at the postal and e-mail addresses below

Mailing address:
Dr Lorraine Cale & Jo Harris
School of Sport and Exercise Sciences, Loughborough University, Leicestershire,
LE11 3TU, United Kingdom

E-mail addresses:
Lorraine Cale e-mail: l.a.cale@lboro.ac.uk; Jo Harris e-mail: j.p.harris@lboro.ac.uk

Fax:
01509 226301

Key words
Fitness; fitness testing; children; young people; schools, physical activity promotion
Over ten years on from “The horse is dead…”: surely it must be time to “dismount”?! 

Abstract

This paper represents a response over ten years on, to an editorial piece in this journal by Thomas Rowland in which he debated fitness testing and asked whether the “horse” of fitness testing in schools was dead. Here, we revisit the debate and consider the progress which has been made with regards to fitness testing in schools in recent years. On the basis of the findings from the literature and some of our research, we suggest that accepting the “horse” is dead would not be a bad thing. Our advice is certainly to pull tightly on the reigns, slow the horse down, and not allow fitness testing to dominate schools’ efforts to promote physical activity.
Introduction

Over ten years ago, we read with interest an editorial piece in this journal by Thomas Rowland (46), which debated fitness testing in children. On the back of an article entitled ‘Toward an Understanding of Appropriate Physical Activity Levels for Youth,’ in which the authors, Corbin, Pangrazi and Welk, proposed a shift from a fitness to a physical activity promotion, or the lifestyle physical activity model (LPAM) (20), Rowland questioned the impact of this model on the role of fitness testing of school children. His discussion centred on the role of testing in the promotion of physical activity, arguing that the model served as ‘the best argument for abandoning the routine field testing of children…’ (p. 119). In conclusion he asked ‘…is the “horse” of exercise testing in the schools dead?’ and if so, suggested ‘…perhaps we should “dismount”’(p. 119).

Rowland was certainly not alone in his concerns. In particular, his editorial provoked a response by Corbin, Pangrazi & Welk (20), in which they supported and elaborated on many of the points Rowland made. In addition, issues have been debated and concerns expressed over fitness testing children by numerous other researchers and organisations over the years (e.g., 5, 7, 8, 14, 16, 24, 26, 29, 30, 35, 43, 47, 51). More recently, Keating has claimed that unless youth fitness testing actually improves fitness and increases involvement in physical activity, the need for it questionable (35).

Despite the concerns and Rowland’s suggestion over ten years ago to perhaps “dismount” the fitness testing horse in schools, testing has and continues to be common place within the PE curriculum (6, 13, 17, 28), with most secondary school programs including compulsory testing (13, 17, 28, 45). Advocates of such testing claim that it promotes
healthy lifestyles and physical activity, motivates young people to maintain or enhance their physical fitness or physical activity, facilitates goal setting, self-monitoring and self-testing skills, promotes positive attitudes, and enhances cognitive and affective learning (41, 55).

Within this paper we revisit the debate and consider the progress which has (or has not) been made with regards to fitness testing in schools. In particular we consider whether the routine field testing which typically takes place within curriculum time has a role to play in the promotion of physical activity and is compatible with the LPAM. Our discussion endeavors to present a United Kingdom (UK) perspective and focuses on a few key issues and the related evidence presented in the literature. In addition, to further inform the debate and exemplify key points, we draw on selected findings of two research projects we have been involved in recently. The first project was a ‘monitoring study’ which involved a national survey of fitness (as well as physical activity and health) monitoring in over 300 secondary schools in England, plus follow up interviews with PE teachers from approximately 20 schools (17). The second was commissioned by the National Assembly for Wales and aimed to determine the feasibility of carrying out a research project on the fitness levels of Welsh children (31). This ‘feasibility study’ was based on a world wide review of literature plus a survey and interviews with 35 key stakeholders (including experts). On the basis of the above, we consider whether or not it is time to take the horse by the reins, accept that the “horse” is dead and “dismount” once and for all.

Methodological limitations

One argument in favour of “dismounting” the “horse” and which was alluded to by Rowland (46), relates to the methodological problems associated with fitness testing children.
Fitness tests are plagued by methodological limitations and the validity and reliability of some tests, test batteries and in applying norm and/or criterion referenced standards with children has been questioned (21, 26, 35, 44, 46, 47, 48). These limitations stem from the many factors that influence performance and are reflected in fitness test scores (22, 38) (e.g., the environment/test conditions, test protocol, motivation, skill, heredity/genetic potential, maturation).

Such limitations were also widely acknowledged by stakeholders in the feasibility study. Comments included:

‘…The extrapolation of field test data is fraught with problems’.

‘Even in a well-equipped sports science laboratory using ‘scientific’ tests, one might consider the error to be around 10%. In the field situation using simpler methods (e.g. sit-up tests or shuttle runs), the error is likely to be huge’.

Problems concerning the measurement of physical fitness in young people have led some to conclude that tests suitable for use in the school environment and which provide valid and objective measures of fitness are simply not available (8, 10, 43). Given such limited faith can be placed in the data, we argue that their use in promoting physical activity is also limited.

The relationships between physical fitness, physical activity and health

Fitness testing in schools may have assumed such popularity due to the growing concerns over young people’s physical fitness, physical activity and the implications of these for their
health. However, some of the concerns are ill founded. It is still often incorrectly assumed that children’s fitness is primarily a reflection of the amount of activity they do, and that those who score high on fitness tests are active and those who do not are inactive (38). This belief was evident amongst some of the teachers in the monitoring study, illustrated by the following remark:

‘Because the more active the children are, the better their fitness level is going to be’.

Whilst physical activity is an important variable in physical fitness development for adults (38), the relationship between fitness and activity among children is low (10). The reason probably lies in the low level of physical activity of most young people (10). Problems can arise if fitness test scores are linked to activity levels. An active child who scores poorly on a test may become demotivated and ‘turned off’ activity because he/she feels it does not ‘pay off’ (19), whilst an inactive child who scores well may be delighted with the outcome, conclude that everything is alright when it is not, and may not be motivated to change.

In addition, efforts to advocate fitness testing due to concerns over young people’s physical fitness and the implications of this for their health are misguided. According to Cale & Harris (16), much media attention and ‘hype’ has been afforded to young people’s fitness, with messages leading us to believe that today’s youth are unfit, unhealthy, and far less than fit than in previous decades. Yet, following a review of the related literature, they declare that ‘there is no evidence to suggest that low levels of aerobic fitness are common amongst young people’ and ‘no convincing evidence to suggest that young people’s aerobic fitness has declined over time’ (p. 32). On this issue, Corbin (19, p. 139) suggests
that the media ‘likes bad news’ and that ‘much talk about lack of fitness of our youth is hyperbole...’

The data from both the monitoring and feasibility studies also revealed some confusion in this respect. Teachers commented:

‘Young people aren’t as fit as they should be’.

‘It is obviously a government issue at the moment...the whole range of fitness and how unfit we are as a society. As PE educators we need to address that and try to make children aware of it’.

Further, the implications of children’s fitness and fitness test scores to their health are not well established. In 1995, Rowland highlighted how the link between performance fitness and health outcomes was weak and how testing was therefore ‘archaic and inconsistent with our current understanding of the exercise-health connection’ (46, p. 199). To date, the link between fitness in youth and health status is still not supported in the literature (26). There still remains only weak evidence that physical fitness is related to health in young people (53) and little or no direct evidence that physical fitness during childhood and adolescence is related to adult health (52, 53). Again, confusion was evident amongst some individuals in both studies with one stakeholder claiming that ‘in terms of the health of our nation, we really do need to look at the fitness of our youngsters very, very seriously...’

The above, plus evidence that a sizeable proportion of young people are inactive and lead sedentary lifestyles, suggests there is a need to focus attention on influencing young
people’s physical activity behaviour (16). In other words, to adopt the LPAM in which emphasis lies on the ‘process’ of physical activity rather than on the ‘product’ of fitness. From a physical activity promotion and LPAM perspective, a number of good reasons have been cited for focusing on physical activity rather than fitness (15, 19, 38, 46). For example, physical activity influences children’s health more so than fitness, promoting physical activity is more likely to be acceptable to the general public and particularly to those who are sedentary or have low fitness, the focus on fitness which was common practice for many years was largely unsuccessful, and increased physical activity is an outcome that can be achieved by all children.

The influence of fitness testing on young people

From the perspective of the LPAM, the influence of fitness tests on young people’s physical activity behaviour, attitudes and motivation towards, and their knowledge and understanding of physical activity and physical fitness would seem to be crucial. Yet to date, research has largely ignored the effects of youth fitness testing in schools (35). Keating claims that in the US three facts cast doubt on the role of fitness testing in promoting physical activity (and improving youth fitness): a) children have failed to show improvements in fitness and have become less physically active; b) the percentage of overweight youth has increased substantially in recent years; and c) the proportion of inactive adults has also increased dramatically (35).

The potential motivational role of fitness testing on young people was acknowledged by individuals in both the monitoring and feasibility studies. For example, a stakeholder felt that fitness testing:
'can be motivational if health-related, linked to physical activity, and used in the right way'.

whilst a teacher explained:

'I think on the whole it’s proved to be quite positive. Generally they like to look at their progress, and over the year they do three or four tests and then they can see if they improved. So they usually quite enjoy it on the whole',

Yet, despite anecdotal support for this notion, concern has been expressed that fitness testing may be repetitive and boring (35), de-motivating and counterproductive to the promotion of active lifestyles in young people (20, 22, 46). This point was made most strongly by Rowland (46) in his editorial who argued that fitness tests are antithetical to the goal of promoting physical activity in so far as they can be demeaning, embarrassing and uncomfortable for children (often those about which there is most concern), and may reinforce the notion that exercise is competitive and unpleasant. Keating (35) warns of problems relating to lack of privacy with testing and test results, whilst Corbin, Pangrazi & Welk (20) warn that testing that is done improperly may turn many youngsters ‘off’ rather than ‘on’ to activity.

The potential negative consequences of fitness testing were also commonly reported by teachers and stakeholders. Teachers expressed concerns over the negative impact of testing on lower ability and overweight children in particular, whilst stakeholders warned how tests can de-motivate, be perceived as threatening, cause discomfort, stress, label and embarrass children, make them look ‘daft’ and ‘turn’/‘switch’ them off. Comments from a teacher and stakeholder respectively included:
‘...if they learn from it in a positive way, it’s good. If they learn that they are unfit, overweight, can’t do PE, then it’s a very negative way. It has to be used with considerable caution...’.

‘PE teachers who use fitness tests regularly on any and every group of children should be encouraged to stop! Children often hate and dread them, they don’t tell us very much, and why should children be forced to endure them?’

Despite the above, little attention has been paid to the motivational effects of fitness testing (25, 34) or fitness test awards on young people (35) and the need for such research has long been recognized (40). Studies that have been conducted (1, 27, 36, 54) have revealed variable results and it has been concluded that motivational enhancement from testing cannot be taken for granted (27) and that there is no empirical data to indicate that students value fitness test awards (35). Attitudes towards fitness tests have been found to be unfavourable (36) and some youngsters (and teachers) have been accused of not taking testing seriously (35). The motivation of young people towards testing has been found to be influenced by feedback following tests, perceived competence (54) or perceived success, and achievement goal orientation and performance in the tests (27). The variable influence of fitness testing on some youngsters was recognized in the monitoring study with one teacher making the following observation:

‘The top half of the group, they’re much more into the testing. I think it does influence some of them...’.
Pate (40) believes that the most important reason for administering fitness tests with children is to facilitate learning in the cognitive (as well as the affective) domains in terms of equipping them with the knowledge and understanding required for participation in a physically active lifestyle. Regrettably however, he and others suggest that these aspects are also usually given the least attention and often overshadowed in fitness test programs (35).

PE teachers and stakeholders similarly acknowledged the value of fitness testing in enhancing learning with responses including ‘to aid learning’, and ‘to inform pupils about their own relative fitness levels’. One teacher remarked that through fitness testing students were able to:

‘…learn a lot about the effects of exercise on the body, for example recording their heart beat and recovery rate and pulse testing’.

However, even less research attention has been paid to the educational role of fitness testing. Only one study can be cited which investigated what children ‘thought, felt and knew about’ the mile run test and which revealed they generally showed little or no understanding of why they were being asked to complete the test (33).

So, is the horse dead and is it time to dismount?

From the preceding discussion, it seems that fitness testing in schools may not be compatible with the promotion of active lifestyles and the LPAM. There is little evidence that fitness tests provide meaningful data, positively influence children’s physical fitness,
physical activity, motivation or knowledge and understanding, or that meaningful relationships exist between physical fitness and physical activity or physical fitness and health. Keating agrees claiming that unfortunately little empirical evidence is available to show that fitness programs really work for children and youth in schools (35). To the contrary, fitness testing can be counterproductive in that it can be unpleasant, embarrassing and meaningless for many young people, and scores can be inaccurate, misleading, unfair and demotivating (16). In this respect, we agree with the sentiments of Rowland (46) that fitness testing in schools should be carefully re-examined. What is disappointing though is that, over ten years on, and certainly in the UK, many of us seem none the wiser or clearer on the issue.

This said however, if appropriately employed, there is no reason why fitness testing could not play a role in supporting active lifestyles and the LPAM. In 1995, Corbin, Pangrazi & Welk (20) were likewise of the opinion that fitness testing might survive if it could be shown to promote the right philosophy, suggesting that ‘there may be some life left in the horse’ (p. 350) if certain considerations are made before testing. They asked, ‘is it the testing itself that is ‘bad’ or the way in which it is done?’ (20, p. 348). An interviewee in the feasibility study held similar views commenting:

‘I’m not against knowing where we are with regard to our children’s fitness but I am against how it could possibly be done, and I’d hate it to be mis-used along the way’.

Concerns have been expressed over the way in which fitness tests are often implemented in schools with too often tests been an almost irrelevant adjunct to the curriculum or else dominating or constituting the entire fitness education program (41). Indeed, the amount of
time schools spend on fitness testing without necessarily positively influencing young people’s activity levels or their attitudes has been criticized (15, 30, 35), and it has been suggested that PE time could be used more wisely (16).

Pate (41) noted how, despite its popularity over a number of years, there is little scientific evidence to guide us in deciding how best to incorporate fitness testing into PE. Over ten years on, little has changed, at least in the UK. Recommendations concerning the implementation of fitness testing with young people have been made by a number of researchers and organizations (e.g., 2, 3, 4, 5, 16, 20, 29, 32, 41, 42), but these have been based more on common sense than on scientific evidence. We believe that if fitness tests are to have a role in schools, teachers need clear guidance, support and training in the implementation of fitness testing and in particular in how to use fitness testing to achieve cognitive, affective and behavioural objectives. This view was also supported by the teachers and stakeholders, with several highlighting the need for appropriate guidance and support materials to assist teachers, including programs to work from. One teacher commented:

‘…I think we need more formal guidance some sort of national program…with a common structure across all schools…’.

One stakeholder in the feasibility study reported:

‘there is a need to convert the extensive detailed scientific knowledge about fitness testing into appropriate educational tools so that teachers (and coaches) can be provided with material that allows them to offer children a contemporary understanding of the facts and issues’. 
On this, there have been very positive developments in the United States with the production of fitness resources for teachers such as ‘Physical Best’ (2, 3, 4) and ‘FITNESSGRAM/ACTIVITYGRAM (18; www.fitnessgram.net). The resources represent comprehensive fitness education programs which recognize the importance of physical activity and the LPAM, as well as fitness, by seeking to develop the affective, cognitive and behavioural components associated with physical activity participation. The latest version of FITNESSGRAM/ACTIVITYGRAM (8.0) includes fitness and activity assessments and reporting programs, and the accompanying reference guide provides guidance on the appropriate and inappropriate use of the resource. Teachers in the UK require and would welcome an equivalent resource or resources.

In summary, and from the viewpoint of promoting physical activity and the available evidence, we feel that accepting the fitness testing “horse” is dead and “dismounting” once and for all would not be a bad thing and that the time often devoted to testing could be better spent. That said, research in this area is still relatively scarce. More studies on the effects of fitness testing in schools are therefore needed to guide the nature and implementation of future testing. However, until such a time and until a stronger evidence base and more definitive guidance becomes available, our advice is certainly to pull tightly on the reins, slow the horse down, and not allow fitness testing to be the front runner and dominate schools’ efforts to promote physical activity. Instead, we would welcome more time and energy being afforded to teaching young people the knowledge, understanding, physical and behavioural skills and fostering the attitude and confidence associated with current and lifelong participation in physical activity, via actively involving them in a range of purposeful activities.
References


Rowland’s call to abandon the routine field testing of physical fitness among youth in schools was made more than a decade ago (4). Rowland’s paper highlighted the many potential abuses of fitness testing, the lack of reliability and validity of many youth fitness tests, and the lack of evidence supporting links between health and fitness among youth. Because Rowland advocated an emphasis on a lifetime physical activity model (LPAM), a model that I had jointly proposed with colleagues Bob Pangrazi and Greg Welk, my colleagues and I were asked to respond to Rowland’s editorial. In that paper we supported many of the points made by Rowland. We particularly agreed that those who use fitness tests should focus on promoting lifelong physical activity (1). At that time we suggested that, done properly, fitness tests could be of value in promoting lifelong physical activity and resulting adult fitness and health. The association among fitness, activity and health is much stronger among adults than among youth.
Now Cale, Harris and Chen have again addressed the limitations of youth fitness testing as a component of physical education suggesting that “…accepting that the fitness testing ‘horse’ is dead and ‘dismounting’ would not be a bad thing and that the time often devoted to testing could be better spent.” Ultimately, however, Cale Harris, and Chen do not call for an end to youth fitness testing in school physical education, but rather urge that we “,,,pull tightly on the reins , slow the horse down, and not allow fitness testing to be the front runner and dominate school’s efforts to promote physical activity.”

While I agree with the ideas expressed by Cale, Harris and Chen, just as I agreed with Rowland’s ideas in 1995, I now would like to suggest that we accept the fact that the “horse is alive, and ride it to our best advantage.” Physical fitness is but one of many important physical education objectives, and should not dominate schools efforts. However, if I have learned anything in my 40+ years of research and teaching, it is that fitness testing will live on regardless of what I think. Fitness testing is a tool that can be used for good or for bad, depending on how it is done. Our challenge is to make sure that it is used in the best possible way. So I urge, as Cale, Harris and Chen do, that we do everything in our power to make sure that fitness testing is used to meet the needs of youth and to promote lifelong physical activity. It is my view that informed teachers have become less likely to use tests inappropriately especially those who use the HELP philosophy and guidelines for appropriate and inappropriate practice developed by FITNESSGRAM® (2,5). I also believe that more and more informed teachers are teaching students to do self-testing and to use fitness tests as a basis for program planning (appropriate uses) rather than as primary methods for grading or determining program success (inappropriate uses). Finally, I believe we have been able to prevent the use of fitness tests as “high stakes physical education tests”, the ultimate misuse of fitness testing in my opinion. Still,
there are efforts by the uninformed to continue to press for the use of fitness tests as statewide achievement tests. These efforts must be rebuffed.

I urge professional organizations and governmental agencies to develop fitness testing policy for use by teachers that is based on a sound philosophical base, that has established guidelines for appropriate and inappropriate uses, that has scientific evidence to support test standards and test items, and that provides scientific support for the use of fitness awards, if they are to be given.

Further, I urge professional organizations and governmental agencies to provide in-service education to teachers to make them aware of the best information and to monitor testing programs to see that “best practices” are being used. Unfortunately, only those teachers who attend professional meetings and read professional journals and websites are likely to be informed. So, special efforts are necessary to inform the uninformed.

I also urge professional, corporate, and governmental organizations to work together so that research and educational efforts will be successful. The fractured approach of governmental, professional and commercial groups working toward separate competitive ends, must be replaced with a unified effort that benefits all youth. Also, as Cale, Harris and Chen point out, much more research is necessary if we are to answer important questions and use fitness tests effectively.

While modest in nature, the development of a free Reference Guide available to all users of FITNESSGRAM® (5) and the publication of a special supplement of the Journal of Physical Activity and Health (3) designed to enhance our knowledge of fitness assessments (5) are steps in the right direction. The special JPAH edition also contains an updated article on appropriate fitness test uses (2).

The horse (fitness testing) is not dead. The horse lives, so let’s do our best to use the horse to benefit youth and to meet important physical education objectives.
Ken Fox

University of Bristol

This paper takes me back to the heated and rather polemic debates that seemed to appear annually at the AAHPERD conference throughout the mid 80s. The arguments were fired by strong personalities taking positions on the value of fitness testing batteries and award schemes as standard fare for American school children. The Presidents Council award scheme was
usually the target of attention. At the same time, the question and merits of fitness testing in schools was being addressed rather more quietly in the United Kingdom through its Physical Education journals and conferences. I was involved in these moves on both sides of the Atlantic and at the time would never have guessed that 20 years on we would be facing similar conversations.

For me the issues and solutions have not changed. Physical fitness testing in schools should be judged on its educational merits, something several of us tried to articulate fairly early on in the context of the emergence of health-related physical activity as curriculum content in Physical Education (2-7). In the absence of specific research evidence on fitness testing at the time, these ideas were inspired by our somewhat limited knowledge of social psychological theories of motivation, persuasive arguments by pioneers in health–related exercise such as Chuck Corbin, Russ Pate and Bob Pangrazi, and our own hands-on experience as teachers.

Our thesis was quite simple. Similar to sport, fitness testing should be seen as an educationally neutral event. The question is whether it can be used wisely to improve the education and motivation of our children to be active. We were convinced that the fitness awards/standards approach that was almost always delivered en masse in a very public manner to children worked to disenchant a large sector of children (8). In essence it encouraged those already athletic, sporty and active and discouraged the ‘health needy’ therefore failing in its primary objective of promoting activity for health among our children. In contrast we argued that there was a place for learning about fitness tests as part of a sequential education process, a concept that had already been well developed by Corbin (1).

I feel this is still the appropriate stance and what little research that has emerged on competence, intrinsic and achievement motivation supports it. On the one hand, fitness award schemes are likely to discourage those who need help most and for me they sit very
uncomfortably in a curriculum designed to promote lasting motivation in children to be active. On the other hand, fitness tests do provide information, they are publicly used in many fitness and health club settings, and our youngsters are likely to be exposed to them at some point, particularly if they choose to join a formal sport or fitness/health club. Furthermore, they can be used to show personal improvement. Surely if the role of physical education is to ‘physically educate’, then we must teach our youngsters how to test for different aspects of fitness, how to critically appraise their reliability and validity, and how to interpret the information they provide with regard to their own fitness and activity levels. This should be achieved in the context of a programme that educates youngsters about the importance of physical activity and the different elements of fitness to their health and well-being.

Fitness test batteries and award schemes are still being presented in schools on both sides of the Atlantic, and this seems to expose an overarching problem. Given the commentaries that have emerged over the past two decades, the delivery system of physical education must be inherently flawed if it perpetuates their use. We therefore need to investigate where the problem with the system lies, and in my view it rests with the quality and orientation of our physical educators. In the United Kingdom, physical education teacher preparation has deteriorated to the level of a one year apprenticeship in schools. There are no longer any mechanisms for innovation and change. As I understand it, physical education also remains at very low status in the USA. Well educated and motivated teachers can adapt mediocre ideas to make them work well. Poorly educated teachers can turn great ideas into disasters. Yes, we do need more research to inform us on how to help coaching and teaching produce long term benefits in the activity, health and well-being of our youngsters. However, although researchers probably hate to admit it, most of what happens at the coal face in teaching is unaffected by research and results from policy and personal aspirations and motives of teachers. Until we can convince
politicians and policy makers that we need better educated and qualified physical educators in schools and give some modicum of status to what they are trying to achieve then we will continue to support the kinds of uninformed practice that we see with the use of fitness testing.

References


*       *        *       *       *       *       *

Jim Morrow

University of North Texas

I must confess that much of this fitness test debate is “Greek to me.” Clearly, it seems that the fitness testing Hydra is once again raising its multiple heads. As one who has taught
“tests and measurements” for more than 30 years and been involved in fitness testing activities with the American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD), The Cooper Institute’s FITNESSGRAM/ACTIVITYGRAM (1), and the President’s Council on Physical Fitness and Sports President’s Challenge (8), it is always interesting to see that this monster has not been slain but continues to battle. Frankly, it appears that the “monster” only sleeps briefly and then rises to once again wreak havoc, discontent, and mayhem across the globe.

Fitness testing is not without controversy and change over many years (5, 6, 10, 11). Assessment of youth fitness levels in the United States was made over 20 years ago through the two National Children and Youth Fitness Surveys (3, 4) and the National School Population Fitness Survey (7). Since that time, there has been much hyperbole about the status of youth physical fitness and it has been suggested that the current status is largely unknown (2). The National Health and Nutrition Examination Survey includes measurement of fitness parameters on children and youth. Is it important to know youth fitness levels? Has status changed? What should be measured? How accurately can these measures be taken and how generalizable are these estimates? Is it physical activity or physical fitness level which is of most interest/importance? The answers to these questions are not simple. Clearly, youth fitness assessment on a large scale is a Herculean matter.

With this as a background, I suggest that we once again take a good look at the youth fitness testing horse (9) to determine if it can enter the battle. Some considerations include:

1) Making evidenced-based decisions with evidence.

2) The purpose of fitness testing is NOT to improve fitness – it is to assess the level of fitness.
3) Is it health-related fitness components or performance-related components that should be measured? If health-related, why is it that so many fitness tests around the globe continue to include performance or sport-related items on their fitness test batteries?

4) Testing does not necessarily promote healthy lifestyles. So what?

5) Failure to assess fitness levels of individuals suggests to them that fitness is NOT important.

6) Is the important issue one of individual monitoring of fitness status or the monitoring of secular changes across populations with surveillance methodologies that stand the test of time?

7) Should physical activity and/or physical fitness be assessed? Is it process or product? Is it both?

8) What are the best aspects of fitness to assess?

9) There is national interest in education testing (which might be good and bad). Why test only in the cognitive domain and ignore the important psychomotor/health-related domain?

10) The argument that students will be “turned off” to fitness testing if they do poorly is a tale suggestive of Sinon, the great liar. Is this true of math, or science, or English, or any other content area?

11) Fitness testing need not be “demeaning, embarrassing, and uncomfortable” although learning about ourselves is sometimes a bit of all of these.

12) Certainly scores can be misleading and cause confusion. This is a result in ALL testing. False positives and false negatives have consequences. Fitness test results, like all test results, are ESTIMATES. They should be interpreted and used that way.
13) Fitness testing can inform one much about oneself and inform large populations about secular changes. Is this bad?

Assessing youth fitness status is a Herculean task. However, assessment of youth fitness status is an important concept. Let’s continue the fight; ever learning the importance, the status, and role of fitness testing.

References

The current paper repeats the call for eliminating physical fitness testing in the schools espoused by Rowland in this journal in 1995. In the process it supports the concept of physical activity (PA) and the lifetime physical activity model (LPAM), but does not suggest monitoring PA in place of physical fitness (PF). The implication is clearly that a choice must be made between PA and PF and that PA is a more than ample surrogate for PF without the baggage of PF testing.

A recent data based review (8,9), however, leads to the conclusion that PF and PA may be independent risk factors for health outcomes and that both need to be monitored. Both PA and PF lead to health benefits and decreased negative health outcomes in adults and children/adolescents (C/A). It is true that the health benefits are more indirect for C/A. Morbidity/mortality in C/A are primarily the result of diseases not impacted by lifestyle (e.g. cancer, congenital problems). However, atherosclerosis, hypertension and obesity are examples of health problems that originate in C/A whose medical complication are not manifested until adulthood. Therefore, the realistic endpoints for epidemiological studies in these age groups are risk factors not health per se. A 2002 summary of five longitudinal observational studies (11) comparing the relationships between PA and/or cardiovascular PF during youth and cardiovascular risk factors later in life showed that irrespective of the methodology of the study, PF was predictive of a healthy cardiovascular risk profile at the later age, but PA was not. The authors speculated that the predictive value of PF might be caused by its relationship to percent body fat (%BF) and/or a healthy lipid profile. Given the obesity epidemic (6) indirect or direct knowledge of body composition is important, and, as the Arkansas process has shown, can reap
definite benefits. Many AK residents responded negatively when told that student BMI reports would be sent home, but 57% of state physicians have had at least one parent bring that report to them for help (1).

When risk reduction (in terms of mortality or risk factors) from PF or improvement in PF is compared to risk reduction from PA or improvement in PA the data suggest that fitness is more important (8,9). This conclusion is often accompanied by the caveat that the measurement of PA is less accurate than PF—leading to more misclassification and weaker association with health outcomes. Well accepted criterion measures and valid field tests are available for cardiovascular and body composition measures; neuromuscular components are not so clear cut (5,12). Conversely, there is no single criterion measure for PA (10). Direct measures (accelerometry, pedometers, observations, etc) involve time, training, money and equipment. Indirect measures (primarily questionnaires) although frequently reliable, are only sometimes objective and rarely truly valid. Thus, measurement problems are not solved if a switch is made from PF to PA.

Data at this point in time do indicate that there is little evidence PA directly influences PF in young children. However, youths classified as fit have tended to be more active. (8,9). Data on the impact of PF results on motivation for PA are sparse. One study (4) investigated the impact of scores from the AAHPERD Youth Fitness Test obtained when the children were either 10-11y or 15-18y on activity as young adults. Physically active adults had better standardized test scores as children on four of the six tests (50yd dash, sit-ups, shuttle run, and 600yd run). The risk of adult inactivity increased with the number of low fitness scores as a child. Individuals who scored in the bottom quintile (0-20%ile) were almost twice as likely to be physically inactive as young adults as those who scored in the highest quintile (80-100%ile).

The magnitude of the problem of individuals being misclassified by PF tests in relation to their activity level is unknown, but probably small. A study by Martino et al. (7) directly measured \( O_{2\text{max}} \) in 1900 healthy young adults. Six participants (\( \leq 0.3\% \) of this sample) were identified as having very high (\( \geq 62.5\text{ mL.kg.min}^{-1} \)) values with absolutely no history of training. Conversely, evidence is abundant (8,9) that individuals respond differently to the same activity
or exercise program and some who are consistently active will score poorly on fitness tests—or for that matter on risk factors such as HDL levels (2). It now appears that genetics impacts not just the results of fitness tests, but also at least some risk factor levels and the inclination to be active or inactive. It therefore seems important to monitor both PA and PF in schools so that each individual student can be guided in interpreting his/her propensity for activity as well as his/her individual pattern of adaptation to various types/levels of activity for each of the major fitness components.

How students are prepared for fitness testing, how fitness tests are administered, how test results are communicated and utilized are all critical factors. Fitness testing needs to be part of a sound physical education program. What must to be eliminated are bad pedagogical practices. The authors acknowledge the “need for appropriate guidance and support materials to assist teachers…” and speak positively about FITNESSGRAM/ACTIVITYGRAM and Personal Best (PB)(3). This advisory board member would like nothing more than for the philosophy of FG/AG/PB to be adopted worldwide. Surely distribution rights could be negotiated for these programs.

Finally, in terms of the horse analogy, horses are herd animals and they directly reflect the training and skill of their riders. Ridden alone the horse called fitness is not entirely comfortable. Rather than dismount, re-tool and give him company on the school trail ride with horses called activity, nutrition, medical screening, knowledge, skill, fun, support, etc. Everyone will benefit.

References
1. Associated Press. Obesity report cards put brakes on bad habits: Benefits seen two years after Arkansas schools began weighing students. 
3. Cooper Institute. FITNESSGRAM®/ACTIVITYGRAM®