

Evaluating Sport and Physical Activity Interventions

A guide for practitioners

Dr Lindsey Dugdill & Prof. Gareth Stratton



the design & print group, university of salford T: 0161 295 2630 21023/07

Dr Lindsey Dugdill
Reader in Exercise & Health
Centre for Public Health Research

The University of Salford
Salford, Greater Manchester
M5 4WT, United Kingdom

T: +44 (0)161 295 2365
E: L.Dugdill@salford.ac.uk

Prof. Gareth Stratton
Professor in Paediatric Exercise Sciences
Research Institute for Sport
and Exercise Sciences

Liverpool John Moores University
Liverpool, Merseyside
L3 2ET United Kingdom

T: +44 (0)151 231 4334
E: G.Stratton@ljmu.ac.uk





Foreword

The guidance document on “Evaluating Sport and Physical Activity Interventions” is aimed at physical activity practitioners, managers and commissioners from Local Authorities, Primary Care Trusts and Sport and Physical Activity Alliances.

We hope you will find it a valuable and practical resource to support you when monitoring and evaluating your physical activity and sport interventions and programmes more effectively.

The main difficulty facing people working in the physical activity field is the complexity of evaluating a health intervention, such as physical activity, for which there is no single validated tool to measure all types of physical activity, in all settings, for all ages. However, there are a range of different quantitative and qualitative tools which

can be used to look at different elements of physical activity under different circumstances.

This guide is not mean to be prescriptive about evaluation, but it should be seen as an aid when making decisions about the evaluation of physical activity and sport interventions.

We hope you will find it useful and we welcome your feedback.

Jackie Brennan
Regional Health and Physical Activity
Coordinator North West Public Health
Group

Shileen Tarpey
Sport, Physical Activity and Health
Manager Sport England



Introduction

The benefits of a physically active lifestyle in health promotion and disease prevention are well-documented (DH, 2004a). Despite such convincing evidence, sport and physical activity levels among the UK population are low (DH, 2000; Active People Survey, 2007) and there is a need to increase and sustain physical activity levels in order to reduce preventable deaths from chronic diseases and promote good health (DH, 2003).

The Chief Medical Officer's (DH, 2004b) recommendations for public health are that adults participate in “a total of at least 30 minutes a day of at least moderate intensity physical activity on five or more days a week.” (piii). For obesity prevention it is recommended that adults participate in 45-60 minutes of at least moderate physical activity each day. Children are recommended to accumulate at least 60 minutes of daily, moderate to vigorous physical activity. The Department of Health has a joint public service agreement with the Treasury, the Department for Education and Skills and the Department for Culture, Media and Sport (DCMS, 2002) to halt the year-on-year rise in obesity among children

under 11 by 2010, in the context of a broader strategy to tackle obesity in the population as a whole. Sport England is contributing to the CMO's physical activity recommendations by enabling adults to do 3 x 30 minutes of sport and active recreation per week (with a 1% year on year increase until 2010). Progression towards these targets are being measured through the Active People Survey (Sport England 2005-2006). School children should be receiving 2 hours of physical education and school sport per week by 2010. The Active People Survey showed that only a fifth of adults achieved 30 minutes of moderate intensity sport or active recreation on at least three days a week and data on children are unclear and patchy.

Aims of the Guide

The World Health Organisation recommend that between 10-20% of total intervention costs should be spent on evaluation.

However, evaluation is sometimes seen as problematic and time consuming and takes lower priority compared to intervention delivery (Stratton et al., 2005). In general, evidence on the effectiveness of physical activity interventions across England is scarce and interventions with sound evaluations are required (NICE, 2006). (for NICE criteria see:- www.nice.org.uk/).

In response, this guidance aims to address key issues that need to be considered when evaluating the impact of interventions on sport and physical activity. It is aimed at practitioners who are working with “Choosing Activity: a Physical Activity Action Plan” (DH, 2005) and who are attempting to evaluate strategies (Dugdill, 2001) developed within local area agreements, wider physical activity interventions or within health care or education settings. The guide discusses principles of good practice when designing evaluations and suggests tools that might be used. To assist practitioners a glossary of terms has also been produced (see appendix 1).

Evaluating Interventions

Evaluation is made up of a number of progressive steps, the most important of which is the collection of appropriate data that is subsequently used to make a judgment about the value of an intervention. The evidence from an evaluation should enable practitioners to produce more effective interventions for participants. An evaluation should measure progress towards meeting the expressed aims and objectives of an intervention. Also, as the resources and skills available for evaluations are often limited, a choice of what to measure and why should be taken early on in the evaluation process. Thus if an intervention aims to increase the number of people walking in a park, then counting the number of walkers around the park using an observational tool and a representative sampling period (morning, midday, afternoon and evening on week and weekend days) would suffice.

Evaluation should always be seen as an integral part of the process of intervention design – especially new interventions, however, it is important to choose methods and type of data carefully. An appropriate evaluation should instigate improvements in intervention effectiveness or aid intervention sustainability. Furthermore it is also useful to create a framework around evaluations as more than one evaluation may be required by an organization (local authority, PCT etc.) at any one time.

An Example of an Evaluation Framework:- RE-AIM

The RE-AIM framework (for more detail see: www.re-aim.org) is perhaps a useful way for practitioners to think about structuring their evaluation.

- Reach - Who did the intervention reach? – e.g. monitoring of participant numbers through registers, post codes, questionnaires, facility usage
- Effectiveness - How effective was the intervention at meeting its aims and objectives? – e.g. physical activity increase, decrease in body mass index or increasing the contemplation to become physically active.
- Adoption – Have significant parts of the intervention been adopted elsewhere? -e.g. exercise referral interventions have been adopted by most local authorities in the UK
- Implementation - How was the intervention implemented and managed? How was the intervention funded. What skills did the staff have?
- Monitoring: What were the monitoring and evaluation strategies used to assess the quality of the intervention? Is the intervention sustainable?

This framework also allows a reporting structure to be used between partnerships. Thus if 10 sport / physical activity interventions were in operation in one neighbourhood, each would report against the 5 RE-AIM sections allowing an overall report to be constructed using the RE-AIM headings.





Stages of Evaluation

Evaluation should measure all of the phases. Outcome evaluation alone is not sufficient as it does not explore the reasons why an intervention has been successful. We must ask what, how and why a sport/physical activity intervention is successful if the evaluation is to be meaningful.

In addition to the stages illustrated in figure 1 well-designed evaluation studies should follow a series of principles of good practice outlined below. Monitoring and measuring should receive as much thought as designing the physical activity intervention. The best way to consider this is to understand that each time there is an action there is evidence to collect. This evidence can be collected, collated and disseminated at various stages:

Evaluation can be described in a series of stages and these are illustrated in figure

- 01 Needs analysis stage (formative evaluation) of the planning cycle informs the strategic development of physical activity programme action planning.
- 02 Process evaluation gives evidence as to the effectiveness of the planning and implementation phases of the intervention.
- 03 Impact evaluation describe the immediate benefits at the end of a health intervention.
- 04 Outcome evaluation describes longer term effects and attempts to examine the sustainability of the intervention.
- 05 Dissemination (reports following RE-AIM framework).

STAGE ONE Planning

■ Stakeholder involvement and participation - evaluation should be carried out by a team who represent all stakeholder views. A balance between insider (management) and outsider involvement (consultants, academics) in evaluation is the ideal and can help keep the balance between subjectivity/objectivity and rigorous practice.

■ Needs analysis – helps prioritise issues in conjunction with stakeholders:

- ✓ it begins the evaluation cycle and provides baseline data
- ✓ it can establish the organisation or intervention's view and philosophy
- ✓ it identifies related problems/issues that exist and why
- ✓ it identifies current good practice but also gaps in provision
- ✓ it determines which issues are priorities for different sectors of the population
- ✓ it helps to establish how and when to intervene
- ✓ it sets up links with different groups and encourages participation

■ Clarifying aims of the intervention – stakeholders can use the needs analysis to decide on priorities for action, intervention aims and goals and implementation delivery.

■ Setting evaluation objectives – this involves several stages:

- i. Deciding the evaluation questions – once the aims and objectives of the evaluation are decided it should be reasonably straightforward to clarify the questions that the evaluation will address. Clarity at this stage can ensure that only relevant, rather than redundant information is collected.
- ii. Choosing appropriate indicators - indicators chosen to reflect stakeholder agendas will produce data of real interest and this is more likely to lead to action as a result.
- iii. A balance of process, impact and outcome indicators are required to avoid a focus on short-term impacts rather than longer-term outcomes. To be effective, evaluation should have a balanced range of all indicators. For instance, failure of an intervention may be due to the manner of implementation rather than the content of the intervention – only process indicators will detect this. Short-term impact measures are important to give participants and stakeholders a view as to the progress being made and can keep people engaged with the process. Outcome measures help to identify if an intervention is still on track at a later time and whether change has been maintained or lost. The boundaries between process, impact and outcome indicators can become blurred particularly with ongoing interventions due as the cyclical nature of action, evaluation and reflection can overlap. This is where structuring a number of intervention evaluations according to an evaluation framework (for example RE-AIM) can give direction to multiple interventions aimed at increasing activity levels.

■ Other factors influencing evaluation design

- ✓ Target group e.g. age, gender, disability
- ✓ Seasonality, climate
- ✓ Geographical location of intervention e.g. rural, urban
- ✓ Resources, skills for evaluation
- ✓ Current policy supporting health practices
- ✓ Environmental, cultural and socio-demographic issues.

Table 1 Strengths and weaknesses of physical activity measures. Direct (objective) and indirect (subjective) measurement – table of tools

Physical Activity Measure	Instruments available	Web site	Strengths	Limitations
Self report	IPAQ GPAQ GPPAQ ZDPAR Active Peoples Survey PAQ-C/A (children/adolescents)	www.ipaq.se www.who.int/cnp/steps/GPAQ http://www.dh.gov.uk/assetRoot/04/14/08/80/04140880.pdf www.rohan.sdsu.edu/faculty/sallis/measures www.activepeoplesurvey.com N/A	Captures qualitative and quantitative information Inexpensive. Low participant burden. Possible to estimate energy expenditure from compendium of activities.	Reliability and validity problems. Misinterpretation of questions-language problems. Problem for use in children <10 years. Potential recall bias
Accelerometer	MTI Actigraph RT3 triaxial	www.thetractigraph.com www.casa.ucl.ac.uk/capable/project/RT3.htm	Objective indicator of body movement (acceleration). Provides a measure of frequency, intensity, duration. Non-invasive Measuring frequency in seconds or minutes/ Large storage capacity (weeks) Easy data manipulation and analysis. Useful for analyzing activity in structured sessions such as exercise classes/walking activities.	Expensive Less useful for detecting upper body movement, incline walking and cannot be used in water. Problems with the placement of the monitor during extended monitoring periods.
Heart rate monitoring	Polar team Polar	www.polar.fi/polar/channels/uk	Indirect physiologic measure of activity. Provides a measure of frequency, intensity, duration. Non-invasive Measuring frequency in seconds or minutes/ Large storage capacity (weeks) Easy data manipulation and analysis. Useful for analyzing activity in structured sessions such as exercise classes/walking activities.	Expensive Occasional monitor discomfort. Heart rate affected by arousal, gender, fitness and temperature.
Pedometers	Yamax Digiwalker	www.polygondirect.com	Inexpensive, non-invasive. Can be administered to large groups. Useful in a variety of settings. Useful for goal setting and promoting behaviour change. Good measure walking type programmes	Problems with between instrument variation. Can be tampered with. Lose accuracy during running activity or during intermittent activity such as an exercise class or PE lesson.
Systematic Direct Observation	SOPAR SOPARK SOFT	www.rohan.sdsu.edu/faculty/sallis/measures	Can provide qualitative and quantitative information concurrently. Specific physical activity behaviours coded.	Time to train observers. Time intensive data collection. Observer presence may alter behaviour
Geographic positioning systems (GPS)	Garmin foretrak 201	www.garmin.com	Detects movement, speed of movement and distance traveled whilst outdoors. Download data and map activity patterns to geographical areas.	Expensive. Does not work indoors. Can provide erroneous results (speed and distance whilst in a car). Some participant burdening. Signals can be poor in some built up areas or under cloud or tree cover.

** Indirect calorimetry and doubly labeled water methods are not available to practitioners and will not be discussed here (See Welk 2002 for further detail).

STAGE TWO

Measurement

It is important to choose a measurement method appropriate to the evaluation question and type of intervention. A variety of research methods could be employed to gather “layers” of data on a topic which can be cross referenced to each other to add validity and rigour to the evaluation. Measuring process, impact and outcome will normally entail the collection of both quantitative and qualitative data.

Gaining Ethical Approval

There are a number of key principles that describe ethical protection when measuring physical activity (Stratton, 2006). The most important aspect of ethical evaluation is voluntary and informed consent by participants aged 16 years and over. For participants under the age of 16 (or vulnerable adults) assent from both the participant and parent/carer is required. Participants must receive an information sheet outlining the type(s) of research methods to be used, the time required, alterations to normal routine and potential recompense for participation.

All participants have the right to withdraw from the evaluation aspect of the intervention without penalty and without reason. Data should be treated in accordance with the data protection act (www.dh.gov.uk/en/PolicyAndGuidance/OrganisationPolicy/RecordsManagement/DH_4000489). Guidance on NHS research ethics can be found on the National Research Ethics Service (NRES; www.nres.npsa.nhs.uk). (N.B. An advanced Criminal Records Bureau (CRB) check is required to work with children and vulnerable adults).

Requirements for Ethical Approval

Application for approval of an Investigation for Research and Evaluation involving Human Subjects

Here the project title is stated along with its aims. The details of the project are outlined along with procedures including a statement about the originality of the project and how it will benefit participants or society. Statistical or content analyses should also be included and related to clear hypotheses or research aims.

Confidentiality of Participant Records

This requires a signed statement by the research team stating that they will adhere to the principles outlined in the Data Protection Act

Form of Consent to take part as a Subject in a major Procedure or Research Project

A form of consent for the participant that should also be signed by a disinterested third party/witness. This should include a short paragraph about the aims of the project.

Forms of Consent (Carer)

A form of consent for a vulnerable participant such as a child. This should include a short paragraph about the aims of the project written at an appropriate level.

Participant Information (PI) Sheet

The PI sheet should state “the exact nature” of involvement required of the participant.

Risk Assessment Form

The RA form needs to clearly establish the nature of the hazard, likelihood of injury occurring and procedures taken to minimise the risk. This may need to be signed by a relevant health and safety officer.

NHS Research Governance Proforma (when NHS staff or patients are involved)

The Local Ethics Research Committee (LREC) will also require an application if NHS staff or patients are involved in any phase of the research. These are usually organised by PCTs or NHS trusts.

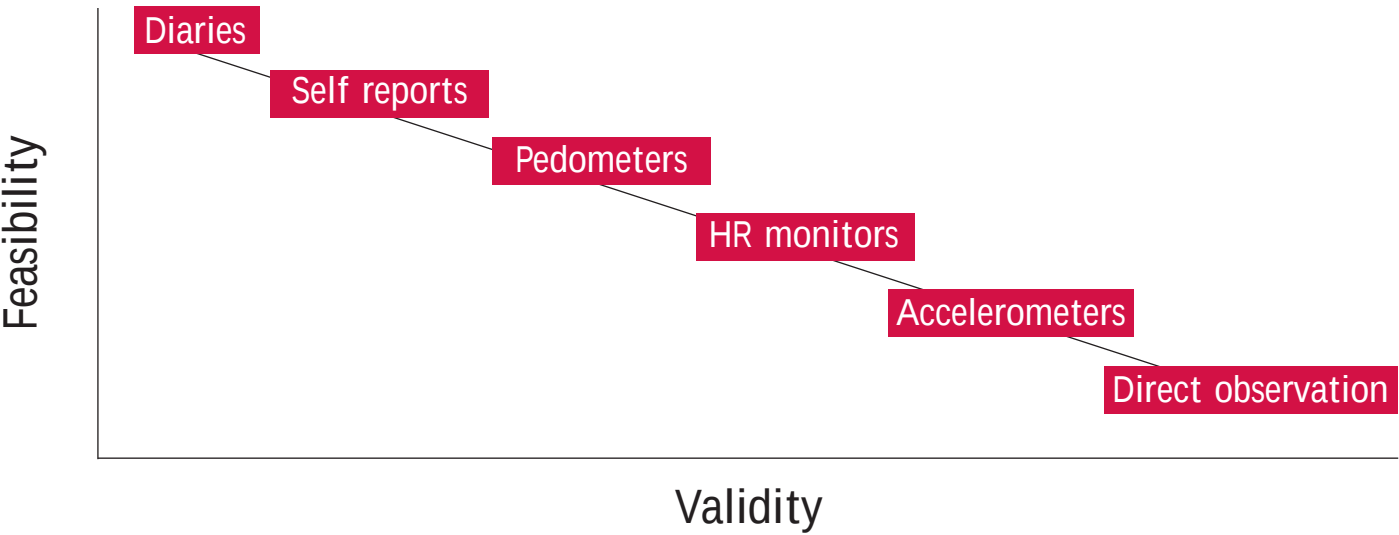
Measuring Sport and Physical Activity

There are a variety of methods available for measuring sport/physical activity levels but there is no gold standard (Welk, 2002). Key parameters for physical activity measurement include frequency (when/how often does it occur), intensity (how hard is the activity), time (the duration) and type of activity (walking, running, swimming etc.). There are over 30 methods available for measuring physical activity and their practicality or feasibility of use is illustrated in figure 2. The challenge is to choose the most valid, accurate and reliable

tool (or portfolio of tools) to measure physical activity within your intervention. (N.B. A valid instrument measures what it purports to measure, whereas a reliable instrument quantifies the degree to which a measure produces stable and repeatable results when used under the same conditions).

The choice of a measure depends on a balance between feasibility (ease and cost) against validity (complexity and expense). Diaries, self reports (questionnaire), pedometers, heart rate monitors,

accelerometers and systematic observation tools move from high feasibility to low validity and low feasibility to high validity. The key issue here is one of practicality and realism. In most cases self-report and pedometers will be the most practical option. Greater expertise and support would be required for more complex evaluations.



Self-report (questionnaire) Tools

The development of an appropriate, valid and reliable questionnaire to measure physical activity is a challenging task. Physical activity for health benefit comprises several components (e.g. intensity, frequency, duration, and type) that can be performed in different domains (e.g. occupational physical activity, transport physical activity, and/or physical activity during discretionary or leisure time).

Self-report tools are probably the most common choice of measuring tool because they are affordable and have low participant burden. Self-report can be used in a number of formats such as diaries, questionnaires (interview, self administered or proxy-administered, e.g. where parents may report on the activity of their children). Self report is influenced by the ability to comprehend a survey question and to recall activity patterns and hence, the most reliable tools tend to be 3 day or 7 day recall (physical activity participation over past 3 or 7 days). These self-report tools are recommended as they have adequate reliability and validity in large populations (Welk, 2002). Some of the most widely used questionnaires are the International Physical Activity Questionnaire (IPAQ), and the tool used in the Active People Survey (www.activepeoplesurvey.com). Population self-report of sport and physical activity has been inconsistent across the UK. Various tools have been used to gather data and all use slightly different measurement parameters (making trend analysis difficult) and methods (e.g. Health Survey for England, 2003; Active People Survey, 2005).

Active People Survey

The Active People Survey used a specially developed questionnaire (telephone interview survey) to provide base line measurements (2006) in order for Sport England to monitor sport and active recreation. Regular participation in sport and active recreation is defined by Sport England as taking part in moderate intensity sport and active recreation on at least three days a week (at least 12 days in the last 4 weeks) for at least 30 minutes continuously in any one session. Participation includes recreational walking and cycling. In the North West region, 20.6% of the adult population took part regularly in sport and active recreation (national average = 21%). Twenty seven per cent of people in the North West built some sport or activity into their lives, whereas 52.4% of adults did not taken part in any moderate intensity sport and active recreation of 30 minutes duration in the last 4 weeks. (www.sportengland.org/active_people_results). Professionals can register to use the Active People Diagnostic (APD) tool to find information on sedentary levels, recreational cycling and walking, sport and recreation participation and volunteering (www.activepeoplesurvey.com). Whilst it is important to note that the APS only records sport and recreation activity that lasts for over 30 minutes they provide the most up to date data on baseline participation rates in sport and active recreation. The APS does not assess occupational, household, gardening or transport-related physical activity (unless walking and cycling as a form of transport is a part of a healthy lifestyle choice).

International and Global Physical Activity Questionnaires

(IPAQ: www.ipaq.ki.se) (GPAQ: www.who.int/chp/steps/GPAQ)

One of the most widely used self-report (telephone or self administered) tools is the International Physical Activity Questionnaire (IPAQ), of which there are short and long forms. Both forms include an assessment of walking, and moderate and vigorous physical activities (MVPA), and are recommended as a viable method of monitoring population levels of physical activity globally for populations aged 15-69. Some authors have found over reporting of physical activity by the IPAQ (Rzewnicki et al., 2003) whereas there is conflicting evidence on its reliability which has been reported as poor (Rutten et al., 2003) and acceptable (r=.8). Low criterion validity results were found between IPAQ and an accelerometer (Craig et al., 2003) whereas Ekelund et al. (2006) found similar criterion validity scores for the short IPAQ, suggesting that the specificity to correctly classify people achieving 30 minutes of physical activity per day was sound. In summary the IPAQ was not initially designed for research and evaluation purposes or in small scale interventions.

However Ekelund et al's findings suggest that it is possible to use the long version of IPAQ to identify people who achieve activity guidelines (30 minutes per day). Therefore it is our view that the long version IPAQ can be used for evaluation purposes and in representative samples (further guidance is available at www.ipaq.se). Technical expertise may be required with the cleaning and analysis phases of IPAQ use.

Comparing the IPAQ and GPAQ

The GPAQ (2nd version) is very similar to the long version of the IPAQ. Both assess frequency, duration and intensity of physical activity. The IPAQ includes MET values of 3.3, 4.0 and 8.0 for walking, moderate and vigorous activity respectively whereas the GPAQ only includes the 4.0 and 8.0 MET values. The IPAQ requires respondents to recall physical activity over the past 7 days compared to a “typical week” in the GPAQ. The GPAQ has 15 questions in three domains (work, travel, recreation) and is validated for use in 16-84 year olds compared to 25 questions across 4 domains in the IPAQ (job, transport domestic/gardening, recreation) which is validated for use in 15-69 year olds. The GPAQ data can also be cleaned and analysed using free public health analysis software (www.cdc.gov/epiinfo) Recently the GPAQ has been selected as the tool for the physical activity element in the North West Public Health Observatory lifestyle questionnaire.



Physical Activity Questionnaire for Older Children and Adolescents (PAQ-C/A)

The Physical Activity Questionnaire for Older Children (PAQ-C) and Adolescents (PAQ-A) are validated (Crocker et al., 1997), self-administered, 7 day recall questionnaires, which assess general levels of physical activity in 9 to 15 year old children. There are no valid questionnaires for children under the age of 9.

IPAQ and GPAQ and PAQ-C/A are examples of a number of forms of self-report or interview administered questionnaire. Many others can be found in a special edition of Medicine and Science in Sports and Exercise: A Collection of Physical Activity Questionnaires (Pereira et al., 1997).

Accelerometry and Heart Rate Monitoring

Accelerometers and heart rate telemeters provide objective measures of physical activity and will be discussed together. It is unlikely that practitioners will use these because of the cost, technical expertise required for use and the complexity of data analysis. Nevertheless a brief explanation of these monitors and a case study will be described here.

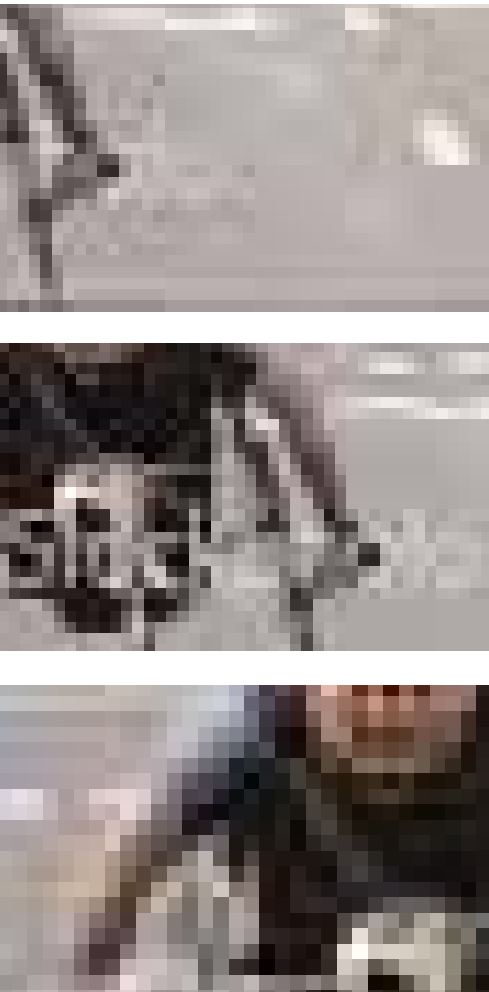
Heart Rate Monitoring

Heart rate monitoring usually uses a telemetric system (<http://www.polar.fi>) with a transmitter which comes in the form of a belt and fits around the chest and detects electrical impulses from the heart and converts these to beats per minute. These data are either stored in the belt or transmitted to a receiver in the form of a wristwatch. Heart rate telemeters can be programmed to record heart rate second to second or minute to minute (recording interval is called an epoch) continuously for weeks.

Accelerometry

Accelerometers (www.theactigraph.com) are small box devices usually placed on the waistband (on the wrist in wheelchair users) which record the vertical (uni-axial) or vertical, horizontal and diagonal (tri-axial) acceleration of the body. These accelerations are subsequently converted to gravitational counts per epoch duration. These instruments can record in second by second or minute-by-minute epochs.

The main advantage of heart rate monitoring and accelerometry is the relatively low participant burden and the relative ease by which data can be collected and more importantly analysed. Both heart rate telemeters and accelerometers are first initialised in conjunction with their respective software programmes when names, monitor numbers and start and end recording dates and times are set up for each individual. On completion of data collection the instruments are interfaced with a PC and data downloaded. Data can then be analysed by frequency, intensity and time and patterns of activity assessed across each minute of the day (www.theactigraph.com; www.polar.fi/polar/channels/uk). Data produced by heart rate and movement methods is extremely valuable and is not beset by problems of recall, although most studies also use objective methods alongside a diary which requires participants to record getting up and going to bed times. For example high heart rates at unusual times may be recorded in a diary and allow the evaluator to interpret the results with greater validity.(e.g. heart rate may be elevated due to stress rather than physical activity).



When is it best to use heart rate telemetry and accelerometry?

From a practitioner perspective, heart rate telemetry and accelerometry may be most useful when a precise quantification of physical activity is required e.g. in a Tai Chi for Arthritis Class (PA Audit 2005, Salford LA.) These methods could also be used in exercise referral or any other structured intervention. Data from the instruments would then be downloaded from a relatively small number of participants and the activity contribution of the intervention analysed and reported. In one study (Ridgers et al., 2005, 2006) activity levels were monitored in primary school playgrounds using both heart rate and accelerometry. This was entirely appropriate as the physical activity had a clear focus (play), it occurred in one setting and involved a group (young children) whose accurate recall of the frequency, intensity, duration and mode of activity is generally problematic. Moreover the data could be analysed by discrete time periods across a school morning or lunch break.

Pedometers

Many physical activity interventions and health promotion messages promote walking as a healthy and free form of physical activity available to all those physically able to take advantage of it. Pedometers can provide useful information on ambulatory walking however there are a number of problems with their use. Individual data such as stride length, body weight and age can be input into some pedometers. However data are sometimes rounded to the nearest 1 kg or 10 cm causing an immediate source of inaccuracy. The incorrect input of stride length is arguably the largest cause of error in estimating physical activity energy expenditure and distances covered during walking. The best use of pedometers is for recording steps and pedometers should always be manually checked for counts by using a calibrated shaker table or by hand (by counting each shake 1,2,3 etc and checking against the display). For representative data to be obtained participants need to wear a pedometer for 3 days (Tudor-Locke et al., 2005). Pedometers are a low cost method of generating accurate and reliable data (depending on the quality of the pedometer; Schneider et al., 2004). A range of pedometers can be viewed at a number of different web sites (www.walking.about.com; www.polygondirect.com). The recent range of newly developed piezo-electric pedometers are probably bests for research and evaluation purposes (www.new-lifestyles.com).

The daily target for physical activity is 10,000 steps per day (Tudor-Locke and Bassett, 2004). However 15,000 and 12,000 steps have been recommended for boys and girls respectively (Tudor Locke et al., 2004). The key aspect for activity intervention is not necessarily the debate over number of steps but whether total steps increase as a result of engaging in an activity intervention. Recent pedometer evaluations in schools have suggested that pedometers work as motivational tools (Butcher et al., 2007) and stimulate increases in physical activity.

Systematic Observation

Systematic observation involves a trained observer coding predetermined physical activity behaviours (sitting, walking, running) in various settings, (playgrounds, parks, homes etc). undertaken by participants over set time intervals. McKenzie has designed a number of systematic observation systems and some methods are (SOPLAY, SOFIT) are available on the internet (www-rohan.sdsu.edu/faculty/sallis/measures.html). One of the most comprehensive instruments to assess walking and cycling is the SPACES system (Systematic Pedestrian and Cycling Environmental Scan, Pikora et al., 2002). Systematic observation requires specific training of observers and can be undertaken live or by reviewing video media. This technique is extremely time consuming and should only be used on small groups in specific settings. Systematic observation tools are accessible for practitioners as they provide powerful data by combining both context and behaviour.

Recent Technologies (GPS, Mobile Telephones)

Recently, geographic position system (GPS) and mobile phone technology have been proposed as tools for physical activity measurement. The Nokia 500 Sport (www.techdigest.tv/2006/05/nokia_5500_spor) has a pedometer integrated into the phone and can also be used to develop physical activity interventions. Sending prompting SMS messages to participants enrolled on interventions to remind them to walk at lunch break or to take the stairs and not the lift is also a good way of promoting physical activity in a cost effective manner.

Example: In an unpublished pilot study 2 focus groups of 5 participants discussed their physical activity behaviour with an intervention lead. Each participant signed up to some simple behaviour change agreements. The intervention lead was allowed to send daily SMS texts to remind them of the behaviour changes. Examples of these were an early morning breakfast SMS to walk to work and/or take the stairs not the lift. In addition SMS messages stating, “put your pedometer on” or “text me your total step count” at the end of the day were also used. Data were then entered into a spreadsheet by hand and analysed at the end of the physical activity prompting intervention.

The Garmin System (www.garmin.com/products/forerunner201/) is a GPS monitor that detects outdoor activity and geographic positioning. The GPS system is useful for calculating distances and speeds as well as the location of activity. The integration of the system with appropriate software also allows mapping of activity. The limitations of GPS is that it only works outside and in areas where interference (cloud cover, trees, high rise buildings) is at a minimum. GPS can be used usefully with activity diaries and accelerometers.

Example: Tracking the use of a skate park by adolescents. Participants agreed to wear the GPS system whilst outdoors before and after the introduction of a skate park in their neighbourhood. The data from the GPS system provided evidence of the geographical location of the skateboarders, their distance travelled and speeds attained not only whilst at the venue but also during the travel to and from the venue. Usually GPS are combined with heart rate monitors or accelerometers and/or qualitative data from focus groups, interviews or photographs of locations.

Cognitive and Psycho-social Measures

Control over exercise and physical activity (Kerner and Grossman, 2001), enjoyment of exercise (Kendzierski and De Carlo, 1991), cost benefit (Marcus, et al., 1992), intention to exercise (Kerner and Grossman, 2001), self efficacy (Sallis et al. 1998) and social support (Sallis et al., 1987) also represent other health markers susceptible to change as a result of participating in an activity intervention and are worth considering for use within physical activity interventions.

Physical Environment Measures

There are many useful measures of physical environments relevant to physical activity. The Neighbourhood Environment Walkability Scale (NEWS, Saelens et al., 2003), Home Environment Scale (Sallis et al., 1997), Perceived Environment Scale (Ball et al., 2001), Awareness of Physical Activity Facilities Scale (Leslie et al, 1999), and Personal, Media, External Environments and Local Opportunity Scales (Stahl et al., 2001), are readily available and, after familiarisation, are appropriate tools for practitioners to use in their evaluation.

Abstracts on cognitive and psycho-social and physical environment measures are available from www.pubmed.org or by contacting the authors direct.

Other Outcome Measures

Changes in physical fitness scores, height, weight, body mass index (BMI), skin fold measures, blood pressure, run/walk times, weight lifted, flexibility and ratings of perceived exertion may also be useful evaluation measures within specific interventions such as exercise referral. Health outcomes such as perceived health status, increased feelings of well-being and improved mental health can also provide useful evidence of programme impact (Bowling, 1997).

Qualitative Research Techniques

Participation in sport and physical activity is a complex behaviour and is not always easy to capture accurately using some of the techniques we have seen so far. Qualitative research techniques allow behaviour to be studied in depth and often in an ongoing way. In order to enhance the rigour of qualitative research, various techniques can be employed such as triangulation (where quantitative & qualitative data are used together in an evaluation design); purposive sampling (where specific cases are selected); and sampling until saturation is reached (i.e. no new themes are emerging when you interview more people) can be used.

Interviews: One-to-one, Focus groups

Interviews are useful for tracking and exploring complex issues. A semi-structured interview format is often the best approach as it allows enough focus for specific questions to be asked by the evaluator however, it allows enough scope for the interviewee to give free-ranging opinions on unintended benefits and outcomes of the intervention under investigation. Focus groups (Morgan and Krueger, 1998) are a type of group interview, usually comprising about 8 people, which allow very specific questions to be explored. These are a useful tool to measure ongoing processes during an intervention, or as a reflective tool to look back retrospectively on an experience of an intervention.

Image-based techniques: Videos, video-diaries, disposable cameras

Images are increasingly being used to measure health-related behaviour as they allow the reality of the ongoing behaviour to be captured (Bauer and Gaskell, 2000).. They can also be used by the participants themselves and hence provide an insider perspective on an intervention. Pictures and images are a powerful mechanism for getting participants to reflect on behaviour and they can also be used as a trigger for a focus group conversation. Drawings (e.g. Draw and Write method) are a useful tool to use with children to engage them in the research process and overcome language barriers.

STAGE THREE

Data Analysis:
Data should be organised according to the type of response.

Qualitative analysis - If qualitative responses are generated then these will need to be coded. Coding is where similar thematic areas arising from within the textual evidence (interview transcript) are grouped together e.g. all factors relating to barriers to physical activity could be grouped as one thematic area. Coding allows patterns within the data to be defined and described. This data can be organised by a paper-based or electronic method (N-VIVO; www.qsrinternational.com). Key points for respondents will be recorded from which themes related to activity behaviour will emerge.

Quantitative analysis – When handling quantitative data a spreadsheet or database (Excel/Access/Epiinfo tool[www.cdc.gov/epiinfo]) will have to be designed, into which the data can then be inputted. From this, statistical analyses can be performed, and general descriptive statistics produced (means, ranges, standard deviations) on the outcome measure (physical activity). These data may be related to sub groups (sometimes called independent or grouping variables) such as gender age, SOA (Super Output Area-Post Code). Any use of data should conform to the Data Protection Act (www.data-protection-act.co.uk) where in dividual identity and related data should be coded appropriately and kept in different electronic files. These data will then be available for use in an overall evaluation of sport / physical activity programming.

STAGE FOUR

Dissemination:
In order to ensure action results from the evaluation it is vital that all data collected during the evaluation is analysed, interpreted and disseminated in an appropriate and timely manner.

Recent use of geographic information systems (GIS) may enable some data to be mapped. Some data may be released in the short-term to keep interest maintained. Ensure results are summarised in a succinct and meaningful way, statistics are fully explained and qualitative findings kept brief. A balance of both positive and negative aspects should be highlighted. Examples of good practice such as case studies, qualitative quotes or photographs may enrich the results and make the summary more interesting.

Case Study 1 (figure 3)

The Liverpool Sporting Playgrounds Project
Figure 3 reviews the Liverpool Sporting Playgrounds Intervention in the “stages of evaluation” model.

Needs analysis

Children’s activity during school play was low. Thus a major initiative was developed by the DfES and Nike. Schools in deprived areas applied for £20k funding to redesign their playground. Twenty one schools received money to do this. Partners involved in the study, such as the Sport Action Zone (SAZ), the Youth Sport Trust (YST), local education authority (school improvement officer), local authority architects and schools councils (teachers, pupils, parents, governors) were involved in discussion regarding the nature of playground redesign and the research approach.

Process evaluation

There were regular meetings held with architects, school councils, Youth Sport Trust, Sport Action Zone manager and the school improvement officer and research team. Evidence gathered using semi-structured interviews and notes from conversations and meetings as well as systematic observation (video) data from the school playgrounds permitted a detailed analysis of the process of change and implementation.

The Stages of Evaluation: The Liverpool Sporting Playgrounds Project

Summary Findings

The result of this rigorous evaluation was that children’s activity was increased and sustained for between 6 and 12 months. This resulted in all primary school in the unitary authority receiving funding to renovate their school playgrounds

Interpreting data using RE-AIM
Because of the data generated the programme will now “reach” all children, it “effectively” increased physical activity. All schools are “adopting” the basic intervention design. “Implementation” lessons have been learned and added to the process. “Monitoring” demonstrated a significant effect and the continued need for evaluation.



- 01 Needs analysis stage (formative evaluation) DfES invest £20k into 470 school play areas. Redesign involves multi-colour markings and playground structures Discussion with stakeholders: School councils (children&teachers), SAZ, School improvement officer, local authority architects, YST about nature of project and key aims
- 02 Process evaluation Regular meetings with architects, school council ,YST, SAZ and school improvement officer Evidence gathered using semi-structured interviews and notes from conversations and meetings as Well as systematic observation (video) data from the school playgrounds
- 03 Impact evaluation Measures increase in physical activity from accelerometry and heart rate measures Assesses physical activity and positive and negative behaviour after playgrounds were redesigned
- 04 Outcome evaluation Physical activity assessed at 6 weeks, 6 months and 12 months post interventionData analysed by age, gender and BMI group Findings disseminated to stakeholders.
- 05 RE-AIM

Outcome evaluation

Activity levels increased by about 4 percent in boys and girls although over half of this was vigorous activity. This increase was sustained 6 months after the intervention although effects dropped off after 12 months post intervention. The intervention effect was independent of gender, BMI category and age. Longer play duration had an additional positive effect on activity levels. Children, teachers and parents were proud of the playground and children were more alert in class. There were some difficulties in constructing some of the playground designs

Case Study 2

Stepping Out LEAP

(Local Exercise Action

Pilot) Evaluation:

An exercise promotion project aimed at people aged 50+ in Wigan borough. This evaluation study aimed to assess the impact of a variety of primary-care led, exercise interventions, targeted at the 50+ population in Wigan borough, on exercise behaviour. “Stepping Out” was one of ten LEAP (Local Exercise Action Pilots) demonstration projects in England. A range of interventions were implemented including a marketing and communication campaign, community outreach programme, senior peer mentoring, ‘Steps to Health’ exercise referral scheme, ‘Next Steps’ activity opportunities, Healthwalks, falls prevention and ‘Moving more often’ - targeting activities in social care settings.

Process evaluation

Regular steering group meetings were held between agencies and the over 50s forum in Wigan to discuss project delivery. Other processes such as the training of peer mentors were evaluated by conducting focus groups at intervals during the project’s time-course.

Impact evaluation

Exercise levels were measured using the self-report, 7 day short form IPAQ, at baseline, 6 month and 12 month intervals. The measures for this were decided by the national LEAP evaluation team (DH, 2006) and one key limitation was the lack of triangulation using an objective measure of exercise. Despite this, a large exercise data set was gathered in Wigan borough (n=961), although 12 month follow-up only yielded a response rate of 30%. Participation was unevenly distributed across the borough with low participation in the evaluation from areas of highest deprivation. The data has been mapped by gender and postcode to assess the relationship between deprivation and physical activity behaviour across the Wigan borough.

Table 2: A Review of the Whole Evaluation process

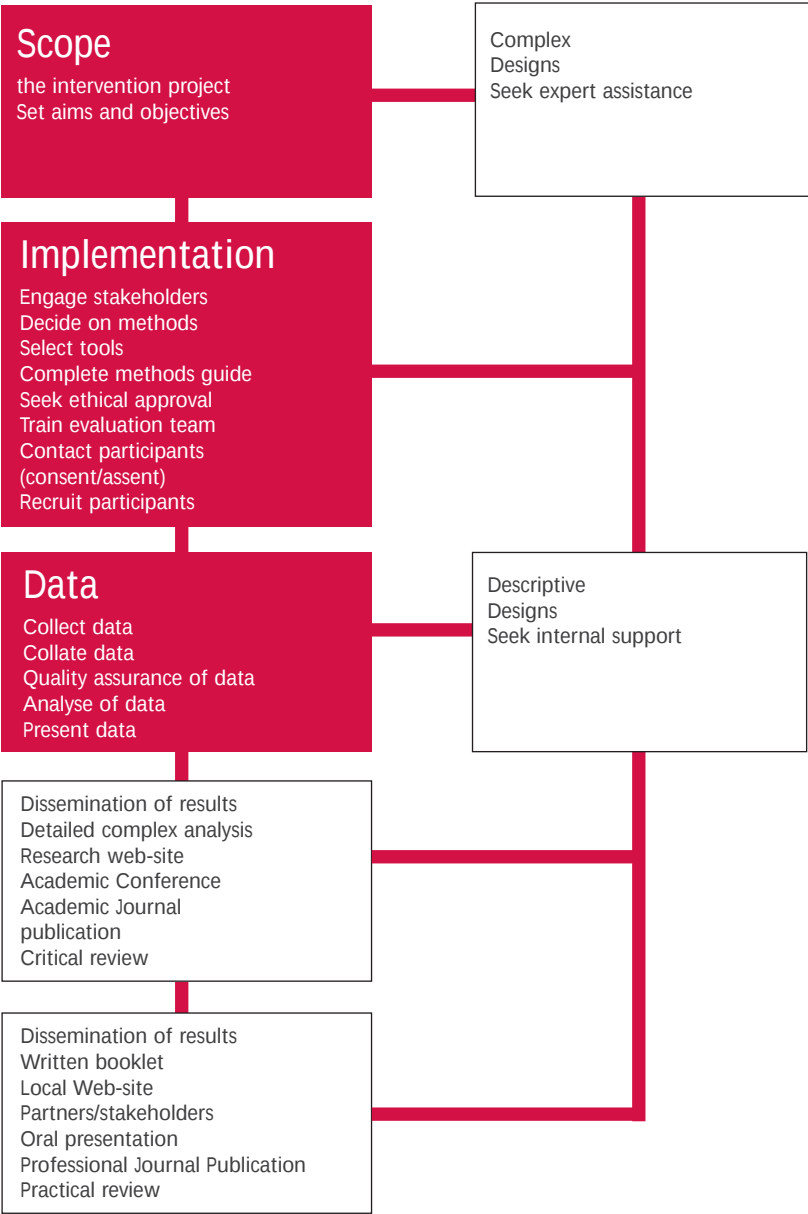
Programme stages	Level of measurement	Stages of measures	Examples of measures
	Design of PA programme	Formative evaluation measures	Responses of target group to testing of PA messages or programme materials; perception of stakeholders of programme's likely success
Programme started	Population reached	Process and implementation measures	The proportion of people attending the programme event; Or % of health professionals participating; The programme was delivered as intended; Environmental changes were carried out as planned; inter-agency planning / partnerships materialise and are maintained
	Proximal effects	Individual level measures	Awareness of PA health benefits Cognitive changes such as self-efficacy, intention to be more active, beliefs Social supports; enhanced social influences; social environment; social capital
		Inter-individual Measures Individual level measures	PA behavioural changes – increased walking, Increased moderate and vigorous activities; Decreased sedentary behaviour or screen time; Increases in incidental PA, active transport
	Health promotion impact measures	Physical environmental measures	Changes to physical environments implemented/completed
Programme sustainability		Community level change ##	Policies developed and implemented; Programme elements institutionalised; Programme elements self-sustaining without the presence of programme initiators
Programme derived policy	Long term outcomes		Reduced disease incidence or mortality from inactivity-related conditions; improved well-being/ quality of life/ social capital
		Other [non health] outcomes ##	Environment which facilitates PA better, such as improved public transport, better parks, urban planning, cultural norms and values changed to demand PA infrastructure Sustained policy changes to facilitate PA enhancements
Dissemination			Spread of PA promoting culture and its policy and resources for effective programmes to all regions where inactivity is a problem.

Summary of evaluation process

Figure 4 overviews the evaluation process and dichotomizes projects into complex and descriptive designs. Descriptive evaluations involve straightforward analysis and reporting of data. In descriptive designs numbers/ percentages of participants, quotes from interviews are reported whereas complex designs involve more advanced statistical content analysis of quantitative and qualitative data respectively.

In descriptive evaluations expertise within your organization or partnership may be sufficient whereas for complex designs, external support and expertise may be required. Table 2 (Bauman et al. 2006) reviews the whole evaluation process

Figure 4 An Overview of the Evaluation Process



References

Ball K, Bauman A, Leslie E and Owen N (2001): Perceived environmental aesthetics and convenience and company are associated with walking for exercise among Australian adults. Preventive Medicine. 33, 434-440.

Bauer MW and Gaskell G Eds (2000) Qualitative Researching with Text, Image and Sound. Sage; London.

Bauman A, Phongsavan P, Schoeppe S and Owen N (2006) Physical activity measurement- a primer for health promotion. IUHPE – Promotion and Education, 13, 92-103.

Booth ML (2000). Assessment of Physical Activity: An International Perspective. Research Quarterly for Exercise and Sport, 71 (2): s1 14-20

Bowling A (1997) Research Methods in Health: Investigating Health and Health Services, OUP: Buckingham

Butcher Z, Fairclough S, Stratton G. and Richardson D. (2007) The Effect of Feedback and Information on Children's Pedometer Step Counts at School. Pediatric Exercise Science, 19(1), (pages)

Crocker PR, Bailey DA, Faulkner RA, Kowalski KC, McGrath R. (1997) Measuring general levels of physical activity: preliminary evidence for the Physical Activity Questionnaire for Older Children. Med Sci Sports Exerc. 29:1344-9.

Department of Culture, Media and Sport. (2002) Game Plan: a strategy for delivering Government's Sport and Physical Activity Targets. Cabinet Office;London.

Department of Health (1998) Health Survey for England. Department of Health Publications; London.

Department of Health (2003) Health Survey for England. Department of Health Publications; London.

Department of Health (2004a) Choosing Health? Choosing Activity: A Consultation on how to increase physical activity. Department of Health Publications; London.

Department of Health (2004b). At least 5 a week: Evidence on the impact of physical activity and its relationship to health. A report from the Chief Medical Officer. Department of Health Publications; London.

Department of Health (2005) Choosing Activity: a physical activity action plan. Department of Health Publications; London.

Dugdill L (2001) Evaluation Supplement: Framework for Action. Health Development Agency, London, U.K.

Health Development Agency. (2004) The effectiveness of public health interventions for increasing physical activity among adults: a review of reviews. Health Development Agency; London.

Hillsdon M, Foster C, Thorogood M (2005) Interventions for promoting physical activity, Cochrane Database of Systematic Reviews, Issue 1. Art No: CD003180.

Kendzierski D and DeCarlo KJ (1991):Physical Activity Enjoyment Scale: Two validation studies. Journal of Sport and Exercise Psychology, 13(1); 50-64.

Kerner MS and Grossman AH (2001): Scale construction for measuring attitude, attitude, beliefs, perception of control, and intention to exercise. J Sports Med Phys Fitness 41(1): 124-31.

Laporte RJ, Montoye HJ and Caspersen CJ (1985) Assessment of physical activity in epidemiologic research: problems and prospects. Public Health Rep, 100:131-46.

Leslie E, Owen N, Salmon J, Bauman A, Sallis JF and Kai Lo S (1999): Insufficiently Active Australian College Students: perceived personal, social, and environmental influences. Preventive Medicine. 28, 20-27.

Marcus BH, Rakowski W and Rossi JS (1992): Assessing motivational readiness and decision-making for exercise. Health Psychology. 11, 257-261.

McKenzie TL, Sallis JF, Nader PR (1991) SOFIT: System for observing fitness instruction over time. J Teach Phys Edu, 11:195-205.

Morgan DL and Krueger RA (1998) The Focus Group Kit. Sage; London.

Ovretveit J (1998) Evaluating Health Interventions. Oxford University Press; Buckingham.

Sport England Active People Survey (2005-2006) http://www.sportengland.org/index/get_research/research/active_people/aps_results.htm (accessed 5/12/06)

Stratton G, Ridgers ND, Gobbii R, and Tocque, K (2005) Physical Activity Exercise, Sport and Health: Regional Mapping for the North-West. (www.nwph.net/pad/accessed 13/3/07)

Pate RR, Trost SG, Felton GW, Ward DS, Dowda M and Saunders R (1997). Correlates of physical activity behavior in rural youth. Research Quarterly for Exercise and Sport, 68, 241-248.

Pereira MA, FitzerGerald SJ, Gregg EW, Joswiak ML, Ryan WJ, Suminski RR, Utter AC, and Zmuda JM. (1997) A collection of Physical Activity Questionnaires for health-related research. Med Sci Sports Exerc. 29: S1-205.

Pikora TJ, Bull FCL, Jamrozik K, Knuiman M, Giles-Cortie B and Donovan RJ (2002): Developing a reliable audit instrument to measure the physical activity environment for physical activity. American Journal of Preventive Medicine. 23 (3), 187-194.

Richardson M, Ainsworth B, Jacobs D, and Leon W. (1996) Validity of the Godin physical activity questionnaire in assessing leisure time physical activity. Medicine and Science in Sports and Exercise, 28, S32 (Abstract).

Ridgers ND and Stratton G (2005). Physical activity during school recess - The Liverpool Sporting Playgrounds Project. Pediatric Exercise Science, 17, 281-290.

Ridgers ND, Stratton G and Fairclough SJ (2006) Assessing physical activity during recess using accelerometry, Preventive Medicine, 41, 102-107.

Saelens BE, Sallis JF, Black JB and Chen D (2003) Neighborhood-based differences in physical activity: and environment scale evaluation. American Journal of Public Health. 93(9), 1552-1558.

Sallis JF and Saelens BE (2000) Assessment of physical activity by self-report: status, limitations, and future directions. Research Quarterly for Exercise and Sport, 71:409.

Sallis JF, Johnson MF, Calfas KJ, Caparosa S, and Nichols JF (1997) Assessing perceived physical environmental variables that may influence physical activity. Research Quarterly for Exercise and Sport. 68(4), 345-351.

Sallis JF, Bauman A, and Pratt M (1998): Environmental and policy interventions to promote physical activity. American Journal of Preventive Medicine. 15 (4), 379-397.

Silverman D (2006) Interpreting Qualitative Data; 3rd edition. Sage; London.

St_hl T, R_tten A, Nutbeam D, Bauman A, Kannas L, Abel T, L_schen G, Rodriguez DJA, Vinck J and van der Zee J (2001): The importance of the social environment for physically active lifestyle – results from an international study. Social Science and Medicine. 52, 1-10.

Stamatakis E, Ekelund U, and Wareham NJ. (2007) Temporal trends in physical activity in England: The Health Survey for England 1991 to 2004. Preventive Medicine. Jan 8; [Epub ahead of print].

Stratton G (2006) Ethics: a primary consideration for BASES members. The Exercise Scientist. BASES. Leeds UK. Tudor-Locke C and Bassett DR Jr (2004) How many steps/day are enough? Preliminary pedometer indices for public health. Sports Medicine. 34:1-8.

Tudor-Locke C, Burkett L, Reis JP, Ainsworth BE, Macera CA, and Wilson DK (2005) How many days of pedometer monitoring predict weekly physical activity in adults? Preventive Medicine. 40:293-8.

Welk, G.J. (Ed) (2002): Physical activity assessments for health-related research. Human Kinetics;

Physical activity measurement:-

grants.nih.gov/grants/guide/pa-files/PAR-06-104.html (accessed 3/3/07)

IPAQ questionnaire

www.ipaq.se (accessed 4/3/07)

GPPAQ questionnaire

www.dh.gov.uk/assetRoot/04/14/08/80/04140880.pdf (accessed 7/3/07)

Active Peoples Survey

www.activepeoplesurvey.com (accessed 6/3/07)

Pedometers walking.about.com/cs/measure/tp/ pedometer.htm

www.polygondirect.com accessed 4/3/07)

San Diego State University

Various tools:

www-rohan.sdsu.edu/faculty/sallis/measures

Systematic observation

7DPAR

Accelerometer

www.theactigraph.com (accessed 11/3/07)

Heart rate telemetry

www.polar.fi/polar/channels/uk (accessed 11/3/07)

Research triangulation

www.tele.sunyit.edu/traingulation.htm (accessed 9/4/07)

Appendix 1.

Glossary

Cardiorespiratory fitness (also called aerobic endurance or aerobic fitness) Cardiorespiratory fitness is the ability of the body's circulatory and respiratory systems to supply fuel and oxygen during sustained physical activity.

Exercise Exercise is physical activity that is planned or structured. It involves repetitive bodily movement done to improve or maintain physical fitness.

Household physical activity Household physical activity includes a large number of domestic activities such as washing, gardening and cleaning.

Inactivity Inactivity is not engaging in physical activity beyond daily functioning. Sometimes used interchangeably with sedentary (sitting) activity such as watching televison, sitting and reading.

Kilocalorie The amount of heat required to raise the temperature of 1 kg of water 1° C. Kilocalorie is the ordinary calorie discussed in food or exercise energy-expenditure tables and food labels (3,500 kilocalories of food energy = 1 pound of body weight) 1 Kcal = 4.2 KJ (Kilojoules)

Leisure-time physical activity Leisure-time physical activity is exercise, sports, recreation, or hobbies that are not associated with activities as part of one's regular job duties, household, or transportation.

MET (standard metabolic equivalent) One MET = the energy (oxygen) used by the body whilst sitting quietly. Activity intensity can be presented in METS. ∑ Any activity that burns 3 to 6 METs is considered moderate-intensity physical activity. ∑ Any activity that burns > 6 METs is considered vigorous-intensity physical activity.

Moderate-intensity physical activity Moderate-intensity physical activity refers to a level of effort which may stimulate mild breathlessness ≥3-6 METs or 3-7 kcal_min-1.

Physical activity Physical activity is any bodily movement produced by skeletal muscles that results in expenditure of energy:

Physical fitness Physical fitness is a set of attributes a person has in regards to a person's ability to perform physical activities that require aerobic fitness, endurance, strength, or flexibility and is determined by a combination of regular activity and genetically inherited ability.

Reliability Reliability is the "consistency" or "repeatability" of the measures being used. Reliability does not imply validity.

Occupational physical activity Occupational physical activity is completed regularly as part of one's job. It includes activities such as walking, hauling, lifting, and packing

Transportation physical activity Transportation physical activity is walking, biking or wheeling (for wheelchair users), or similar activities to and from places such as: work, school or community facilities.

Triangulation Triangulation is the application and combination of several research methodologies in the study of the same phenomenon. (See: www.tele.sunyit.edu/ traingulation.htm for further detail)

Validity Validity determines whether the research measures that which it was intended to measure or how truthful and trustworthy the results are. Valid measures are reliable measures

Vigorous-intensity physical activity Vigorous-intensity physical activity may be intense enough to cause breathlessness and significant physical effort: ≥6 METs or 7 kcal_min-1

Weight-bearing physical activity Physical activity that moves a load or has impact on the skeleton (such as jumping or skipping).

Acknowledgments

The authors would like to acknowledge the support of:

Professor Les Burwitz, Manchester Metropolitan University, and the British Association of Sport and Exercise Sciences (BASES) - the original idea for this document was initiated from a BASES “Measuring Physical Activity” workshop, 2006.

Jackie Brennan (North West Regional Health and Physical Activity Coordinator) and Shileen Tarpey (Sport and Physical Activity lead, Sport England North West) for their constructive comments on drafts of this document.

Liam Curtis (The Design and Print Group, University of Salford) for designing the document.