Methodological Issues in Measuring and Evaluating Physical Measures to Promote Walking and Cycling

John Preston (University of Southampton) and Jane Powell (UWE Bristol) in conjunction with Christian Brand (University of Oxford), Fiona Bull (Loughborough University), Ashley Cooper (University of Bristol), Nanette Mutrie (University of Strathclyde), David Ogilvie (MRC Epidemiology Unit, Cambridge) and Harry Rutter (University of Oxford).


Preliminary Draft – Not to be quoted.

1. Introduction

This paper represents some of the initial thoughts of the iConnect (Impact of CONstructing Non-motorised Networks and Evaluating Changes in Travel) consortium which is funded by the Engineering and Physical Sciences Research Council and consists of eight academic groups, based at the Universities of Bristol, East Anglia, Loughborough, Oxford, Southampton, Strathclyde, West of England (UWE) and at the MRC Epidemiology Unit in Cambridge. It brings together expertise in energy, environmental, physical activity, public health and transport research, including health economics, transport economics and computerised urban modelling. With researchers from backgrounds in the physical, life and social sciences, it aims to be genuinely interdisciplinary with a commitment to multi-method approaches.

The consortium’s aim is to measure and evaluate the changes in travel, physical activity and carbon emissions related to Sustrans’ Connect2 project, which won the public vote in the Big Lottery Fund’s People’s £50 Million Competition in December 2007. Connect2 is an ambitious UK-wide project that aims to transform local travel in 79 communities by creating new pathways, crossings and bridges to overcome barriers such as busy roads, rivers and railways, giving people easier and healthier access to their work, schools, shops, parks and countryside. The first scheme, a refurbishment of a viaduct across the River Nith in Dumfries, was completed in July 2008. The objectives of the consortium are:

(1) To develop an evaluation framework and measurement tools for use in assessing the impacts of infrastructural interventions on travel behaviour, physical activity, carbon emissions and energy use (and their inter-relationships) at the individual and population level.

(2) To apply the framework and tools thus developed to quantify the changes in travel behaviour, physical activity and carbon emissions and the changes in putative mediating factors associated with infrastructural interventions at a purposive sample of case-study

1 See www.sustransconnect2.org
Connect2 sites and to explore why these interventions are (or are not) effective, in what ways, for whom and in what circumstances.

(3) To determine by means of a randomised controlled trial whether an additional tailored self-help intervention delivered at the level of the household can enhance the effects of an infrastructural intervention.

(4) To evaluate the Connect2 programme in terms of its economic performance at a broad strategic level.

The overall structure of the iConnect research programme is indicated by Figure 1. iConnect is a five year research programme, with work commencing in May 2008 and still at a very preliminary stage. Initial emphasis is being placed on developing appropriate measurement tools and evaluation frameworks – which provides the focus for this paper.

**Figure 1: Structure of the iConnect Research Programme**

This paper considers some issues in accurately measuring walking and cycling in section 2. In section 3 methodologies for assessing the relationships between context, mechanism and contexts are considered. The role of randomised control trials is briefly
examined. In section 4 issues surrounding evaluation will be considered, with particular reference to existing tools such as HEAT and NATA that have been used to assess health and travel benefits respectively. Finally in section 5 some preliminary conclusions are drawn.


UK National Travel Survey (NTS) data show that walking and cycling trips decreased by about 15% between 1996 and 2005 (Department for Transport (DfT), 2006, Chart 2.2a). The National Cycling Strategy aimed to quadruple cycling between 1996 and 2012, but this target was dropped in the 2004 Future of Transport white paper (DfT, 2004), while Traffic Advisory Leaflet 2/00 outlined a framework for local walking strategies without prescribing targets. The need to promote walking and cycling has also been recognised by devolved administrations within the UK, for example in the Welsh Assembly Government’s Walking and Cycling Strategy (2002) and Transport for London’s Walking (2003) and Cycling (2004) Action Plans. To date, however, there is little evidence that these initiatives have been effective in promoting walking or cycling on a national scale, although there have been successes at the local scale, for example, in Hull, London and York (DfT, 2005).

The difficulties of measuring and assessing trends in walking, cycling, and physical activity in general are illustrated by the observations that NTS data suggest a decrease in distance cycled in the first half of this decade whereas traffic counts suggest an increase (DfT, 2005) and that it has been impossible to establish from Health Survey for England data whether the trend in overall physical activity is rising, falling or static (Stamatakis et al., 2007). Accurate measurement of walking and cycling using self-report instruments is notoriously difficult. Short trips tend to be under-recorded in travel diary surveys. For example, the NTS only records short trips (less than one mile) on the seventh day of a week long survey and only includes trips made on a public highway (Christophersen, 2008). Some groups associated with high levels of cycling (e.g. young men and students) tend to have low response rates to surveys. Many traditional physical activity questionnaires are designed to ascertain leisure activities rather than walking or cycling as utilitarian modes of transport. The EPSRC funded CAPABLE project shows that self report instruments are particularly problematic for children (Mackett et al., 2007). Automated counts, based on pneumatic tubes, inductive loops, piezoelectric pressure mats or infra-red beams, are often inaccurate. Manual methods, including CCTV analysis, are usually more accurate but also more expensive, although image recognition technology offers some prospects of automating some of this work in the future. Pedometers, cycle computers and personal GPS devices can provide accurate objective measurements and indicate the extent of under reporting in travel diary surveys (e.g. Stopher and Greaves, 2007).

At this stage the iConnect research project is proposing to produce two principal measurement tools: a longer postal questionnaire, mainly intended for use in researching the effects of interventions at the community or population level, and a shorter user intercept questionnaire, mainly intended for distribution by hand to pedestrians and
cyclists using a particular route or facility as part of the local evaluation of specific initiatives. Such user intercept questionnaires are already routinely used by Sustrans, e.g., to assess the uptake of new sections of the National Cycle Network. In order to do this, we will collate instruments currently available for measuring travel behaviour and physical activity and for imputing carbon emissions and critique their fitness for purpose in terms of their measurement properties, their test–retest reliability, their concurrent validity relative to objective measures of activity, movement and energy consumption (e.g., pedometers, cycle computers, accelerometers, GPS devices, odometer readings and car fuel purchase receipts), and their face validity for application in the UK. We will also assess the predictive validity of routinely collected data (such as traffic counts and automated bicycle counters) for use as a potential proxy indicator of the population impact of infrastructure interventions on travel and physical activity behaviour as an alternative to collecting survey data. We will draw on the expertise and previous work of the investigators and their institutions, e.g. on integrated emissions profiling at the personal level (Brand, 2006) and on the development and validation of new physical activity questionnaires (PAQs) such as G(Global)PAQ (Armstrong and Bull, 2006), I(International)PAQ (Craig et al., 2003), E(European)PAQ (Wareham et al., 2002) and R(Recent)PAQ (Besson et al., 2002). We will also incorporate the emerging findings from the ALPHA project, a European review of methods for assessing physical activity currently in progress. As well as measuring the behaviours and primary outcomes of interest, the longer questionnaire will also include items to assess characteristics of the intervention itself, the context of the intervention, and putative mediating factors. These items will be selected on the basis of a collation and synthesis of currently available items and the development of an interdisciplinary consensus as to which are most fit for purpose in this project, again drawing on the investigators’ links with initiatives such as the International Physical Activity and Environment Network (IPEN). Working closely with the Sustrans Research and Monitoring Unit, we will then pilot the feasibility and acceptability of both questionnaires and assess their test–retest reliability and concurrent validity relative to objective measures before full scale deployment. This will involve a total pilot sample of around 600 usable responses in two waves.

3. **Contexts, Mechanisms and Outcomes**

Our approach is informed (but not defined) by the realist approach to evaluation. Pawson and Tilley (1997) argue that evaluative research should be designed to test theories about how interventions work rather than to test whether a given intervention ‘works’ in an aggregate, generalisable sense. We suggest that Connect2 should therefore be regarded as a set of multiple unique implementations of a minimally-specified principle, summarised on the Connect2 website as ‘creating new routes for the local journeys we all make every day’. From an evaluative perspective, Connect2 is therefore a completely different proposition from interventions such as TravelSmart or Walk In To Work Out (which could be regarded as more-or-less standard packages, albeit implemented with a degree of tailoring to each context), or even some ‘hard’ measures such as speed cameras (whose appearance and function are more-or-less invariant, even though the contexts into which they are inserted may vary).
We propose that the Connect2 case study sites are used to examine two sorts of theory:

- A general theoretical model of how the physical environment is supposed to influence active travel.
- A ‘middle-range theory’ that explains how this particular type of intervention is supposed to work.

For the first, we propose a simple socio-ecological model would suffice such as that of Saelens et al. (2003), whose model shows individual, psychosocial and environmental factors interacting to influence walking and cycling as a mode of transport and (perhaps differently) as a leisure activity (see also Figure 2).

**Figure 2: A Socio-Ecological Model of Walking and Cycling**

This type of model has permeated much previous work including that of PEACH (Personal and Environmental Associations with Children’s Health) and other work by Ashley Cooper and colleagues (Cooper et al. 2005) as shown by Figure 3. This indicates that contextual factors can be categorised into generic and intervention specific. These generic factors may be seen as having a distal causal influence, whilst the intervention specific factors may be associated with more proximal causation.
The second, ‘middle-range theory’ about how Connect2 is supposed to work would be elicited through interviews with senior Sustrans personnel (for the Connect2 programme in general) and with key local stakeholders at each case study site. This would form an important part of the ‘baseline contextual fieldwork’, which would also include collation of quantitative data from the Census, the Index of Multiple Deprivation, accessibility measures etc. By combining the first (general) and second (specific) theoretical perspectives we ought to be able to articulate a set of hypotheses — expressed in terms of context, mechanism and outcome; the ‘CMO configurations’ advocated by Pawson and Tilley — which, when tested, would help to confirm or disconfirm components of these theoretical principles. This would be the basis of our claim for making generalized causal inferences from a set of highly contextual case studies whose selection exemplifies what Shadish et al. (2002) have called the ‘purposive sampling of heterogeneous instances’. To this end we are proposing the following case studies that include a mix of large urban, small urban and rural locations and that also include a mix of utilitarian and recreational travel:
1. Bridge to Nowhere (Glasgow)
2. Whitlingham Country Park (Norwich)
3. Banbury (Oxfordshire)
4. Itchen Walkway (Southampton)
5. Ottery St Mary (Devon)

We will undertake a longitudinal cohort (panel) observational study of the local resident population in the vicinity of each site, and compare and interpret the findings drawing on the realist evaluation perspective as discussed above. A common core research module will be applied at each site. This will consist of: (a) a baseline (‘before’) postal household survey before the new infrastructure is opened to collect demographic and socioeconomic characteristics and baseline data on mediators and behaviour; (b) baseline fieldwork to establish an understanding of the pre-intervention context (field visits, photographs, observations of travel behaviour at key locations or along key routes, and interviews with key local stakeholders); (c) a follow-up (‘after’) survey of the original cohort of survey respondents to identify changes in primary and secondary outcomes of interest after one year (to control for possible seasonal variation in travel behaviour), and (d) repeating parts of the baseline fieldwork to assist with interpretation and explanation of the quantitative survey findings. A total sample of 6,000 useable responses is envisaged from the baseline survey and 3,600 responses in the follow-up survey. Using multivariate regression modelling and/or multilevel modelling as appropriate, we will analyse changes in the primary and secondary outcome measures of interest and their relationships with (a) demographic group (age and sex), (b) socioeconomic status (e.g. household income and car ownership), (c) proximity to the intervention (e.g. by assigning respondents’ home addresses to concentric buffers around access points for the new infrastructure), and (d) changes in putative mediating factors (psychosocial characteristics, perceptions of the local environment, and accessibility of local services and opportunities).

Each case study site will also be used to collect additional, in-depth data related to impacts on either (a) travel and carbon emissions or (b) physical activity. At the ‘travel and carbon emissions’ specialist sites (provisionally Banbury and Southampton), we will undertake a fuller assessment of changes in travel behaviour and their net impact on energy use and carbon emissions, e.g. by collecting GPS data, odometer data, fuel purchase data, technical details of displaced modes of transport, etc. from a subset of the cohort of survey respondents. At the ‘physical activity’ specialist sites (provisionally Norwich and Ottery St Mary), we will undertake a fuller assessment of the impact of the intervention on overall physical activity, e.g. by using more detailed physical activity questionnaires and accelerometers in a subset of the cohort to establish baseline levels of physical activity and track changes in overall energy expenditure or substitution effects between domains of physical activity (transport, occupational, domestic and leisure), or by using GPS devices in a subset of the cohort to track changes in patterns of use of green space.

The literature suggests that a supportive environment may be a necessary, but not a sufficient, condition to encourage regular physical activity and that people with a
supportive environment and positive cognitions about walking may walk up to eight times as much as those without (Giles-Corti, 2006). We will therefore establish a randomised controlled trial (RCT) at one further Connect2 site, provisionally Glasgow, to examine whether a tailored self-help intervention delivered at the level of the household can enhance the effects of an infrastructural intervention.

The first task of the RCT will be to develop the promotional intervention. There is evidence that appropriately tailored self-help materials can be effective in promoting both active commuting (the Walk In To Work Out trial: Mutrie et al., 2002) and walking in general (Ogilvie et al., 2007). The materials used in Walk In To Work Out have been disseminated through Scotland’s Health at Work scheme and the DfT. We intend to develop these materials in three ways: first, by tailoring them to the specific infrastructural changes at the chosen intervention site; second, by making use of new visualisation technology which can provide animations of possible new routes to be viewed using devices such as mobile phones, iPods or DVD players; and third, by including the potential motivating factor of individuals’ responsibility to reduce their carbon emissions.

The second main task will be to evaluate the effectiveness of the additional promotional intervention in a randomised controlled trial. We propose three arms to the trial, in which local households will be exposed to (a) the Connect2 infrastructural improvements only; (b) Connect2 infrastructural improvements plus a low-technology intervention using traditional low-cost printed self-help materials; or (c) Connect2 infrastructural improvements plus a high-technology intervention using the visual imaging techniques described above to personalise the opportunities for walking and cycling provided by the new infrastructure. Households in the study area will be allocated to one of these three arms using a randomisation procedure clustered by small area (e.g. by unit postcode) to minimise the likelihood of contamination between groups. The self-help intervention will be delivered shortly after the infrastructural intervention has been completed. In effect, this trial will constitute a ‘special case’ of the intervention case studies. We will collect and analyse baseline and follow-up data in the same way as for the other detailed case studies. We will collect additional qualitative data to gain further insight into participants’ reaction to, engagement with and use of the promotional materials and the facilitators of and barriers to their taking up the opportunities provided by the infrastructural intervention. We will also conduct an economic evaluation of the costs, resource savings and effectiveness of additional promotional interventions drawing upon previous work in this area (Kelly et al., 2005).
4. Towards an Evaluation Framework

Increasing walking and cycling in the population as a whole would help to address goals in three complementary areas of public policy. First, with respect to travel, traffic congestion has continued to increase; the air quality, noise and other environmental impacts of road traffic continue to cause concern; and the rate of improvement of road safety has slowed, particularly with respect to the incidence of fatal road accidents (DfT, 2007). Transport research on Smarter Choices policy interventions has demonstrated that a modal shift from car travel to more benign modes of transport (walking, cycling or public transport) of around a half could be achieved in the DfT’s three Sustainable Travel Demonstration Towns (Sustrans, 2007). Individualised marketing of these modes of transport to households has been found to be effective in promoting their use among motivated participants (Cairns et al., 2004, Ogilvie et al., 2007). A shift towards walking and cycling in the population at large would also reduce the socially divisive and inequitable effects of a transport system dominated by less sustainable modes (Woodcock et al., 2007).

Second, with respect to physical activity, walking and cycling offer an ideal opportunity for people to incorporate more moderate-intensity physical activity into their daily lives; two-thirds of UK adults would benefit from being more physically active, thereby reducing their risk of chronic diseases such as diabetes and coronary heart disease (Chief Medical Officer, 2004). A growing body of evidence, mostly from cross-sectional studies, suggests that people who live in supportive environments, such as areas with safe, convenient walking and cycling routes to local destinations, are more likely to walk or cycle than those who do not (e.g. Committee on Physical Activity, Health, Transportation, and Land Use, 2005); improving the infrastructure for walking and cycling has also recently been identified as one of the most important strategies for tackling obesity in the UK (Foresight, 2007).

Third, with respect to carbon, the current draft climate change bill commits the UK to reducing greenhouse gas emissions by 60% between 1990 and 2050, but it is widely accepted that this will not be achieved by technological innovation alone (e.g. Potter, 2007). Cars largely dominate surface transport carbon emissions, with a highly unequal distribution of emissions amongst the population (Brand and Boardman, 2008). Targeting the high (car) users is key to successful changes in travel behaviour and related carbon emissions.

Our evaluation framework will be designed around the need to collect data on context, mechanism and outcomes. Under the heading of ‘mechanism’, we propose to collect data on putative mediating factors comprising psychosocial characteristics (such as self-efficacy), changes in perceptions of the local environment, and changes in objective measures of accessibility of local destinations. Under the heading of ‘outcomes’, we intend to examine changes in one core primary outcome — walking and cycling behaviour — and a variety of secondary outcomes: (a) trip generation or trip suppression and modal shift in overall travel behaviour, (b) household fuel use, fuel and other transport expenditure and carbon emissions, and (c) overall physical activity. We will
also critically examine the use of GIS-based measures of accessibility (such as the Accession accessibility planning tool), a measure of general health and wellbeing such as the Short Form (SF-8), and routinely-collected injury data from hospitals (HES for England, SMR01 for Scotland) or the police (STATS19).

We will also examine economic outcomes in terms of cost-per-QALY (quality-adjusted life years) from a cost-utility perspective, cost-benefit ratios from a societal perspective and costs and consequences from a realistic evaluation perspective. Methodological guidance on the economic appraisal of health effects related to walking and cycling, entitled the Health Economic Assessment Tool (HEAT) has recently been released under the auspices of WHO Europe (Cavill et al., 2007, Rutter et al. 2007) (see also Figure 4). At present, HEAT can be applied to estimate the outcomes, benefits and resource savings of interventions that encourage cycling and it is already being adopted by DfT as a tool within the latest version of WebTAG (Transport Analysis Guidance) the online resource for the conduct of economic appraisal in transport. As part of this work package we intend to test and refine HEAT for cycling and to initiate the development and testing of HEAT for walking.

Figure 4: Health Economic Assessment Tool for Cycling
Current guidance on the costs and benefits of transport schemes exists in the form of the New Approach to Appraisal (NATA) which is currently being reassessed by the DfT in a ‘refresh’ exercise. In quantitative terms, appraisal benefits are dominated by users’ travel time and costs and the evidence-base for the valuation of pedestrians’ and, particularly, cyclists’ time is limited (see, for example, Wardman et al. 2007). NATA includes valuations of carbon and quantification of physical fitness, but these will be reassessed. For example, the validity of the 30 minutes threshold for daily physical activity will be considered, and alternative continuous measures assessed. Valuation evidence exists in NATA for factors such as journey ambience, whilst townscape, accessibility, social inclusion and integration are seen as areas in which valuation should be developed (DfT, 2007: Figure 4.4). Methodologies for including such ‘soft’ composite variables in economic appraisal will be reviewed. Current detailed recommendations for appraising cycling and walking facilities are given in Transport Analysis Guidance (TAG) Unit 3.14.1. This was informed by work by Wardman et al. (op cit.) (drawing on earlier work by Wardman and Hopkinson (1996) and Wardman et al. (1997, 2005) to determine the values cyclists place on journey time and the provision of facilities and hence the extent of user benefits. Similar values of walking are provide by Heuman (2005) based on a study of the Strategic Walk Network in London. Estimate for valuing health benefits are based on work by TfL (2004) and only cover three diseases: coronary heart disease, stroke and colon of the cancer, whilst the use of thresholds rather than a continuous dose-response approach can cause potentially significant underestimation of benefits (SQW, 2006). Benefits from reduced absenteeism through increased walking and cycling are also based on TfL (2004) who in turn use data from the USA on the impact of physical activity on absenteeism (WHO, 2003). Impacts on accidents are measured using the safety in numbers hypothesis that has been confirmed by Jacobsen (2003). Three case studies are presented that indicate large benefit cost ratios of between 11.7 and 22.1.

The key question for the economic work of iConnect is thus: ‘how and why does (or why doesn’t) Connect2 produce economic returns and in what ways are these returns maximized (minimized), for whom and in what circumstances?’ The economic framework for Connect2 will build upon the critical realist framework adopted for the evaluation at the five case study sites. This approach to economic evaluation could be termed ‘realistic economic evaluation’ and is retrospective (ex–post) looking back at the infrastructural interventions in the case studies and examining their economic effectiveness. Economic appraisal, exemplified above, looks ahead (ex-ante) to the ways in which economic effectiveness can be appraised and this is also an important part of the approach intended in this research. The iConnect economic evaluation framework required thus needs to embrace flexibility to satisfy the questions the iConnect team have posed for this research. Therefore, our economic evaluation of Connect2 will encompass a range of policy perspectives including the 'societal' and 'government agency' perspectives. In order to cover all policy perspectives of potential importance, we intend to work with robust economic evaluation research designs including cost-benefit analysis and cost-utility analysis that can stand up to scrutiny in a harsh policy environment and meet the requirements of common-currency resource allocation arguments. We will develop a strategic evaluation framework incorporating the main areas identified above to
include definition of scope, economic evaluation criteria, time horizon, requirements for longitudinal data collection, calculation of net effects and output indicators (such as cost-effectiveness analysis, cost-benefit analysis, improved accessibility of services and social cohesion). We will assemble the necessary data from a combination of primary and secondary data sources. Primary data will be collected using the user intercept questionnaires to enhance Sustrans’ routine data collection which is to be carried out at all Connect2 sites. This may be complemented by more qualitative data from interviews with local stakeholders. Secondary data will be derived from the literature in transport, health and climate-change economics, including government guidance (WebTAG, WHO), research and consultancy reports (e.g. Stockholm Environmental Institute, 2005).

Economic appraisal and evaluation frameworks that pertain across the travel and transport sectors tend to follow a mechanistic, linear, reductionist logic that applies step-by-step rules for the calculation of benefit cost ratios (or similar indicators) connected with comparative ways of making investments. Our iConnect economic evaluation framework will encompass this type of thinking at one level but will also try to go beyond this by ‘rocking the boat without capsizing it’. We propose a decision analytic model that incorporates our knowledge of context, mechanism and outcome and our understanding of the population and how the entire Connect2 infrastructure programme is working, including who it is (and is not) working for, to what extent, and in what circumstances.

5. Conclusions

Connect2 involves a variety of different physical interventions to promote walking and cycling in a wide range of contextual setting. To provide a general evaluation of such a varied investment programme presents a number of research challenges. We will use a purposive sample of heterogeneous case study sites to develop a general theoretical model of how the physical environment influences active travel and a middle range theory that attempts to explain how physical interventions work. We will study context, mechanism and outcome to determine why these interventions are (or are not) effective, in what ways, for whom and in what circumstances. We will pay particular attention to the accurate measurement of behavioural change through the use of longitudinal household surveys and user intercept surveys, supplemented by counts and other observational data. We will build on existing knowledge of the advantages and disadvantages of travel diaries, physical activity questionnaires and emission profiling. A randomised control trial will examine the impact of complementary interventions. A realistic economic evaluation framework will be developed and applied to all Connect2 sites. This will build on existing tools such as NATA and HEAT but will be expanded to consider contexts and mechanisms as well as outcome to answer the question: how and why does (or why doesn’t) Connect2 produce economic returns and in what ways are these returns maximized (minimized), for whom and in what circumstances.
References


Sustrans (2007) Low carbon travel: Reducing the climate change impact of road transport, Information Sheet FF44. Bristol: Sustrans


World Health Organisation (WHO) (2003), Health and development through physical activity and sport.