Advice on selecting indicators for sustainable transport

DistillATE

ADVICE ON SELECTING INDICATORS FOR SUSTAINABLE TRANSPORT

Why read this guide?

What makes a good indicator? Even when the topic for measurement is clear, e.g. congestion, the exact nature of the indicator is more difficult to specify (delay per person?, per vehicle?, over what period?, regular or unexpected delay?). For some policy areas (e.g. health and transport) identifying the right indicators to measure can be even more problematic as the relationships between interventions and outcomes can be weak.

Choosing the wrong indicator can have significant downsides. Apart from being a waste of scare monitoring resources, poor indicator selection can mis-inform decision making, distort priorities and ultimately discredit the policies selected.

This guide describes the key issues surrounding the selection of individual indicators. It presents an audit process which it is suggested is followed prior to the adoption of indicators into a monitoring strategy. It also discusses some of the key issues which surround indicator use such as the difference between scheme and strategy level indicators, statistical robustness and target setting. The guide uses case studies to illustrate the principles in practice.

What is in this guide?

More detail?

This report brings together the findings from two larger in-depth research reports:

→ Marsden, G. and Thanos, S., (2008) Measuring wider economic benefits of transport: A case study in good practice for indicator selection, Deliverable C2, DISTILLATE Project

→ Marsden, G., Kelly, C., Snell, C. and Forrester, J. (2005) Sustainable Transport Indicators: Selection and Use, Deliverable C1, DISTILLATE Project

Both reports can be downloaded from the DISTILLATE website **www.distillate.ac.uk**

1. What makes a good indicator?

1.1 Introduction

Mitchell et al. (1995) state that indicators are needed to make sense of the 'complex systems' that we live in. In particular they identify four main reasons for using indicators to do this, which are:

- They allow the synthesis of masses of data;
- They show the current position in relation to desirable states;
- They demonstrate progress towards goals and objectives; and
- They communicate current status to stakeholders so that effective management decisions can be taken that lead towards the targets.

For example, Traffic 411.com in the US summarises the delays on major routes in cities across the West of the US with circles shaded different colours to reflect delay. The colours act as an indicator of delay and more detailed information is available by clicking on the links for those users affected by the problem.

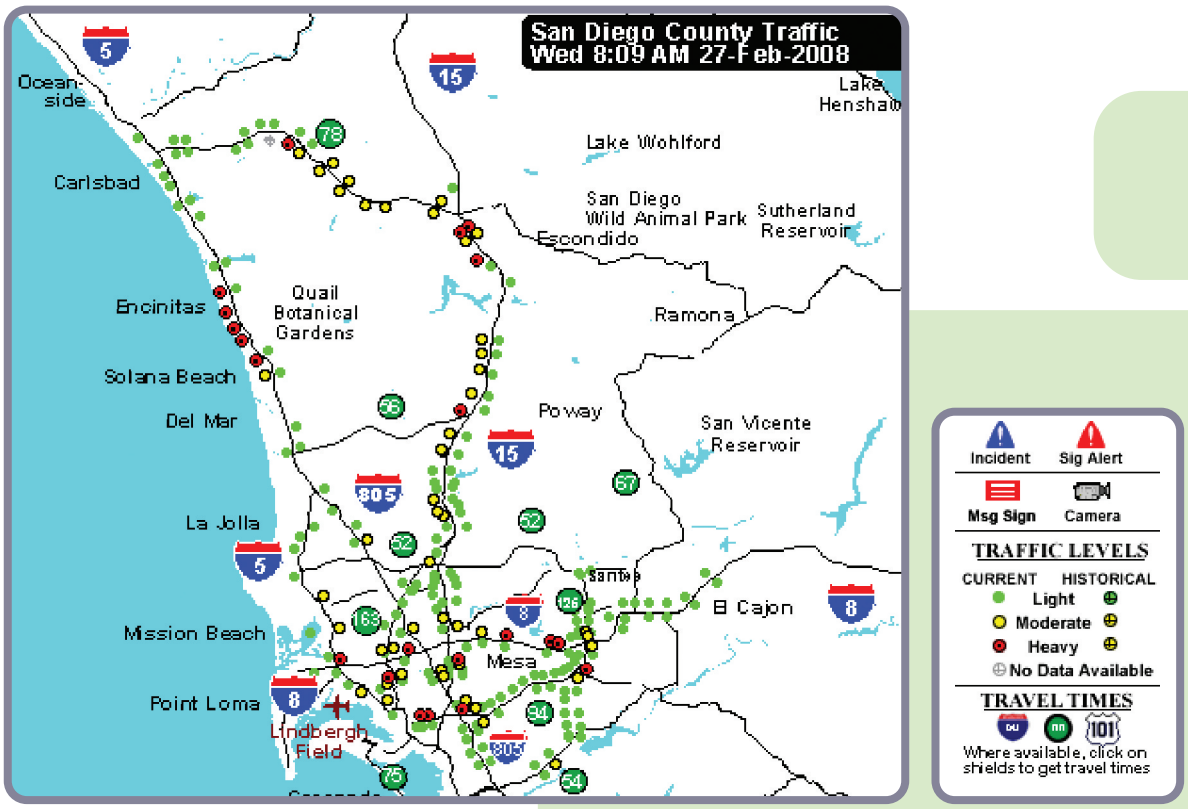

Figure 1: Real-time indicators of delay in San Diego county (courtesy of Iteris, Inc)

The indicators in Figure 1 are designed to communicate specifically with the general public where, at a glance, problems on a route can be identified. Such a scale would however be too coarse to be useful to decide which links are most in need of remedial action to cut congestion. A similar colour coded delay scheme is in operation for traffic information on the UK trunk road network (http://www.trafficengland.com/TCC/). A different indicator is required however for prioritizing sites requiring remedial action to tackle congestion. For these purposes, the Highways Agency has a target to improve reliability, defined as "the average vehicle delay on the 10% slowest journeys" (www.highways.gov.uk).

This guide deals with the issues surrounding the choice of an individual indicator once it is known what role that indicator will perform. So, for example, should the main indicator of congestion be based on all journeys or the slowest 10% of journeys? Should service satisfaction be measured by users only or by users and non-users?

Guidance on how to assemble a package of indicators into a coherent strategic monitoring framework is provided in a companion guide (Designing a monitoring strategy to support sustainable transport goals). That guide also considers the importance of the end use of indicators.

1.2 Types of indicator

There are several different types of indicators that can be measured as part of a monitoring programme. The UK Audit Commission (2000) suggests the following broad categorisation:

- Cost – the money spent to acquire the resources (e.g. transport plan expenditure)
- Input – the resources employed to provide the service (e.g. amount of tarmac laid)
- Output – the service provided to the public (e.g. the number of bus miles run)
- Intermediate outcome – the changes to the transport system that can be observed (e.g. the number of bus users, the number of kilometres travelled)
- Outcome – the actual impact and the value of the service delivery – i.e. achievement of objectives (e.g. delay per person kilometre, fatalities)

These indicator types will be referred to throughout the guide.

1.3 Characteristics of a good indicator

The literature provides many lists that help to define the characteristics that a good or robust performance indicator should include (e.g. see PASTILLE (2002), Carlin (2004), Markless and Streatfield (2001) Audit Commission (2000)). The four examples provided in Table 1 provide a good range of characteristics and indicate that the generic requirements are very similar across a range of disciplines. A short summary of the key characteristics follows to assist with interpretation of the findings.

Table 1: What makes a 'good' indicator?

Characteristics	Audit Commission (2000)	Local Authority Carlin (2004)	Education Fitz-Gibbon (1996)	Sustainability indicators PASTILLE (2002)
Relevant to the organisation/ strategy	✓	✓	✓	
Clearly defined/ easy to understand/ transparent	✓		✓	✓
Based on available data/measurable				✓
Controllable/ attributable	✓	✓		
Cost-effective	✓		✓	
Limited in number		✓		✓
Responsive/timely	✓	✓	✓	✓
Statistically/ scientifically valid	✓		✓	✓
Comparable/ consistent over time	✓	✓		✓

Relevant to the organisation/strategy Whilst this appears obvious, it should be clear to decision-makers and those involved in data collection what the end purpose of the indicator is. Evidence from research on local transport plans suggests that indicators fall into disuse where this is not clear (Marsden et al., 2007).

Clearly defined/easy to understand/transparent This requires the indicators presented to be fully defined. AM peak hour traffic levels would, for example, require the specification of the period of the AM peak hour time period (e.g. 0730-0930), count area (e.g. crossing the inner ring road) and whether in-bound only or two ways across the cordon. An indicator has to be clearly defined to be understandable but being easy to understand goes beyond the definition. Composite indicators for example (e.g. NPV or sustainability indices) fold a lot of information into one measure and can be difficult to interpret. The ability of decision-makers to understand the information provided should be part of the selection process.

Based on available data/measurable This suggests that indicators which can only be estimated through modelling exercises are not suitable as part of a monitoring strategy as there is never any external validation of the model results.

Controllable/attributable This is a critical issue for transport monitoring, particularly in a cross-sectoral policy environment. Indicators such as traffic flow and bus patronage have a long and fairly well established relationship with transport policy interventions such as pricing or supply changes (controllable). The transport impacts of complex interventions such as congestion pricing in London and the Stockholm cordon pricing scheme were predicted with good accuracy (TfL, 2007; Eliasson, 2007), suggesting that the relationship between intervention and outcome is clear (attributable). For other policy areas, the relationship between transport interventions and outcomes is much less clear. For example, improvements in safe routes to school and the implementation of school travel plans have a mixed (but positive) impact on the levels of cycling and walking to school (Cairns and Newson, 2006) but a weak relationship with obesity levels, a key health priority. Until this evidence base is established, it is not sensible to set targets for the extent to which transport will lead to an obesity reduction, but monitoring progress on mode share to school is still a useful contextual indicator (alongside other evidence on diet, exercise etc.). To take another example, the relationship between transport improvements and employment is weak at an aggregate level but scheme level monitoring (for example workplace travel plans) can capture these links. This is discussed further in Section 3 of this report.

Cost-effective The cost effectiveness of indicators should be examined as part of a complete monitoring strategy. The companion guide (Designing a Monitoring Strategy to Support Sustainable Transport Goals) sets out the process for assembling and prioritising indicators.

Limited in number Decision-makers can only process and trade-off so many indicators. The LTP2 guidance suggests a maximum of 40 indicators. It is unlikely that decision-makers can prioritise effectively over such a large number. For example, Nellthorp and Mackie (2000) suggested that most but not all of the 16 NATA indicators were used in decision on the 1998 Roads Review. There are therefore strong arguments for limiting the number selected for top-level strategy monitoring and keeping these focused on key outcomes but selecting further supporting indicators for more detailed decision-support (see companion guide Designing a Monitoring Strategy to Support Sustainable Transport Goals).

Timely monitoring is used for decision-making. It is therefore critical that the information provided is at a frequency which is sufficient for management actions to be taken. This is perhaps most difficult with areas where the time-lags between interventions and changes to outcomes are long. So, whilst an accident black spot intervention should show results over the short-term (1 to 3 years) the development of a cycling network may take a longer time to obtain a critical mass and hence show increases in the numbers of people cycling and subsequently impacting on their health. Careful selection of output, intermediate outcome and outcome indicators is therefore required over a longer period for some interventions.

Statistically/scientifically valid It is important that monitoring programmes are robust and stand up to external challenge. Evidence from the Atkins review of Local Transport Plans suggests some indicators are still open to challenge. For example, two different measurement approaches for cycling were seen to show opposing trends (as shown in Figure 2). The local authority APR method includes counters, twice yearly cordons and questionnaire surveys. SUSTRANS supplemented this with evidence from their electronic count site, route user surveys and school/workplace travel surveys.

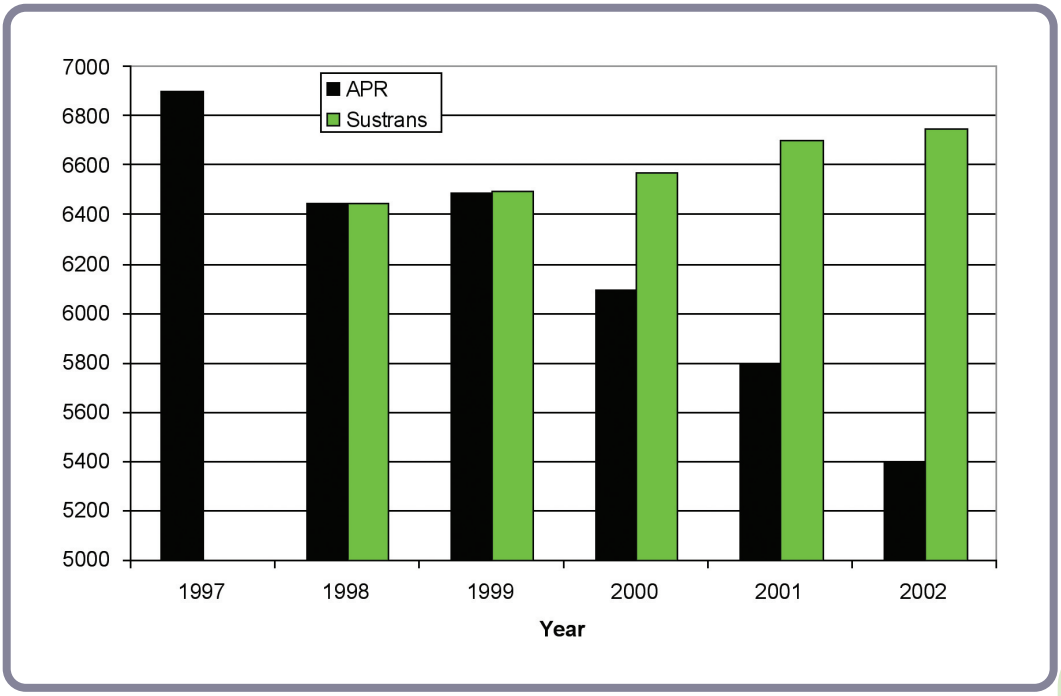

Figure 2: Comparison of a local authority annual performance report and sustrans count of cycle use (based on Atkins et al., 2006)

Issues of how much data to collect and how to be certain over the implications of the data are addressed further in Section 4.

Comparable/consistent over time There is substantial benefit to be had from holding a consistent time series of data against which to judge whether or not there is a change in direction of the data. Changes to measurement equipment, count site location and measurement methods can all disrupt time series data. Whilst external influences such as technological advancements or changes to the road network might make such changes desirable or necessary, these should be minimised. Where possible dual monitoring of data, for an overlapping period, should be conducted to understand differences which arise from the change.

1.4 Criticisms of indicators

It is often suggested that indicators corrupt decision-making processes in a number of ways:

1. By encouraging agencies to modify their measurement approaches to artificially improve performance.

   This has been particularly true of intermediate outcome based measures (e.g. in the health sector where wait times in accident and emergency wards were cut by re-engineering the check in routine). The more outcome focused nature of transport monitoring reduces these risks as does central government guidance on, and audit of, measurement approaches.

2. By artificially focusing efforts on the detail of what is measured.

   This might for example apply to a route-based measure of congestion where management action focuses on improvements to the route that is measured at the expense of competing routes. Whilst such risks exist, it is not the indicator that creates the risk but its application. It may be desirable, from a public communication perspective, to provide information on key routes rather than aggregate city wide delay figures. Where such indicators are fully embedded in a performance management monitoring strategy then their role should be seen as part of the bigger picture (see Designing a Monitoring Strategy to Support Sustainable Transport Goals).

3. By focusing efforts on what is measured to the exclusion of what is not.

   This is a potential criticism of any performance monitoring framework. It it is not a criticism of individual indicators but rather whether the framework is comprehensive in its coverage of policy objectives. The main concern from a transport perspective is that many of the cross-sectoral policy interventions of which transport contributes to, such as obesity, do not have a strong evidence base of effectiveness upon which to select indicators. There is a risk that such policies are then excluded due to uncertainties over future performance levels. A companion guide (Monitoring Across Sectors and Spatial Levels for Sustainable Transport: A Good Practice Guide) provides some evidence to support the development of better cross-sectoral monitoring. The application of these types of indicators is discussed further in Section 3 of this report at a scheme level.

In summary, there is always scope for strategic game playing with indicators. In the context of local transport planning in England, these risks appear to be limited in nature due to the focus on outcomes, the degree of central government guidance and the limited financial gains to be made by game playing with individual measures. Those risks that do exist can be designed out through careful assembly of a monitoring framework. This suggests that scrutiny of the monitoring framework is important. The next section of this report defines an audit process for new indicators which encompasses the key elements of best practice described in Section 1.3.

2. An audit process for new indicators

2.1 Background

There is a tradition of monitoring transport patterns within the UK. There are therefore well established methods for collecting traffic speed, flow and accident data. In the past couple of decades air quality monitoring has become more important and other issues such as noise and local economic performance are now emerging. There is therefore pressure for an evolution of monitoring to keep pace with the changing understanding of the role of transport in supporting sustainable development. A snapshot of local authority views of the degree of importance and satisfaction with a range of transport indicators from the first DISTILLATE survey is shown in Figure 3.

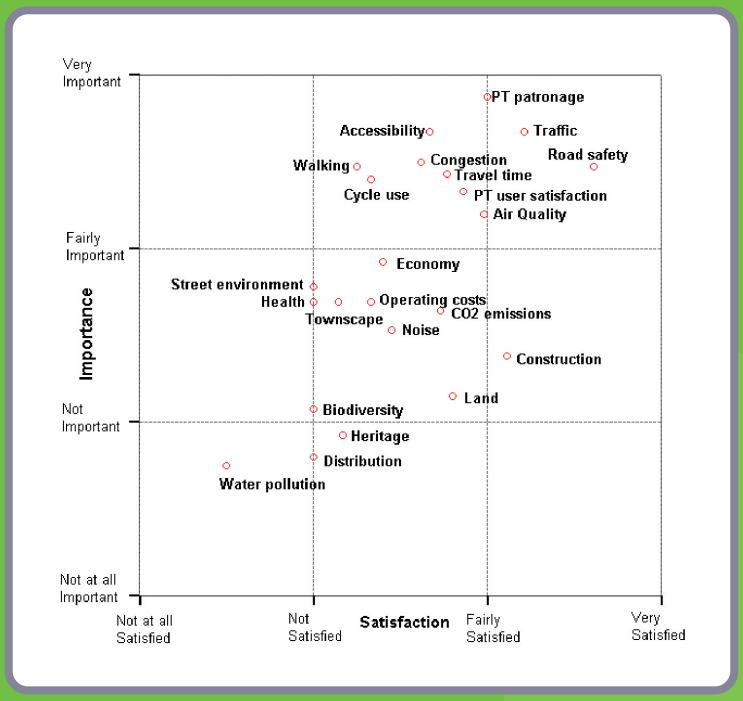

Figure 3: Importance and satisfaction with indicators (2004)

There appears to be a range of indicators which relate to sustainable development which are moving up the policy agenda but which are currently difficult to monitor (such as economic impacts, health and street environment). The principles of good practice for indicator selection set out in Section 1.3 have been reviewed and a six criteria audit process has been established to assist with the selection of new indicators. This is set out below and a case study application presented in Section 2.3.

2.2 Six criteria audit

The criteria of a good indicator reviewed in Section 1.2 have been condensed into a six criteria audit process shown in Table 2.

Table 2: Audit process for new indicators

Criteria	Comments	Yes	No	?
Clearly defined	The details of exactly what is to be measured are clear.			
Controllable	The impacts of transport policy interventions are likely to be significant on the indicator and unlikely to be dwarfed by changes in external factors.			
Measurable	If the indicator cannot be directly measured then it should be rejected in favour of a proxy indicator or outright.			
Responsive	If an indicator is unlikely to respond in the short-term to policy changes then meaningful collection periods should be considered or the indicator rejected. Long delays in between monitoring can lead to other external factors changing.			
Easy to understand	The indicator should be examined to ensure that it is presenting information in a format suitable for the target audience. High degrees of aggregation of information can reduce the comprehensibility of an indicator and increase the risk of double counting of 'hidden' elements of that indicator.			
Cost-effective	The benefits of collecting the data are sufficiently high to justify the cost of collection relative to alternative solutions.			

Once the audit process has been completed, an evaluation of the responses should be conducted. Given the variety of indicators that could be brought forward and the range of issues which may be associated with any indicator in a given set of local circumstances we do not proffer guidance on the right balance between yes, no and do not know (?) responses which make an indicator worth selecting. The audit process of itself should however be sufficient to alert monitoring and policy officers to the difficulties of the indicators selected. The audit process is applied to a Case Study in Section 2.3 on page 11.

Aspects of good indicators that have been omitted are relevance to the organisation, limited in number, statistically valid and consistent over time. The first two should be dealt with through the application of a systematic process of indicator selection (such as that set out in Designing a Monitoring Strategy to Support Sustainable Transport Goals), i.e. individual indicators should only be subject to an audit of suitability when there is a clearly established need for them. The other two should be borne in mind in applying the selected indicator. Provided the measurement techniques are robust and a sufficiently large sample of measurements is taken statistical validity can be achieved. This is more therefore a question of whether such sampling is cost effective. The issue of statistical validity is addressed further in Section 4.1. It is not possible at indicator selection stage to review whether or not measurement approaches will be consistent over time. This is always an aim but it should not prevent advances in measurement techniques.

2.3 Case Study: Indicators to measure productivity

2.3.1 Definitions

The wider economic effect of transportation is an important policy issue, to which attention has been given, especially since the SACTRA Report (DETR, 1999) recommendations. More recently, the UK Department for Transport has introduced new guidance to estimate additional productivity benefits accruing from transport investments which are over and above those captured by travel time savings. Productivity is often defined as the amount of output per unit of input achieved by a firm, industry, or country.

The underlying theoretical basis for productivity gains is complex as is the basis for calculating adjustments (see DfT, 2006). The DfT guidance suggests that productivity gains should be estimated for major infrastructure investments. This section examines productivity as an emerging policy field and asks whether changes to productivity could be monitored over time as part of a Local Transport Plan rather than simply for individual schemes. The case study description is based on a more detailed technical analysis available on the DISTILLATE web site (Marsden and Thanos, 2008).

Four main types of wider economic benefits relating to productivity are discussed by the Department for Transport.

WB1: Agglomeration economies

Agglomeration economies occur as larger markets allow wider choice and a greater range of specialist services. So, larger cities have clusters of specialist industries which bring associated benefits such as larger pools of skilled labour, knowledge transfer and support services. However, these gains are limited as further agglomeration gives rise to diseconomies due to congestion and pollution.

It is possible to distinguish the importance of transport in large concentrations of population, with congestion being a reason for "agglomeration diseconomies" arising. It seems that "agglomeration economies" is one of the links in the "transport-productivity" relationship.

WB2: Increased competition as a result of better transport

DfT (2006) do not normally expect to find significant wider benefits owing to increased competition as these are normally well functioning markets or competition authorities are in place to curb market powers. This is not addressed further here.

WB3: Increased output in imperfectly-competitive markets

Transport appraisal captures benefits to firms by estimating the time savings for travel undertaken in the course of work. Firms will respond to such cost savings by reducing prices and increasing output. Where there is imperfect competition in a market, the value placed on an additional unit of production (price) is normally higher than its (marginal) production cost. Firms and consumers would therefore be jointly better off if firms were to increase production. If better transport induces firms to increase production, there are precisely such benefits – the value placed on the additional production is higher than the cost of producing it. Since these second round benefits would not fall to the firms that receive the transport benefits, the value attached to time savings would underestimate the true benefits (DfT, 2006).

Conventional transport appraisal understates the transport benefits, by an "up-rate factor" to the direct cost savings to firms, ie business time savings and reliability gains. However, as this would simply be a multiplication of an already calculated journey time saving this does not suggest that any new monitoring is required.

WB4: Economic welfare benefits arising from improved labour supply
Commuting costs (including the cost of time and the inconvenience of overcrowding etc) are one of the effects that may limit how much, and how many, people work. These changes are not fully captured by an individual's time-cost trade off as individuals work on post-tax income so the impacts on the exchequer can be underestimated. Reducing costs of commuting can bring wider economic benefits and these benefits can be split into three categories according to DfT (2006):

- More people choosing to work due to changes in effective wages (GP1)
- Working longer hours in current job (GP2)
- Working in more productive jobs (GP3)

2.3.2 Indicators for assessment

It seems therefore that the following two indicators are the only ones appropriate to be taken forward in this analysis:

- Generalised costs as an input to agglomeration economies from transport improvements (WB1)
- Generalised costs and commuting flows as an input to new entrants in the labour market (GP1, part of WB4)

The Equations which relate the measures listed above to the wider economic benefits are shown in Marsden and Thanos (2008).

2.3.3 Audit

Although both apparently complex relationships, the key transport change which underpins both measures is the change in zone-zone generalised costs and, for GP1 commuter flows. The Indicator audits are show in Tables 3 and 4.

Table 3: Indicator audit WB1

Indicator	Generalised cost as an input to WB1 (agglomeration economies)			
Criteria	**Comments**	**Yes**	**No**	**?**
Clearly defined	Generalised cost is the major component for calculating WB1. This is a well defined concept combining out of pocket costs with the monetised value of travel time costs. The current advice on wider economic benefits is to use ward to ward generalised cost estimates although the time periods are not clearly specified. Wards are well defined units so this also seems feasible. Graham (2005) used daily averaged generalised costs, provided by DfT. Guidance on the calculation of generalised costs is given in DfT(2005b) and TAG (2007). The other elements (GDP, employment) are sufficiently defined.	✓		
Controllable	External effects in GDP can be controlled for; external effects in the employment of the area may be an issue if they are significant (but then they can be identified as an error source). However, it is expected that generalised cost will produce the most significant changes in WB1. Transport policies are one of the dominant influences on zone to zone generalised costs.	✓		
Measurable	All the elements of WB1 should be available from DfT and ONS, except generalised cost. Ward centroid to ward centroid generalised costs are typically estimated through transport models rather than measured. Measuring centroid to centroid journey times requires a series of journeys to follow that type of journey which is difficult to establish. There is currently no measured baseline data that is fit for purpose. The growth in satellite tracking devices may make this data easier to collect and process should it be seen as critical.	✓		
Responsive	Very substantial changes in zone-zone generalised costs (such as those generated by major transport investment) are likely to be required to see significant productivity impacts. Annual monitoring would therefore be of little use although a less frequent approach (e.g. five yearly) might be appropriate.		✓	
Easy to understand	The element of the generalised cost in WB1 is relatively easy to work with from a technocratic perspective. It has disadvantages in being some way removed from the actual productivity impacts which could be calculated. It is also a very aggregate measure which may mask a mix of gains and losses to different modes. The impact of changes in generalised cost will be on a monetary measure (change in GDP) that is fairly easy to understand and to present to politicians and may be useful for communication with the public. However, understanding the whole procedure of calculating the agglomeration effects is much more difficult, especially for those without a strong background in economics. A degree of aggregation of information may also reduce comprehensibility, but the risk of double counting is very small.			✓
Cost-effective	A bespoke survey of zone-centroid to zone-centroid journey times would be prohibitively expensive given the uncertainty over the responsiveness of the indicator to most LTP interventions. The Department for Transport is making available increasing amounts of data from vehicles equipped with satellite tracking systems to local authorities. The data could be reinterpreted to approximate zone-centroid to zone-centroid journey times at relatively low cost. The use of such data should be piloted however before it can be concluded to be robust enough for tracking change over a five year period across a large area. The data collection costs are currently prohibitive although new satellite tracking measurements may remove this barrier.			✓

Table 4: Indicator audit WB4 GP1

Indicator	Generalised cost and commuter flows as an input to GP1			
Criteria	**Comments**	**Yes**	**No**	**?**
Clearly defined	Generalised cost is discussed above (Table 3) and generally found to not be problematic subject to certain caveats. Commuting flows from zone to zone are available through the census and are clearly defined.	✓		
Controllable	There are some elements in GP1 that may not be the result of transport changes. For example the change in the number of commuters for a given area is supposed to be as a result of a transport project and there are other external factors outside of LTP interventions which will influence commute destination choice (e.g. new developments and fuel prices) but these are likely to be small.		✓	
Measurable	DfT (2006) provides information on where, the elements needed to calculate GP1 are available. Guidelines for calculating the generalised cost of commuting are given in DfT (2005b) and TAG (2007). Commuting flows are currently only measured every 10 years.	✓		
Responsive	Very substantial changes in commuting generalised costs are likely to be required to see significant impacts on labour supply as discussed above. More needs to be understood about the general underlying churn in the labour market.		✓	
Easy to understand	GP1 is essentially the extra tax from more people choosing to work due to changes in effective wages by reduction of transport costs. Whilst generalised costs and commuting flows are easy to understand the translation of these benefits into a tax figure is more difficult and, as the tax revenues do not flow to the local authorities, this would be of limited value in LTP2 monitoring.		✓	
Cost-effective	The same difficulties exist for generalised cost as described above. The tracking of commuting flows on a more frequent basis than the census would be expensive relative to the likely benefits.			✓

2.3.4 Audit results

Neither of the indicators suggested appears to make a strong case for selection. Whilst zone to zone generalised cost appears to have some promise as an indicator, further evidence would be required on the scale of change in generalised costs that needs to occur for there to be substantial productivity gains to be established. If these are only likely to emerge as the result of major schemes then it would not make sense to monitor this as part of local transport planning processes. Substantial investment in data collection methods using satellite tracking would be required to facilitate this type of data collection. At a local level it may be sufficient and preferable to consider trends in travel costs and commuting flows on a 10 year timescale alongside the census rather than as part of an annual policy review process.

This work shows the value of understanding the relationships between the intermediate transport outcomes that can be measured (e.g. generalised cost) and the end outcomes that these are expected to influence (e.g. productivity). In this instance the evidence is still comparatively new and what is available suggests that most local transport initiatives will have very limited impact on productivity. Further understanding will need to be developed, probably through major scheme development examples, before the added value of monitoring any related indicators could be assured. It is unrealistic to expect individual local authorities to take on the costs of developing such new indicators and to experience the risk of using them in assessment frameworks. There is a strong role here for central government in facilitating the uptake of any indicators in this area.

3. Scheme and strategy assessment

3.1 Strategy assessment

The companion guide "Designing a Monitoring Strategy to Support Sustainable Transport Goals" provides a step-wise guide to the design of a cost-effective monitoring system to support an integrated transport strategy. The guidance covers the key issues of:

- Identifying objectives for the transport strategy
- Linking outcome indicators to these objectives
- Connecting changes in the use of the transport system to changes in the outcomes
- Connecting policy interventions (inputs and outputs) to the changes in the transport system or the outcomes

This process develops a causal chain approach to understanding how the policy interventions taken ultimately impact on the key outcomes. Designed well, the framework should provide feedback that is useful to assessing both whether the strategy is working and which elements of the strategy are working well and which may need to be reviewed, i.e. it performs an important decision-support role.

Whilst such an approach should pick up major changes in the system, aggregation of data to a city or sub-regional level (as is the case in metropolitan areas) can hide important variations in behaviour across areas and between different groups. It is also difficult to see the link between some smaller interventions and conurbation or county wide changes in outcomes. For example, conurbation-wide reporting of changes in bus patronage hides the location and nature of increases and declines in particular routes. Where major interventions have been put in place, such as new quality corridor arrangements, then more scheme specific monitoring is desirable. Some indicators are derived from a more site specific measurement and aggregation (e.g. road traffic accidents) and scheme level detail will thus be available to decision-makers.

It can also be difficult to trace the direct relationship between transport interventions and some outcomes, particularly those such as health and employment which are subject to other non-transport impacts (see the discussion on controllability in Section 2.2). For others (e.g. impacts on employment levels) there is a lack of connectivity between transport interventions and outcomes. So, for example, if public transport accessibility to a major employment site is improved yet jobs are lost elsewhere in the city, it will be difficult to chart the success of the transport intervention against the broader picture of job losses. Here, scheme specific monitoring is required.

The Atkins review of LTP2 monitoring recommended that "Authorities should be encouraged to undertake, and disseminate the results of, scheme specific before and after surveys to provide more evidence of cause and effect at the local level and overcome the fact that current aggregate data may not be the most suitable for informing strategy development." (Atkins et al., 2006, pv)

3.2. Scheme assessment

Scheme assessment is desirable for major interventions and policy initiatives, particularly where previous local evidence on effectiveness is absent. Whilst knowledge bases such as KonSULT (http://www.konsult.leeds.ac.uk/) can provide an indication of the likely impacts of a range of transport interventions, these can vary substantially with local context.

Scheme assessment should be related to the main strategy monitoring framework but should be tailored to the characteristics of the individual scheme. Some examples of strategy-scheme relationships are highlighted below.

Example 1: Increasing employment
A key outcome indicator for a strategy might be identified as "increasing employment" for a city. The main intermediate outcome indicators that might link to this through the strategy monitoring framework might be "reducing congestion" and "public transport accessibility to major sites of employment". As suggested previously, improvements to scheduled public transport services might help to facilitate increased uptake of employment, other site or area specific initiatives might also be adopted.

One employment site may for example offer a works bus scheme which would not feature in the scheduled bus service offering. Monitoring of the effectiveness of that scheme would look at the vacancy rates at the site, retention of staff and previous employment status of new starters. This would not be practical or relevant at a city wide level but there would be a direct connection between the more detailed measurement and the outcome of increasing employment.

Another area might consider developing a "Wheels to Work" scheme where job seekers are given free or subsidised hire of scooters or cars for a limited period as they get back into employment. This has been demonstrated to be successful with many people remaining in employment but potentially carrying on travelling by scooter, car or liftsharing with new colleagues. Few transfer to public transport. Here, scheme monitoring of the % of participants remaining in employment would be the main measure. Again, this is not city wide but it does feed straight into the outcome measure of increasing employment.

Example 2: Bus quality corridor
Increasing bus patronage is a key government target and a goal for local authorities. Ultimately increased bus patronage can help to fulfil a series of outcomes such as improved social inclusion and reduced congestion and pollution and is better considered in this context. One of the key policies for local authorities and bus operators over recent years has been the introduction of bus quality partnerships which often revolve around particular corridors. In such partnerships the local authority upgrades the bus stop infrastructure and provides dedicated bus lanes and bus priority (to varying degrees) whilst the operators invest in a better quality fleet for the route and sometimes more frequent services and better marketing.

Bus companies are unwilling, for commercial reasons, to provide route specific information to local authorities for monitoring purposes. It is therefore an important element of understanding the aggregate bus patronage figures to look at the role of new quality corridors in the overall patronage figures. There are also potential knock on effects to other competing bus routes and to the car traffic in the area (with reallocation of road space and new public transport alternatives). Scheme monitoring here is critical to identify the real net improvements to all travellers (rather than the bus routes in question) and to justify continued expenditure on other corridors. The outcomes of the scheme monitoring would connect directly into strategy monitoring of bus patronage, overall traffic delay and bus user satisfaction and provide an extra layer of detail for decision-makers.

3.3 Summary

Aggregate strategy-wide monitoring should be augmented by scheme specific before and after monitoring. Such monitoring will not only provide a better understanding of the detail of the change in area-wide indicators but will also potentially challenge the assumptions about the degree and/or direction of relationships between interventions and indicators. So, where scheme specific monitoring fails to demonstrate impacts on key outcomes then this can be used to help refine and reduce the macro monitoring framework. Similarly, scheme specific monitoring may act as the catalyst to establish that there is a relationship between an intervention and an outcome that it is worth monitoring.

It is acknowledged that funding for ex-ante and ex-post scheme monitoring is often limited or not available. Where possible it should be built into scheme costs and/or funded through associated developer contributions.

4. Issues in monitoring

This section provides a short review of some key issues raised by practitioners involved in DISTILLATE which continue to affect their data collection and management processes. These have not been major areas of inquiry for the project and the purpose of this section is to identify where guidance and good practice exists for those facing these issues.

4.1 Statistical robustness

There is an increasing concern about the statistical robustness of data and of reported changes in data from one data collection period to the next. When is a change in data a real difference and when it is a minor variation?

Texts on statistical robustness can stretch to hundreds of pages. A series of important and accessible resources have therefore been identified to assist those looking for guidance on how to develop and assess effective monitoring strategies.

The Department for Transport commissioned specific research on monitoring and reporting of LTP outcomes. The report (Atkins et al., 2006) specifically addresses the core transport indicators that local authorities have to measure and examines issues such as assessing progress over time, error margins and the impact of monitoring approaches. It is available on-line at **http://www.dft.gov.uk/pgr/regional/ltp/research/ltpoutcomes** free of charge.

Survey Methods for Transport Planning (Richardson, Ampt and Meyburg, 1995) provides a clear and authoritative guide to the planning, design and administration of a range of different types of survey in transport. It is available on-line at **http://www.transportsurveymethods.com.au/Overview.html** free of charge.

Understanding Traffic Systems: Data, Analysis and Presentation (Taylor, Bonsall and Young, 2000) focuses on traffic data analysis and provides a useful complement to the more survey based focus of Richardson et al.'s work.

The Research Methods Knowledge Base also provides a fairly accessible but more generic description of approaches to sampling, data collection methods, survey design and analysis. It is helpful in providing accessible definitions of terms such as population, sample and confidence interval. It is available free of charge at **http://www.socialresearchmethods.net/kb/index.php**

Statistical Software products such as SPSS and Statsoft offer on-line guidance on the analysis procedures and the underlying rationale for different statistical tests. For example, time-series analysis is appropriate for the study of sequences of measurements that follow non-random patterns (e.g. traffic flow) and guidance is available at **http://www.statsoft.com/textbook/stathome.html?sttimser.html&1**.

All of the survey design and analysis sources listed are very clear in the need to identify the exact nature of the data to be collected at a very early stage. Issues such as whether the population to be studied is bus users or all travellers should be easy to establish. Better planning of what is to be collected, why and how it is to be analysed will ensure better value for money from the resources invested.

4.2 Target setting

Targets set out the level of performance that an organisation aims to achieve for a particular activity within a given timeframe. This might be for example a commitment to reduce fatalities on urban roads by 10% over the next five years. DISTILLATE's local authority partners identified target setting as a difficult process, partly due to this being a new approach to managing their performance. A parallel research project has been studying the process of target setting and the impacts that a system of target setting and performance rewards has on the performance of local authorities. Full details of the project can be found at **http://www.its.leeds.ac.uk/projects/opt_incentives/**

This section provides guidance on how to set targets. It is based on 30 interviews in six local authorities and a questionnaire survey of all LTP submitting authorities (38% response rate) reported in Kelly et al., 2006 and 2008.

The questionnaire identified a number of techniques used to develop targets. Good practice approaches identified included:

- Stakeholder consultation – where targets were jointly agreed with other partners such as bus companies or health service partners
- Regional forums were held to discuss targets – to get a better understanding of the underlying trends which were likely to affect progress
- Benchmarking – where performance was compared on a voluntary basis with 'peer' authorities either through benchmarking meetings or the LTPN catalogue of performance measures
- Targets were based on a statistical analysis of trends/ previous performance – although changes to what is measured or how it is measured can disrupt this. For example, measuring accessibility is a new concept for some authorities which made target setting difficult.

Models were sometimes used to support the target setting process. Modelling results were used in quite different ways however. Good practice involved using the model results as a basis to discuss targets with other stakeholders and to examine the assumptions underpinning the forecasts to determine how likely different target levels were to be achieved. In other instances it was observed that the modelling results were more influential and that assumptions were unchallenged due to a lack of in-house expertise in their interpretation. This potentially inflates the risk of non-achievement.

One of the main difficulties expressed by authorities was in interpreting the Department for Transport's guidance that the levels of target attainment should be "challenging but realistic". The questionnaire survey studies the extent to which authorities felt they were likely to achieve their targets and the results are shown in Table 5. The table is split into three columns for each indicator. The first shows the chance of meeting the target for all responses. The second shows that sub-set of responses from those authorities deemed to have set satisfactory but not stretching targets as defined by the Department for Transport's thresholds. The third column shows the data for the sub-set of responses who were deemed to set stretching targets.

Table 5: Average reported likelihood of meeting targets (source: Kelly et al., 2008)

Indicator	Average % chance of meeting target		
	All targets	Satisfactory targets	Stretching targets
Killed and seriously injured	79.2	80.4	78.2
Child KSI	75.8	76.2	74.6
Slight accident rate	76.3	73.9	78.4
Bus satisfaction	72.6	71.9	77.5
Bus patronage	72.2	63.0	76.8
Bus punctuality	70.2	68.3	71.7
Unclassified road condition	74.1	72.3	79.3
Footway condition	74.1	69.3	80.0

It appears that a 70 to 80% likelihood of meeting a target is deemed to be the appropriate level of challenge for local authorities to adopt, on average. Interestingly, indicators for which authorities choose to set stretching levels of achievement typically have a higher associated chance of being achieved which suggests that authorities are pushing themselves in areas in which they know they have a strong delivery potential.

The research also examined the degree to which the local authorities' view of the difficulty of achieving the targets set tallied with that of the Department for Transport. On average the assessments matched well although there were some significant differences when unitary and shire authorities were analysed separately. This points to the need for further and continuing dialogue about the definition of stretching and satisfactory targets. The research concluded that it may be more effective to compare performance amongst a small group of peer authorities for some indicators than to attempt to define national standards.

The overall conclusions of the work suggest that the process of target setting and performance rewards does incentivise greater levels of outcome achievement than setting targets alone. Two important aspects should be considered by authorities:

1. Effort is focused primarily on the targets selected. It is therefore critical that these cover the full range of transport objectives and that, where possible, these are focused on outcomes and not processes to avoid effort adversely skewing priorities and decisions.
2. Being rewarded for setting stretching targets encourages more ambitious target setting and ultimately higher levels of outcomes. Although the outcomes are higher than they might otherwise have been, the targets are not necessarily achieved. This needs to be communicated to politicians in the plan development stage.

4.3 Customer satisfaction indicators

Feedback from local authority partners in DISTILLATE highlighted a concern that outcomes relating to satisfaction with transport network management and service provision were subject to strong external factors which could adversely affect the short-term satisfaction of users, whilst not being indicative of the trend. For example, short-staffing in a bus company or a set of roadworks in the city centre might adversely impact on bus reliability in the period immediately preceding an opinion survey. The corresponding satisfaction ratings might not then be representative of the longer term trends.

Concerns have also been raised about the degree to which users and non-users of transport systems are included in surveys. For example, BVPI104 requires both users and non-users of public transport to be surveyed for satisfaction.

Customer satisfaction measures need to be approached in the same manner as other performance measures. The issues raised above largely relate to the data collection and interpretation process rather than whether or not the data should be collected. For example, if the process of surveying customer satisfaction is undertaken on a rolling basis at several points in the year then the risks of individual events impacting the results are reduced.

Several useful guidance documents on performance measures for customer satisfaction are listed below. The key messages from these guides are:

1. It is important to survey customers (service users) and potential customers (current non-users) to provide information on differences between these groups. Satisfaction of customers is the preferred measure of service delivery quality but would provide no information on how to encourage non-users to change.
2. Segmentation of the sample is important to understand the responses of different sub-groups of respondents (e.g. ethnic minority groups may have a different level of satisfaction with public transport information provision).
3. It is helpful to determine both the importance of a service attribute and the levels of satisfaction with the service. This prevents pre-conceptions of the importance of service attributes from driving decisions.
4. Descriptions of experiences using the service can help with interpretation of results and developing improvements.

The Department of Communities and Local Government has produced guidance on user satisfaction performance indicators jointly with the Audit Commission. This guidance considers specifically the case of how to interpret bus user satisfaction data.
http://www.communities.gov.uk/localgovernment/localregional/servicedelivery/usersatisfaction/

The Local Government Association jointly with the National Consumer Council has conducted a research programme on customer satisfaction with local services. Although generic in nature, the guide provides a short and accessible explanation of best practice approaches to measuring customer satisfaction and provides case study examples.
http://www.lga.gov.uk/lga/aio/22563

The US Transportation Research Board has produced a research review (TCRP Report 47) of how to measure transit system service quality including more detailed discussion about data collection processes, sampling and bias. Case studies are provided of different performance indicators and different analysis techniques as well as how the results were used in practice.
http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rpt_47-a.pdf

4.4 Summary

There are several issues which needs to be understood and addressed prior to the selection and application of indicators. It is critical that the difficulties that inevitably exist with most indicators are understood and communicated to those using the information. This guide has highlighted the most important factors and provided guidance on making better indicator selections. It should be used in parallel with the companion guide on developing monitoring strategies if a coherent authority wide approach is to be adopted.

References

Atkins, PwC and Warwick Business School (2006) Monitoring and Reporting of LTP outcomes, February 2006 **http://www.dft.gov.uk/pgr/regional/ltp/research/ltpoutcomes**

Audit-Commission (2000). On target: The Practice of Performance Indicators. Management Paper, Audit Commission. **http://www.audit-commission.gov.uk**

Cairns, S. and Newson, C. (2006) Making School Travel Plans Work: Effects, Benefits and Success Factors at English schools, In Marsden, G. (Ed.) Wasted Miles, Wasted Money: A less congested, more energy efficient future, CICC Publications

Carlin T (2004) Output based management and the management of performance: insights from the Victoria Experience. Management Accounting research 15 pp267-283

DETR (1999); 'Transport and the Economy', SACTRA (Standing Advisory Committee on Trunk Road Assessment)

Department for Transport (DfT) 2006; Transport Appraisal, Wider Economic Benefits and Impacts on GDP (July 2006), Department for Transport, London

Eliasson, J. (2007) The Stockholm Congestion Charging System – A summary of the effects, World Conference on Transport Research, San Francisco, 24-28 June 2007

Fitz-Gibbon C (1996) Monitoring Education Indicators, Quality and Effectiveness, London: Cassell

Kelly, C., Marsden, G. and Nellthorp, J. and (2008) An assessment of target setting in local transport plans in England, *in preparation, Institute for Transport Studies, Leeds*

Kelly, C., Nellthorp, J. and Marsden, G. (2006) *How do local authorities set their LTP targets?* Proc. Transport Practitioners Annual Meeting, Manchester, 25-26 July

Markless S and Streatfield D (2001) Developing performance and impact indicators and targets in public and education libraries. *International journal of information management*

Marsden, G., Kelly, C., Snell, C. and Forrester, J. (2005) Sustainable Transport Indicators: Selection and Use, Deliverable C1, DISTILLATE Project, **www.distillate.ac.uk**

Marsden, G., Kelly, C., Hull, A.D., Tricker, R., Lucas, K., Brookes, M., Snell, C. and Forrester, J. (2007) Improving Monitoring and Reporting for Local Authorities: Lessons from the Transport Sector, Deliverable C3, DISTILLATE Project, **www.distillate.ac.uk**

Marsden, G. and Thanos, S. (2008) Measuring wider economic benefits of transport: A case study in good practice for indicator selection, Deliverable C2, DISTILLATE project, **www.distillate.ac.uk**

Nellthorp, J. and Mackie, P.J. (2000) The UK Roads Review – A Hedonic Model of Decision Making, *Transport Policy,* **7**, 127-138

PASTILLE (2002) Indicators into action: local sustainability indicator sets in their context 5th framework, London School of Economics

Taylor, M.A.P., Bonsall, P.W. and Young, W. (2000) Understanding Traffic Systems: Data Analysis and Presentation, Second Edition, Ashgate, ISBN: 0 7546 1248 1

TfL (2007) Central London Congestion Charging: Impacts Monitoring – Fifth Annual Report, Transport for London, **http://www.tfl.gov.uk/assets/downloads/fifth-annual-impacts-monitoring-report-2007-07-07.pdf**

This guide has been written by:

Dr Greg Marsden, Institute for Transport Studies, University of Leeds

Key contributors to the technical work are:
Michael Brookes, University of Westminster
Angela Hull, Heriot-Watt University
Charlotte Kelly, Institute for Transport Studies, University of Leeds
Karen Lucas, University of Westminster
Carolyn Snell, University of York
Reggie Tricker, formerly of University of West of England

This work has been conducted as part of the DISTILLATE project funded as part of the EPSRC's Sustainable Urban Environment programme (Contract GR/S90829/01). Full details of the range of activities conducted through the DISTILLATE project are available at **www.distillate.ac.uk**.

The DISTILLATE consortium comprises:

- Institute for Transport Studies, University of Leeds
- Heriot-Watt University
- Stockholm Environment Institute, University of York
- TRL
- University College London

The DISTILLATE research programme has been conducted with a range of local, sub-regional and regional government partners. We are grateful for their collaboration and co-operation in developing and testing the research products and to our steering group.

UNIVERSITY OF LEEDS

Dr Greg Marsden
Institute for Transport Studies
University of Leeds, Leeds LS2 9JT, UK
t +44 (0)113 343 5358
f +44 (0)113 343 5334
e G.R.Marsden@its.leeds.ac.uk
w www.its.leeds.ac.uk