International Transport Forum

TRANSPORT FOR A GLOBAL ECONOMY

Challenges and Opportunities in the Downturn

FORUM 2009 • 26-29 May • Leipzig

WORKSHOP 1

Intermodal Transport & Supply Chains

Improving Reliability on Surface Transport Networks

BACKGROUND DOCUMENT

OECD

WORKING GROUP ON SURFACE TRANSPORT NETWORKS:
IMPROVING RELIABILITY AND LEVELS OF SERVICE

Preliminary Key Messages and Executive Summary

Reliability can be better integrated into transport policy making

The objective of this report is to provide policy makers with a framework for understanding reliability issues and for designing reliability management policies. The report has made significant progress in identifying methodology for incorporating improvements in reliability into project and policy evaluation, while exploring the pitfalls that need to be avoided.

At present, network and service reliability is not systematically incorporated in the transport planning process and thus is not reflected adequately in decision making. Reliability is rarely factored into cost-benefit analysis, the core planning tool for surface transport networks.

Increasingly complex scheduling places importance on reliability

Technological advances and investments in infrastructure have lowered transport costs and increased average transport speeds. This has facilitated and complemented product specialisation. Supply chains are, more than ever, on a global scale, underpinned by global and, often, just-in-time production and distribution systems. This complexity is echoed in passenger movements, both for business and social purposes. These changing patterns have increased the importance of schedules – and of keeping to those schedules. This increases the focus on transport reliability.

These changes have occurred while traffic levels have risen, and, without compensating infrastructure expansion, congestion has grown, undermining reliability. Individuals, companies and infrastructure managers affected by changing reliability can respond in a number of ways; individuals build extra (buffer) time into their journeys to allow for the possibility of delay, companies adapt their pattern and timing of operations, while infrastructure managers often provide traffic flow information to reduce the impact of unreliability.

Reliability improvements can be delivered by both users and network providers. It should not be presumed that the infrastructure (or service) provider/ government always has to be <u>the</u> source of reliability enhancements. The "low-hanging fruit" of cost-effective reliability improvements may come from network users.

Four main instruments to optimise reliability on transport networks

A wide range of instruments is available to manage reliability. The policy framework proposed in this study distils these into four principal options:

- Increasing physical capacity of infrastructure either through supplying extra capacity and/or improving the quality of existing capacity. Capacity enhancements are generally costly, time consuming and often politically difficult. Setting appropriate network standards and improving the robustness of existing capacity (for instance, durability of material) is a decision on infrastructure quality that also impacts on reliability.

- Better management of existing capacity and service can facilitate reliability just as poor management can increase unreliability. Infrastructure managers can improve reliability through better incident management and appropriate scheduling and publicising of maintenance. The core management skills can be supplemented by pro-active network oversight.

- Where feasible, charging directly for reliability would encourage an efficient and an appropriate level of reliability. However, it is often difficult to provide different levels of reliability and to extract different charges for that differential performance.

- Information can be provided to users enabling them to mitigate the adverse effects of poor reliability. This may be a cost-effective way to reduce both unreliability as well as the impacts of traffic incidents on subsequent business and personal schedules.

Incorporating reliability into cost benefit assessments encourages consideration of options for delivering appropriate levels of reliability.

In the absence of a direct market for reliability, cost–benefit assessments can be used to determine appropriate levels of reliability. If a separate market existed for reliability then prices would encourage an efficient level of reliability and would allocate responsibility for reliability to the party that could bear it at least cost. A cost-benefit analysis attempts to proxy such a market. This study has found that reliability is rarely embodied in such analyses.

Projects designed to deliver travel time benefits (such as those arising when congestion is reduced) are sometimes credited with generating reliability benefits. However, standard appraisals fail to unbundle benefits from improved reliability (reductions in travel time variability) from the benefits due to the reductions in average travel time. This omission removes the factual basis for arguing that a project really does improve reliability.

There are ways to measure and value reliability that can be integrated into cost-benefit analysis. These have been used on a pilot basis in a small number of member countries. These approaches provide a foundation for explicitly incorporating reliability benefits into investment appraisals and, consequently, policy frameworks.

Diversity in network user demands for reliability means that no simple mark-up can be applied to incorporate reliability into project assessments

It is difficult to generalise about the value of reliability as it will be project, location, user-, and time-specific. For one project studied, the value of improvements in reliability were found to be negligible, whereas for another project they were found to add 25% to the welfare benefits of time savings achieved. It is important to recognise the ‘granularity’ of reliability, that is, the different values placed on reliability by different network users at different times and for different trip purposes.

Since the demand for reliability varies markedly across users, products, locations and firms, a single monetary value for reliability will be of little, if any use, in project appraisal – a range of values is required that represents the major user groups in each case. Practitioners cannot assume that values used in one study are readily transferable to a project in another situation.

It is also important to avoid potential double-counting when factoring reliability into project assessment. This can arise if the standard values of time used to assess average time savings have incorporated an implicit, crude value for reliability.

Governments often set reliability targets but these need to be applied with caution, distinguishing between the network/operator and the user perspective

Reliability targets and performance indicators for services and infrastructure performance can facilitate discussions between users, operators and decision makers regarding the "right" levels of reliability. But employing fixed targets may be distorting as they can dominate other service characteristics that may be of equal or greater importance. Also, such targets invariably present an average level of reliability not reflecting the diversity in the demand for reliability.

There are also trade-offs to be made. For instance, a rail infrastructure manager may enhance reliability by reducing the number of trains that it operates. The improvements in reliability may then come at the cost of a more limited train schedule and higher overcrowding on the trains. In those cases where only passenger trains face performance targets, network managers may be inclined to give higher priority to passenger trains over freight trains than economically justified. Targets should therefore aim at reflecting both the network and the user perspective.

EXECUTIVE SUMMARY

Most of us face unreliability in our daily lives, with unexpected travel delays leading us to miss a train or arrive late for school or work. Whether it be those social or business events, or the deliveries of goods, reliability is a key quality of movement. However, a review of policies in OECD/ITF countries shows that few countries explicitly incorporate reliability into transport policy making. This report aims at providing policy makers with a framework for understanding reliability issues, for incorporating reliability into project assessment and for designing reliability management policies.

The economics of reliability

Reliability is unanimously regarded as a desirable transport network attribute. There is less unanimity in the transport sphere in defining reliability. Yet the definition adopted has major implications for policy. Technically, a reliable system is one that performs its required functions under stated conditions for a specified period of time. Under this definition, a road system that becomes choked with traffic during peak hour, reducing speeds to a slow 20 kilometers per hour, could be regarded as only "50 per cent reliable" or even as "unreliable".

A less specific definition of reliability draws on the attributes of *predictability*. In this context, a congested road system where the speeds at different times of day and different days of the week are consistent, and hence predictable, would be ranked as "highly reliable". While both interpretations are valid, the focus of this study will be on the second definition.

Like all desirable features of a transport network, reliability comes at a cost. It is subject to the standard rules of supply and demand: the higher the price, the higher the quantity that will be supplied but the lower the quantity that will be demanded. Conversely, the lower the price, the more consumers will demand it. The challenge for policy makers arises in two areas. The first is in formulating the institutional arrangements that impact on the market for reliability. For instance, it would be important to avoid a legal framework that placed impediments to differentiating between services on the basis of reliability. The second is in the treatment of reliability when assessing publicly-funded transport infrastructure projects

In other words, the role of the government is two-fold: encouraging a market for reliability and incorporating reliability into the assessment of transport infrastructure projects. In terms of the first role, it is important to note that, as a service attribute, reliability is often bundled with other attributes such as speed, convenience and cost, making it very difficult to differentiate a separate market for reliability.[1]

An important point that follows from this is that only when, say, two parallel services are provided with reliability being the key differentiating feature is there an explicit market for

1 This is a common feature of all markets, as rarely, if ever, is the array of goods and services so vast that all consumers can select the exact amount of each attribute that they are willing to pay for.

reliability. Without this, there is a major challenge in developing sound estimates of the value placed on reliability by network users.[2]

Ideally, the market incentives would encourage not only an efficient level of reliability but would also allocate responsibility for reliability to the party that could bear it at least cost.[3] This point is also explored in the report.

Second, factoring reliability into cost–benefit analysis is desirable but problematic. Using an incorrect value could result in a worse outcome than a failure to incorporate a value for reliability at all. The values placed on reliability vary from project to project. Cost–benefit analysis, as a set of rationalised economic principles, has evolved over more than a century and useful refinements are unlikely to be developed overnight.

However, this report has made significant progress in identifying possible methodology for incorporating a value for reliability into project evaluation, as well as exploring the possible pitfalls that need to be avoided.

Reliable transport networks and services are required because of more complex and inter-related supply chains and increasingly complex scheduled activities

Changes in commerce and personal travel patterns have increased the importance of a reliable transport system. The physical way that the economy operates has changed facilitated by, and demanding, transport sector enhancements.

Transport productivity has increased markedly, yielding benefits for business through the specialisation of production on a global scale and the spread of just-in-time production and distribution systems. One aspect of that productivity is the reduction in transit time, which expands the market for the goods and services, and broadens the way in which firms can interact. However, the increased interaction between businesses means that firms depend on reliability. In modern dispersed production systems, “time” has become the critical factor where timely delivery of components has replaced traditional stock-holding. These developments have facilitated and accompanied the growing operational sphere of influence for businesses. Businesses themselves have consolidated into larger, but fewer, physical locations, growing the globalised economy. Broadening national and international trade links, with increasing goods movements, has brought greater volumes of goods, moving further and in increasingly complex and—crucially—interdependent ways. That interdependence relies on reliable transport.

There are also changes in personal lifestyles. Passenger movements, both for business and social purposes, have become more complex with increased disposable income, recreational choices and leisure time. Complex commuting and leisure activities have increased reliance on robust network performance. Improvements in road and car quality, and more disposable income and leisure choices, have raised expectations of, and demands for, reliable transport – particularly private transport. These diverse and geographically-spread activities have led to more intensive use of transport systems, bringing greater dependence on transport to be reliable so that delays do not cascade through the busy calendar of events. Thus, the scheduling approach adopted in private lives echo the “just-in-time” deliveries in commerce.

---

2 For instance, if they cannot be charged directly, they are likely to say they place a much higher value on reliability than otherwise.

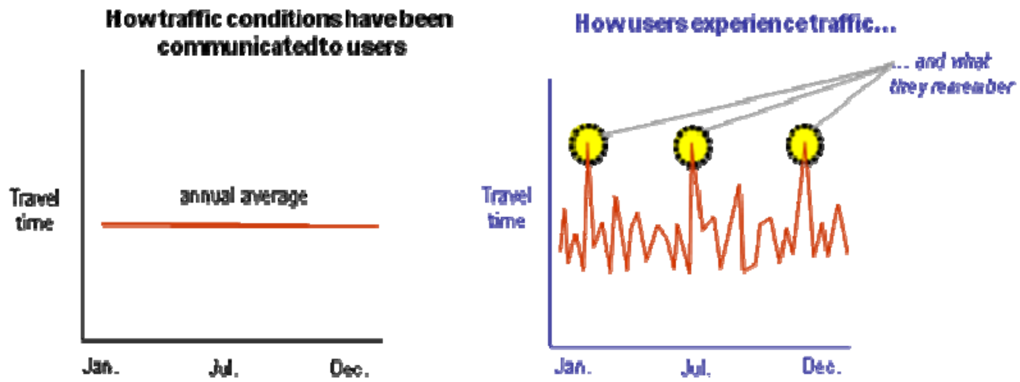
3 In this way, reliability is analogous to risk.

The importance of scheduling in personal and freight activities has grown, so that transport unreliability has an increasingly-marked effect on downstream activities. The expectation from these demand trends is increasingly that transport should provide high levels of reliability.

Unreliability makes trips frustrating

Unreliability makes journeys frustrating and causes stress. The feeling of travelling without control over one's travel time is a disempowering experience and "bad" experiences are remembered by travellers. Looking at Figure 1 below, traffic conditions in the past have often been communicated to travellers only in terms of simple averages (left chart in figure). However, most travellers experience and remember something much different than a simple average of commuting travel time (right chart in figure). Users have deeply negative perceptions of unexpected delays, which colour their attitude to the experience.

Figure 1. **Travellers' perception of traffic conditions**



Source: USFHWA (2006).

Unreliability constitutes a cost

Where performance is inconsistent, network users may simply have to accept the consequences of the delay, albeit that it may have ripple-effects or, worse, snowballing (compounding, or growing) effects, impacting on other activities or stages in the personal or logistics chain, constituting a cost to those involved.

The ripple-effect of delays is an important reminder of the inter-connectedness of many individual schedules. A delay at one stage in a person's schedule of activities can mean delays in later related or unrelated tasks. Similarly, while logistics chains are built in such a way as to reduce their vulnerability to individual events, any delays in individual consignments can still reverberate through the chain. Indeed, because the transport task is part of a chain, a break in any part of it is a break in the entire chain. An assembled television set with only 99 of its 100 components is an incomplete product that can be neither shipped nor sold.

Costs of unreliability may rival those of congestion. Bearing in mind that the results are not transferable across locations, it is nonetheless significant that two recent studies found that unreliability costs caused around half of total underlying delay costs.

Journey-time predictability is a defining feature of reliability

In this report, reliability is defined as:

the ability of the transport system to provide the expected level of service quality, upon which users have organized their activities.

The key word is "expected". According to the definition, reliability can be improved either by supplying a higher level of reliability, or by changing expectations of the level of reliability.

In other words, unpredictability (or inconsistency) of network performance is the defining characteristic of unreliability. The more random (less predictable) the performance, the harder it is for the network user to insure against delays.

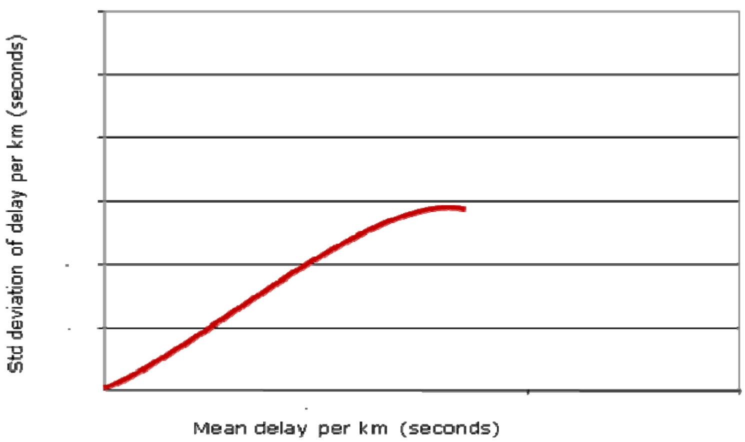
Average travel time between two destinations includes both expected and unexpected delays. It is assumed that network users accommodate expected delays into their travel time through, say, the inclusion of buffer time. However, it is more difficult and costly to incorporate the unpredictable — the unexpected — delays that lead to variation from planned (anticipated) travel time.

Disturbances that cause delays can also be classified as "recurrent" (such as weekday peak-hour congestion) or "non-recurrent" (such as crashes, inclement weather and other events of nature). The essence of the degree of recurrence is that it provides information about the predictability of the event. This report focuses particularly on the non-recurrent events since they are, by definition, unpredictable.

The terms of unreliability and congestion are often used synonymously. However, as follows from the foregoing discussion, a congested network does not have to be unreliable. Unreliability refers to unanticipated delays, and therefore a congested network is not necessarily unreliable because journey time along a congested road can be fairly predictable.

That said, congestion increases the likelihood of unreliability: as traffic levels increase, the time delays due to slight perturbations tend to increase more than proportionately. This is illustrated by one example, a motorway in the United Kingdom (see Figure 2 below), where there is a clear correlation between the level of congestion and reliability until high levels of congestion are reached. That said, it is not possible to say whether the variability of travel time was predictable or not.

Figure 2. **Relationship between reliability (vertical access) and congestion (horizontal access) on the M42 motorway in the UK**



Source: UK Highways Agency.

The distinction between reliability and congestion is important because of the different policy implications. However, it is also recognised that remedial actions directed at congestion can improve reliability and, similarly, actions that improve reliability can reduce congestion. For instance, many of the bottlenecks in international supply chains are located in congested urban areas. Reducing congestion at port and hinterland connections may also improve the reliability of the entire logistic chain. That is to say, there can be overlaps.

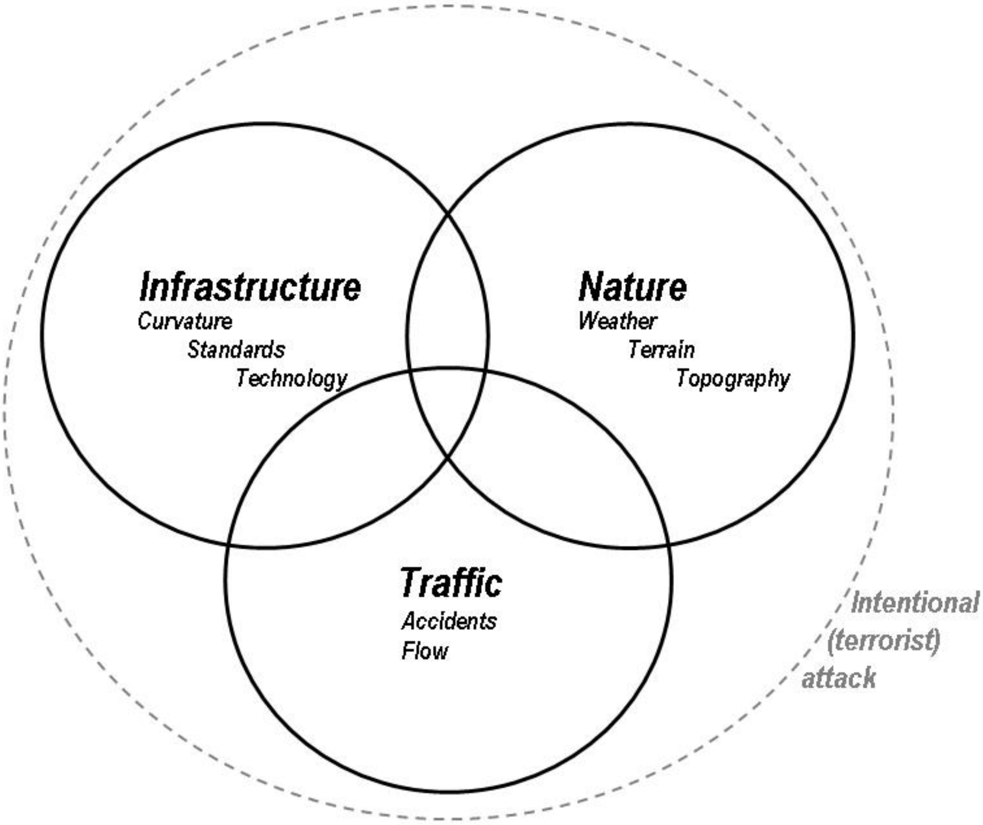
Unreliability arises from multiple sources, each requiring different ways to manage the problem

Unreliability can arise from various activities which are within the control of the network user or provider. Unreliability of the transport infrastructure network arises from two primary sources:

- Unpredictable demand-related traffic interactions between users (congestion).
- Unanticipated supply-related:
  - traffic incidents (accidents and vehicle break-downs);
  - natural events (*e.g*. floods and earthquakes);
  - network maintenance (causing temporary reduction in supply); and
  - mismanagement in infrastructure supply, which can also include inappropriate maintenance programs.

Mismanagement of road and railway networks can reinforce other sources of unreliability. It is possible that an uncongested road can be unreliable if the network is poorly-managed; similarly, a congested road with poor management is likely to magnify the unreliability. This observation is represented by the intersection of the circles in Figure 3, showing the primary sources of unreliability.

Figure 3. **Primary sources of unreliability and inter-relationships**



The figure above illustrates the interfaces between the various sources of unreliability. For example, low standards of infrastructure are likely to be more prone to unreliability arising from events of nature than if the infrastructure is set to a high standard. This is not to argue for infrastructure to be built to a high standard by default; given prevailing conditions, such as the likelihood of disruption and low levels of traffic, it may be highly appropriate for the infrastructure to be built to a low standard.

Finally, reliability issues are very location- and time-specific – and this affects potential actions to manage the problem, as well as the degree to which costs and benefits from one situation can be inferred to another situation.

In response to unreliability network users develop strategies to deal with unreliability

Individuals and companies affected by deterioration in reliability respond in a number of ways. As a consequence of unreliability, schedules for our commuting or leisure activities become disrupted as does the scheduled chain of flow of goods.

To reduce the risk of being late at the destination, network users may allow more time for the travel (the so-called “safety margin” or “buffer”). This means, in practice, leaving earlier to ensure arriving on-time. Also companies and logistic managers adapt their operations either through changing the way they operate, or by building in buffer stocks of goods. Deliveries can avoid daytime and peak delays, and there has been a growth in evening or night-time deliveries; in some instances companies make greater use of regional depots. Companies also adapt their logistic operations by active traffic management schemes. Increased use of vehicle telematics, routing software and fleet management packages have assisted the adjustment to more congested infrastructure. Minimising the impact of delays on the cost and quality of logistics has become a core skill for freight and logistics managers.

However, each of these options has an associated cost. Leaving earlier to ensure arriving on-time consumes the time available for other, more productive, activities. Holding additional stocks of goods “just in case” involves a capital cost both in terms of the storage facilities and financing them.

Governments have started measuring and targeting reliability

The first step to recognise the importance of reliability is to measure it. A number of countries have been exploring ways to measure reliability. Two distinctive activities are involved here: measuring service reliability and the setting of targets against which the service provider’s actual performance is compared.

A review of existing reliability indicators suggests that the purpose of such measurement is seen as a way of ascertaining the quality of performance of transport service delivery. However there are several shortcomings in some of the reliability indicators currently available:

- Aggregation across users: Most existing reliability indicators monitor the performance characteristics of the whole system rather than satisfaction of users’ needs. That is, whether different users actually receive reliable services.

- Aggregation across time: The indicators normally show overall annual averages only, and therefore mask smaller-scale temporal variations in service standards.

- Reporting partial data: More generally, most of the existing indicators were originally designed to provide feedback to network managers, rather than to measure reliability as perceived by end-users. Thus, the indicators may report operational details such as freight train arrival times rather than those of primary interest to customers, such as the predictability of collection or delivery times.

The initial government response to unreliability in many countries has been to set standards and performance targets. Performance targets are set for three primary reasons:

- Reliability is an important service characteristic in the transport sector.

- The services to which targets are applied often involve monopoly provisions underwritten by the taxpayer. Hence, governments have a vested interest in seeking attractive services and efficient provision.

- Reliability targets are important for initiating discussions between politicians, operators, providers and users on the appropriate delivery of service standards.

Most of the existing reliability targets can be found in the rail sector, a transport mode that seeks to run to strict “working timetables”. Target-setting practice is prevalent in passenger railway for two reasons. First, the scheduling of arrival times readily enables these types of targets to be set (setting actual train arrivals against scheduled arrivals); and, secondly, because the service provider is considered to be a monopoly provider.

To the extent that the provider is perceived to be a monopoly provider of services or facilities, governments often oversee supply standards by monitoring and/or setting performance standards; the target provides a degree of accountability in service quality. Data on service reliability are essential for this oversight and policy stance. A similar approach is adopted in aviation, where airline service punctuality statistics provide bellwethers for regulatory and policy monitoring.

The actual service performance and the performance against targets are often published as a way for regulators and governments to make service providers accountable, and as implicit encouragement to improve services.

The publication of service performance is also relevant for network users to understand the quality of service delivery, enabling users to allow adequate buffer time against delays.

As noted earlier, this study has found that, despite its obvious importance, there is generally no explicit view on what travel time reliability *is* precisely; similarly, there is no consensus on how reliability should be monitored. Various definitions for travel time reliability exist, and subsequently many different relevant indicators are available. Crucially, there is little recognition of the risks in setting targets (or the difficulty in establishing cost-effective targets); too high might distort desirable management decisions, while low targets might make service provision quality too lax.

There are also trade-offs to be made when influencing reliability through explicit targets. For instance, a railway operator may enhance reliability by reducing the number of passenger trains that it operates. The higher reliability then comes at the cost of less-frequent, higher-loaded trains. Similarly, if performance targets were applied only to passenger trains, the network managers may be inclined to give higher priority to those trains over freight trains than economically justified.

A small number of countries have incorporated reliability into project cost benefit assessment but so far fail to reflect diversity in reliability valuations

Case studies reviewed in this study illustrate that some projects are carried out specifically in order to improve reliability. However, there are very few cases where reliability is formally incorporated into the cost-benefit assessment (and hence in the decision making process). Even where decision-making guidelines do incorporate reliability, most of the actual project appraisals do not include monetised parameters for reliability.

When appraising transport investments, projects are often dominated by improvements in safety and travel time. The time benefits are traditionally measured as the improvement in journey time. In incorporating reliability, those time benefits are split into travel time savings and savings in reliability (buffer) time. A monetary value is then given to time. Both would vary across users, trip purpose, and location. Thus, the focus on incorporating reliability into cost–benefit analysis should be on ensuring adequate user granularity, using the value of time as the basis for monetising the reliability benefits.

In a small number of countries (the United Kingdom, Netherlands, Denmark, New Zealand, Norway and Sweden), some project appraisals do incorporate reliability. However, the values used are typically based on values that are considered to be the same for all users. This approach is not fully adequate: the value of reliability is, inevitably, very “granular” — diverse across users — with a wide spectrum of values. It is important therefore to identify values that differentiate between the transport modes and journey purpose/task. Using coarse (or, worse, a single) value for reliability improvement will distort the outcome (and, even more so if it is not location-specific).

Delivering optimal levels of reliability

At present, reliability is generally not taken into account when evaluating a project. If an infrastructure investment, for example, is aimed at improving travel time reliability rather than average travel time, such project merits will be overlooked.

To appraise reliability effects in cost-benefit analysis it is important to measure both average travel time and travel time variability. If the project fails to separate these two measures but argues that the project is indeed improving reliability, the assessment lacks a factual basis.

Average time savings should be split into travel time reductions and a reduction of travel time variation. Both of these components should be identified. An appreciation of the traveller types using the link would then enable appropriate values for the components to be applied. This unbundling enables planners to gain insight into the relative levels of reliability benefits.

Further, *ex ante* cost benefit analysis will require some quantification of the expected reliability effects of policies. This appears to be a poorly documented field. It is therefore necessary to improve the traffic forecasting tools and models currently available. Ideally, these tools should be able to provide estimates of future changes in the standard deviations of travel times on links, and model the influence of such variables on travel and demand and network use.

Above all, because reliability issues are location, user, time etc specific, countries should avoid applying or repeating the use of a single value for reliability, or applying a value that has been used in one study to a project in another situation. For each project, there are differences in the mix of user groups and time/reliability splits.

Integrating reliability in transport policy impact appraisals requires further development on the following fronts:

- Use and derivation of a commonly-accepted measure of reliability (such as buffer time or standard deviation).

- Development of assessment methodology (data and survey requirements, principles for identifying primary user groups).

Options for achieving reliability should be selected on the basis of cost-effectiveness

A key policy challenge is to create incentive structures that encourage selection of the most cost-effective reliability option – that is, adopting the option that delivers a given level of reliability improvement for the lowest cost. The objective is to ensure that option is chosen ahead of the less-effective options, regardless of whether the responsibility for adopting the option lies with the network provider or the network user. For instance, one project scenario might conclude that the cost-effective way for just-in-time shippers to achieve greater reliability would be to hold more stocks than for the network provider to incur incremental infrastructure costs.

In order to be able to take into account reliability in policy impact evaluation, only a cost-benefit assessment framework provides consistency in assessing the societal pros and cons of policy interventions in terms of their positive or negative effects on reliability.

Network/operator perspective and user perspective should be distinguished

For policy making, it is important to measure and report on both network/operator and user perspectives of reliability. The way reliability is measured and communicated provides a policy signal in itself. Also, the better informed regulators are about the appropriate reliability targets, the better the policy.

The Working Group recommends that the distinction is made between the system- and user-perspective of reliability measurers.

1. For a network *provider or operator*, the focus is on:

   – system robustness / vulnerability. Here, a further distinction is made between indicators to be applied to measure developments, either in link or network performance, under changing conditions

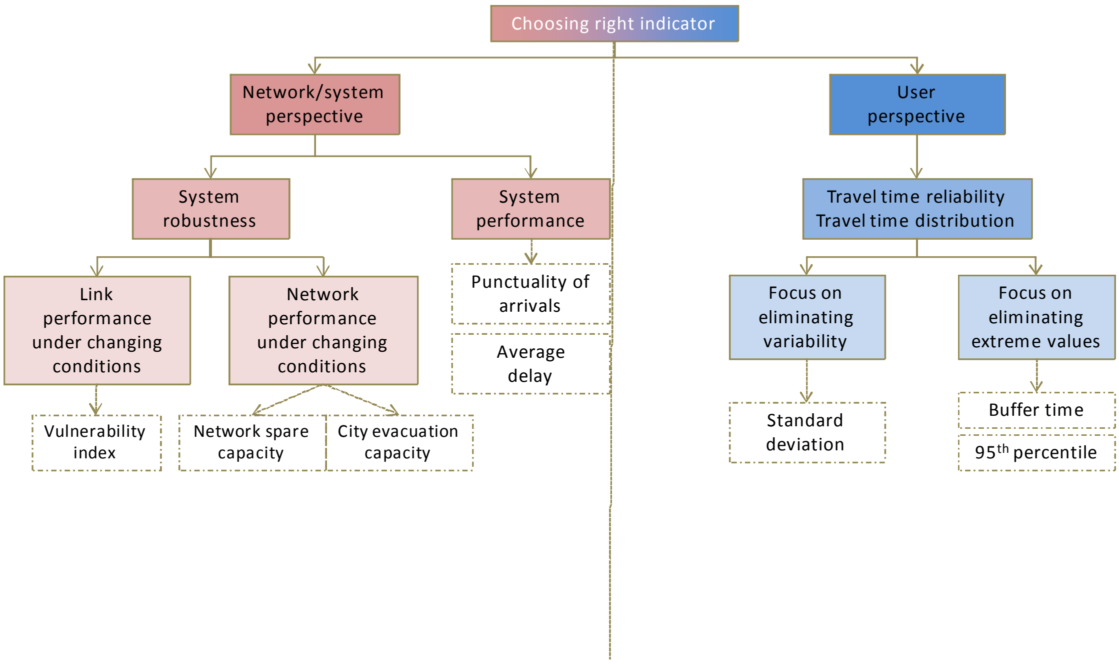
   – system operating performance. Here, the focus is on indicators to describe the performance of a system in terms of deviations from expected, or agreed, levels of service.

2. For a network *user*, the focus is on:

   – the variability of travel times experienced by the user. This provides useful travel-planning information. A further distinction is made between indicators to describe issues regarding general variability of travel times, and issues regarding the elimination of extreme unexpected travel times.

Based on the review of existing indicators, the working group has adopted the following schematic overview of the different purposes for indicator combinations (Figure 4). The main conclusion of the Working Group is that it is extremely important to look at both network and user perspectives, as each has different policy implications.

Figure 4. **Recommended reliability indicators by purpose**



New policy framework

There are many techniques and instruments available that can be used to improve the reliability of the transport network, both individually and in combination with each other. The Working Group has identified four principal policy options available to manage reliability:

- Physical expansion of capacity.
- Better management of capacity.
- Pricing mechanisms to deliver a market for reliability.
- Reliability information systems – these are intended to reduce—mitigate—the adverse consequences (*i.e.* the costs) of unreliability, rather than to reduce the incidence of the unreliability.

In general, these are not necessarily alternative options but, nonetheless, each should be subject to cost-benefit appraisals.

Physical expansion of capacity

On the supply side, infrastructure design and construction can incorporate reliability options. Improving supply-side reliability entails reducing the probability of an unexpected disruption in service. There is a wide array of options to enhance capacity by expanding infrastructure: upgrading and adding line capacity, increasing transport service in corridors and transfer points, construction of new highway lanes and alignment, as well as new rail lines, transit routes and terminals.

Infrastructure can also be built at standards that reduce the need for maintenance or improve the robustness of the capacity. It is notable that these supply side, capital-based "build" options are implemented before any incident takes place. Hence, adaptability of the infrastructure is a key issue.

Supplying new capacity is costly, time consuming and often politically difficult, while setting higher network standards and improving the robustness of capacity may deliver higher reliability more cost-effectively. It is too often the case that additional supply is considered as the only option, whereas it should be considered as an option among others.

Better management of capacity

There is a wide range of techniques and instruments available to better manage network capacity to improve reliability. These techniques address both recurring and non-recurring causes of unreliability.

The use of pro-active oversight and management of transport networks can address both types of unreliability, either through pro-active insight in vulnerable network-characteristics and enhanced incident management. For instance, the impact of congestion on reliability might be reduced by the use of variable road speed limits and the temporary addition of road capacity, using emergency hard-shoulder break-down lanes. Similarly, improved management oversight can also be applied on the railway network. Optimised timetabling, a dynamic rescheduling of rail networks in case of an incident, and advanced train management systems can be used.

Enhanced management techniques can assist in reducing the impact of maintenance on network users' reliability and reduce the cost of the maintenance activity itself. For example, some contracts in Public-Private Partnership projects have included charges for maintenance works to discourage private network owners from adding too many work zones at the same time.

In summary, an important policy focus for delivering reliability is to better manage capacity through dynamic network management. A focus on interfaces, such as border crossings and ports and hinterland connections where unreliability is likely to occur, might also be appropriate.

Developing mechanisms for charging directly for reliability

Charging for transport networks, or portions thereof, is becoming a more common method of managing traffic demand, and consequently traffic flow and network reliability. It should also be noted that it is already possible to charge for information systems, such as GPS, which network users can buy into in order to mitigate against the worst effects of delays. Charges can be applied selectively to segments of the transport network, or more broadly over large sections of the network.

Developments in technology have facilitated an expansion of charging schemes and techniques to strategic parts of road networks that can be used to manage demand on transport networks. Although most of these techniques are directed at cost recovery and demand management, they can also be effective in improving reliability.

There are few situations on road networks where access has been restricted and where charges have been introduced to improve reliability. “Dynamic pricing” on the Interstate-15 in the USA is one of the few instances. In this example, charges are adjusted up and down to ensure a predictable travel time for the 8 miles of road involved.

In principle, because their control of access to the network allows network-link charges, railways are better placed to use charging as a tool to deliver a consistent level of reliability. There are limited examples (in North America and Australia, for example) where high-reliability freight train services are offered. In general, however, freight railways’ profit-maximising strategy is to move large amounts of freight that does not require very high reliability standards; there are only a few exceptions to this. By contrast, high-speed passenger trains/tracks (ICE, TGV, Pendolino, etc) have been built to provide near-exclusive rights for services with low transit time and high reliability. Infrastructure charges for these trains are correspondingly high.

In summary, then, charging directly for reliability by setting differential charges for infrastructure use and service supply, according to the level of reliability, might deliver an appropriate level of reliability. However, it should be noted that it is often difficult or impossible to differentiate charges sufficiently to match the level of reliability demanded by different types of user of transport infrastructure. Also, the cost of a charging system that discriminates on the basis of reliability could outweigh the benefits and must be included in cost benefit assessments of charging systems.

Mitigating the cost burden associated with unreliability using information

There are many ways network users can reduce the cost burden associated with unreliability. Information systems can reduce the consequences or mitigate the effect of network incidents. Network demand can be deflected away from the site of congestion or traffic incidents. The information can also reduce in-vehicle stress associated with unreliability, and work to ease and manage the problems associated with delays to schedules.

As noted previously, travel time reliability depends, to some extent, on the user’s expectation of predictable travel times; this expectation can vary according to the information available. Network providers can facilitate network usage and impact of unreliability by informing users of prevailing conditions. The information does not reduce the incident happening but, rather, reduces the costs that arise from the incident. For example, the widespread adoption of the mobile phone in recent years has provided the network user with the means to alert interested parties (the warehouse, the family) that arrival will be delayed; the latter parties might then be able to reduce the impact of that delay. Hence, information can mitigate the unreliability and reduce the ripple effect or snowballing effect that otherwise would be the result of unreliability.

Information options may be further divided between pre-trip and on-trip measures. Information may be used in different ways to improve reliability depending on whether a traveller has left the origin, whether a traveller can divert to another route, or if the traveller cannot divert but can reduce the ripple effect (consequences). Different tools exist for delivering this information, including variable message signs, car navigators and the internet.

Information can be provided to users to mitigate the effects of poor reliability. This is often a cost-effective way to reduce unreliability costs and the cascading impacts of traffic incidents.

Conclusion

A wide range of policy instruments is available to manage reliability. Because there generally is no direct market for reliability, cost benefit assessment needs to be used to determine appropriate levels of reliability and to select cost-effective policies to manage reliability. Cost benefit assessment has been applied to projects designed to improve reliability in only a small number of countries with techniques that are in some important respects unsatisfactory. The report makes significant progress in identifying appropriate methodology for incorporating a value for reliability into project and policy evaluations, as well as exploring the pitfalls that need to be avoided. This can be used to develop robust and consistent reliability assessments, important for the selection of cost-effective policies and projects and for informing decisions on achieving more optimal levels of reliability on surface transport networks.