Methods, measures and indicators for evaluating benefits of transportation research

L du Plessis[1], JJ Krüger[2]

L du Plessis, PhD candidate at the Centre for Research and Continued Engineering Development, North -West University Suite 93, Private Bag X30, Lynnwood Ridge 0040, South Africa, co-affiliation, CSIR, P.O. Box 395, Pretoria 0001, South Africa

Phone: +27 12 841 292, Fax: +27 12 842 7158, E-mail: lplessis@csir.co.za

Corresponding author

JJ Krüger, Centre for Research and Continued Engineering Development, North -West University, Suite 93, Private Bag X30, Lynnwood Ridge 0040, South Africa.

Phone: +27 87 980 3602, E-mail: jjkruger@ashemia.net

Methods, measures and indicators for evaluating benefits of transportation research

The purpose of this article is to provide updated information by identifying and discussing methods, measures and indicators for evaluating benefits appropriate for transportation-related research facilities/programs. The information has been drawn from within and outside transportation research. The article discusses the sources driving the need for evaluating benefits and describes the challenges confronting the evaluation process. It reviews and compares qualitative and quantitative techniques and highlights previous published work, investigations and case studies.

Many traditional challenges of determining benefits persist, contributing to the gap between the ability to identify non-technical benefits of research and the growing need to demonstrate such benefits. This article aims to stimulate dialogue and investigations to advance the development of an appropriate robust method to determine quantitative benefits stemming from specifically Accelerated Pavement Testing (APT) type transportation research. The ultimate goal is to help better understand, demonstrate and communicate the benefits of APT research.

Keywords: transportation research benefit determination, benefit-cost analysis, accelerated pavement testing

Introduction

The direct economic benefits of an Accelerated Pavement Testing (APT) research program for potential use by government and research agencies were identified in an earlier pilot study (Du Plessis *et al*. 2011, Nokes *et al.* 2011). The need to understand techniques for assessing the benefits of such research continues to grow. Clear (non-technical) justifications that emphasize costs and benefits (such as cost-benefit analyses) may increase public confidence in decision-makers (Baron and Gurmankin, 2013). In this article , we provide a unique description of contemporary techniques and measures for qualitative and quantitative evaluation of transportation-related research.

Accountability for public expenditure has only grown during the recent global economic downturn, leading to demands for greater transparency, more scrutiny of public agency processes and a greater need for the use of state-of-the-practice methodologies. European Co-operation in the Field of Scientific and Technical Research (COST) study suggested that applying methods to evaluate costs and benefits increase funding because of better marketing of transportation research activities and results (European Co-operation in the Field of Scientific and Technical Research, 2005). Substantive and pro-active efforts to demonstrate the value added from investing in road research have documented previous studies showing qualitative benefits as well as quantified economic returns on funds spent on research.

The international scope of evaluating non-technical benefits of research is indicated by the 2008 scanning tour conducted by United States (US) research administrators to review transportation research program administration practices. This included the valuation of research and ways to enhance US transportation research administration (Elston *et al*. 2009). Discussions with senior research program administrators in national governments, the European Commission, non-governmental research consortia, universities and other research organisations in Europe, Japan and South Korea have led to the following findings:

- Unlike in the US, research programmes in many other countries do not have to continually justify expenditures.
- While research programs in all countries have a process for evaluating results, the techniques vary in complexity, effectiveness and success.
- As in the US, research programs in all countries face continuing challenges in quantifying benefits of research. No country has a totally satisfactory method. However, an important difference from the US is that justifying research based on the analysis of benefits is not a critical concern in any of the countries visited.

One of the scan team’s six recommendations was to improve research evaluation processes by promoting systematic and consistent practices. Future international collaboration may require compatibility of research evaluation methods.

Ongoing activities at the US Department of Transportation (US DOT) Research and Innovative Technology Administration (RITA), which focus on compiling results from cost-benefit assessments of Intelligent Transportation Systems (ITS) in the US and abroad, provide potentially useful examples for benefits evaluation of APT. RITA hosts a knowledge resource portal to help measure and document the benefits of ITS within certain goal areas, such as safety, mobility, productivity, energy and environmental impacts (US Department of Transportation, 2016).

These activities by the US DOT are linked to the International Benefits, Evaluation and Costs (IBEC) Working Group, which was created to coordinate and expand international evaluation efforts, exchange information and techniques, and evaluate benefits and costs of ITS

(The International Benefits, Evaluation and Costs Working Group, 2014). IBEC facilitates dialogue about topics of interest to the international community of ITS evaluators and encourages the more effective use of information from evaluations. They aim to bring about better informed decisions about ITS investments.

Because of the intended practical use of these qualitative and quantitative transportation research methods, we focus on practical applications. This article is not intended to focus solely on academic literature in this broad and evolving research field. Also, the approaches and case studies that are described in this article are indicative rather than exhaustive.

Indicators of qualitative benefits are described, whilst quantitative methods and measures are emphasized. This article does not discuss macroeconomic impacts, but focus on direct economic impacts that can be attributed to specific research results instead. Research administration, management and policies are not examined in this article, although they are linked to benefits assessment and should gain from the material herein.

Research Objectives

The objective of this article is to provide updated information by identifying and discussing methods, measures and indicators that may be suitable for evaluating benefits from full-scale APT. The author’s intent is for the information to aid in translating technical pavement measures well-known to APT experts into quantitative measures and qualitative indicators so that public decision-makers can understand and appreciate various returns on investments in APT. The article discusses the sources driving the need for evaluating benefits, describes the challenges confronting the evaluation process, reviews and compares qualitative and quantitative (including direct economic benefits) techniques. It highlights previous reviews, investigations, and case studies. The main research objectives are:

- Identify major areas of benefits, determine common methods, measures and information required to reasonably determine benefits of implementing research results.
- Identify current research evaluation methods and (qualitative/quantitative) benefit metrics used by transportation agencies for determining the value of research results.
- Identify the critical knowledge gaps in the evaluation of research results that require further research.
- Suggest appropriate techniques and methodologies suitable for APT related research benefit determination.

In a recent study published by the Southeast Transportation Consortium (2014) three elaborate surveys were conducted to capture state of knowledge and practice in determining the value of research in DOTs and to collect the best examples for determining value of transportation research. The study field covered the Southeast Transportation Consortium (STC) of the United States and included twelve DOTs of the following States: Florida, Louisiana, Arkansas, Missippi, Alabama, Georgia, South Carolina, North Carolina, Tennessee, Virginia, Kentucky and West Virginia. One of the findings of the report stated:

> "Although several methods are proposed for quantifying the benefits of research projects in the research reports collected in the first survey, there is no formal guideline or formal method to evaluate the quantitative and/or qualitative benefits of research projects in State DOTs".

The above findings point to the same conclusion: although several agencies have different ways and methodologies to measure the effectiveness of their research programmes there is a need to develop or adopt a methodology to quantify the value of transportation-related research. The aim of this article is not to develop a generic acceptable methodology or system to quantify the benefits of research projects. That is dealt by additional publications by Du Plessis (2011), Nokes and Du Plessis (2011) and Du Plessis and Prozzi (2008), but rather to investigate methods, measures and indicators that may be suitable for evaluating benefits as detailed above.

Challenges in evaluating benefits of transportation research

The broadest challenge in evaluating benefits is the broad range of expectations by those who focus only on results analysis (ignoring evaluation processes) for such purposes as budget or program justification, planning or decision-making. A review of the literature suggests three typical and wide spread expectations about the evaluation process:

- The process will lead to the “right” answer.
- The process will produce an “objective” analysis.
- The process will remove discomfort in determining benefits.

Cited information discussed in this article reveal that the expectations listed above are not met by contemporary approaches and their applications. However, the same information can aid those who must evaluate and communicate research benefits as well as those, such as decision-makers, program stakeholders or the public, who wish to understand evaluation results and processes.

A 1986 review of approaches used in US government and by industry to evaluate outcomes of federally funded research remains a landmark study of the past 30 years. The study was performed by the US Congress, Office of Technology Assessment (OTA) to provide information on the feasibility of quantifying outcomes in terms of return on investment (ROI). OTA performed an extensive literature review, conducted their own analysis of quantitative methods and interviewed economists, public policy analysts and research decision-makers in government and industry (US Congress, Office of Technology Assessment, 1986).

While the OTA study focused on basic research, questions were also examined about measuring ROI with regard to applied research and technology development. The study determined that two-thirds of federal expenditures on applied research was related to the production of public goods “whose primary value is not measured in economic terms”. The OTA concluded that viewing research as an investment is conceptually valid but such a view is of limited practical value, because factors affecting evaluation of basic research are too complex. They are subjective, payoffs are too diverse and institutional barriers prevent allowing quantitative models to replace “mature, informed judgment”. However, the OTA study also concluded that quantitative economic assessment was potentially useful for evaluating applied research and development and research facilities within a single, focused discipline. The latter conclusion appears to apply to APT.

The OTA study also examined non-economic measures, including bibliometrics (assessing research outputs in terms of publications) and “indicators” (evaluating research in terms of educational degrees, personnel, awards, etc.) as complementary tools. Despite limitations in bibliometrics, the study acknowledged their utility. The study found the utility and reliability of indicators more problematic and referred to them as “flawed” because of the narrow and subjective assumptions that must be understood to interpret them. The study found peer review (which dominated industry research at the time of the study) to be a necessary complement to the use of bibliometrics and indicators in order to overcome problems stemming from each method separately (US Congress, Office of Technology Assessment, 1986).

The use of various techniques to evaluate research outputs and their benefits continues to be controversial in government and industry. The OTA investigators deduced the following about the use of quantitative methods from interviews with industry research managers in 1986:

> "In industry, where one might expect quantitative techniques to prevail due to the existence of a well-defined economic objective for the individual firm or business, OTA found great scepticism among research managers about the utility of such techniques. Managers found them to be overly simplistic, inaccurate, misleading and subject to serious misinterpretation. There is little systematic data about the use of quantitative techniques. Most articles describe a process adopted by one firm or another without any indication as to how widespread the practice is in industry as a whole."

One of the basic challenges to evaluating research benefits in terms of ROI is the fact (as noted by the OTA investigators) that economic benefits are not the primary drivers for most government research, which focuses more on wide and non-economic aspects of public interests, such as safety, security, environment, health and generally advancing the body of knowledge. It is noteworthy that difficulties in assessing ROI arise in the private sector despite greater awareness and emphasis on economic aspects of research. Whether focused on economic or non-economic types of benefits, many challenges confront investigators when trying to evaluate benefits of research in any sector.

Reviews of approaches for evaluating benefits of research in a wide variety of fields are reported in the literature. Here, we provide an overview intended to assist potential investigations and use by APT owners/operators and researchers. Readers interested in more information about historical reviews may want to examine some of the documents listed in the References section. Descriptions of evaluation approaches reported in the past decade (some of which refer to older studies) are the focus of this article.

Highlights from previous studies show typical concerns, findings and insights about several challenges, including the following:

- Lack of familiarity with this topic;
- Scale of evaluation (i.e. test specific, project specific or programme wide);

- Complexity and context-sensitivity;
- Time domain, and
- Recurring procedural challenges.

The scale of evaluation, i.e. test specific, project specific or programme level, must be decided at the outset. The purpose of the evaluation probably determines its scale. Program-level assessments of research are reported in the literature in terms of determining whether a research program has achieved some preset targets or goals. Review of evaluations of research (inside and outside of transportation research) reported in the past decade emphasises program evaluations on a larger-scale and performance reviews. APT research evaluations described in this article show that assessments have been reported for all three scale levels (test specific, project specific or programme level). Each level presents its own challenges. A unifying approach does not exist. However, developments in recent years provide tools that may be customized for the level of evaluation needed.

Complexity is another challenge. Like the research process itself, evaluation of research benefits is a complex effort with uncertain outputs. Any effort to evaluate benefits from research faces substantial difficulties. The evaluation process can be labour-intensive, takes a long time to complete and requires expertise in the subject, all of which can lead to high costs. Deciding whether to proceed with benefits assessment and the process itself are context-sensitive, reflecting the attributes and constraints of the research products as well as those of the evaluators, the organisation and the users of the results. This partly explains why no universal technique has been found or recommended.

Understanding the challenge of the time domain requires stepping back, taking a long-term view of a research project’s life cycle and examining the typical sequence of actions in publicly funded transportation research. This helps to reveal some of the complications in assessing benefits. Figure 1 outlines the activities, their sequence and cash-flow (both costs and benefits) for a typical research project (Krugler *et al*. 2006).

The figure does not show the costs and processes in identifying a problem, defining and scoping a research project, incorporating the project into the organization’s overall research plan and obtaining funding as well as securing in-house or contract resources. Two aspects of this pre-research phase present additional barriers to performing evaluations. Firstly, the time required to initiate a project could add one or more years to the left of the timeline in Figure 1, thus further extending the period for evaluation. Secondly, the responsibility for pre-research activities is typically led by personnel dedicated to managerial activities and who often do not perform research or implement the results. These research management personnel are tasked to process the research program and its projects and not to evaluate possible benefits emanating from research outputs.

In the Active Research phase the researchers naturally assume lead responsibility for the work, while research management personnel monitor the work and ensure that reporting requirements are met. After completing the Active Research phase, the implementation phase begins. Researchers and research management personnel may be involved in implementation, but often the implementers (typical in operational functions such as design, maintenance and construction) of research outputs assume the lead responsibility for moving the research outputs into practice. After paying the costs of research and implementation, the implementation of research outputs should begin to gradually produce benefits. (Only annual agency cost savings are shown in Figure 1, above the timeline). Safety projects are expected to lead to immediate benefits in terms of lives saved and fewer accidents.

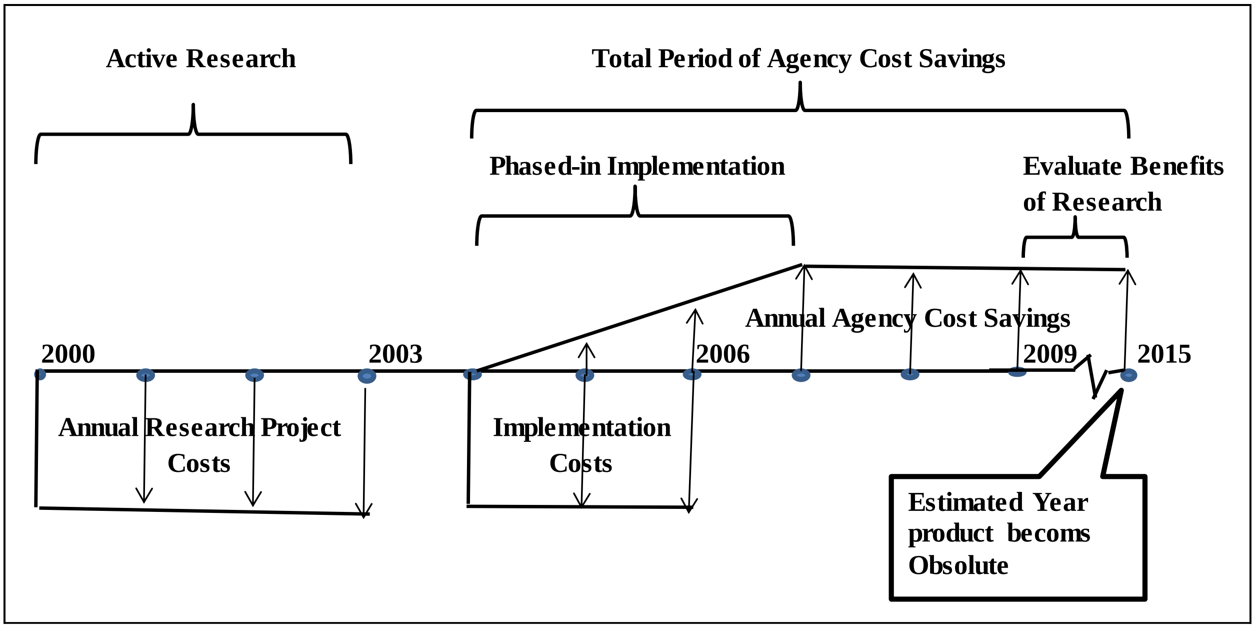

Figure 1: Public-funded project timeline and cash-flow (Adapted from Krugler)

This example illustrates the key points: (1) the research project life cycle is long, (2) monetary benefits accumulate long after research is completed, and (3) lead responsibilities change with each phase of the project. With successfully completed pre-research, research and implementation, the subsequent accumulation of benefits, which may take longer than any of the preceding phases, most likely requires continuing actions by implementers not associated with the original research project. Not shown in Figure 1 are activities (which would appear to the far right in the diagram) in conducting a retrospective evaluation of benefits. For the most credible results, retrospective assessment of benefits typically must await substantial implementation for most, if not all, benefits to accumulate.

The example in Figure 1 shows fairly constant agency cost savings that would most likely diminish gradually over a time horizon such as 5 to 10 years, or longer. As the research product heads toward obsolescence, new problems will be identified and new projects (possibly stemming from the completed research) will start as the cycle continues.

In a more recent study, Southeast Transportation Consortium (2014) reported that a similar research evaluation process will be developed for the Florida Department of Transportation (Figure 2).

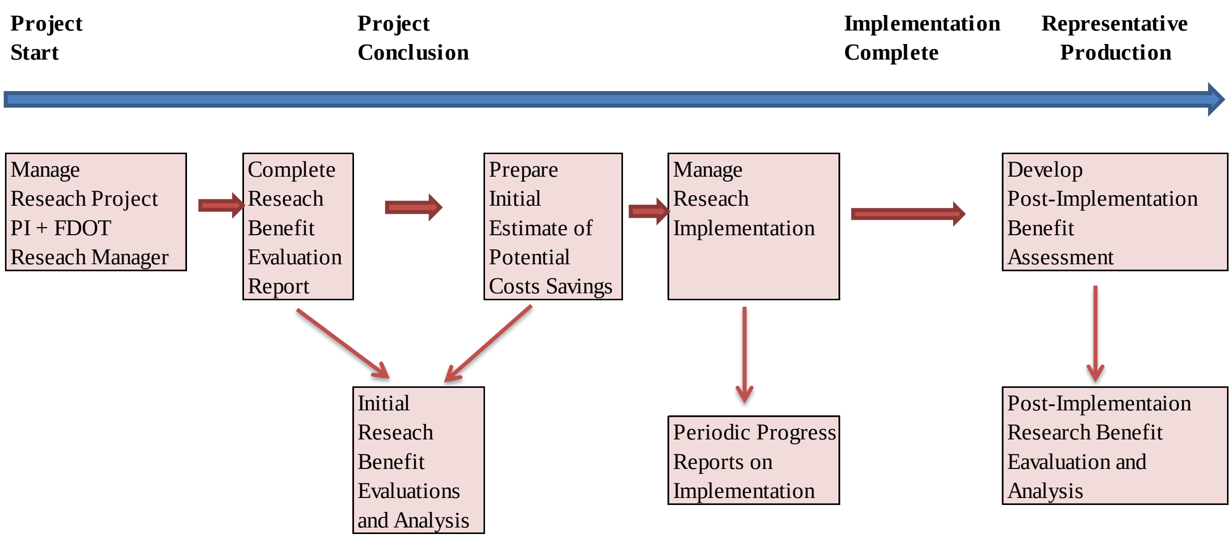

Figure 2: Research evaluation process proposed for the Florida Department of Transportation (Adapted from Southeast Transportation Consortium, 2014).

Another challenge is estimating the useful service life of a research product before becoming obsolete or being superseded by subsequent research projects or other changes in technology, policies, practices, specifications, etc. Estimating this "useful life" duration depends on many factors, including the category (e.g. pavement, bridge) and type (e.g. material, construction) of research product, experience and track record of similar innovations, local conditions and institutional aspects. An indication of the variability in estimating the useful life of various categories and types of research products is the range of useful life estimates reported by Krugler *et al*. (2006) and selected categories summarized in Table 1.

The values shown in Table 1 represent the number of responses (in each category) from a survey conducted at the 2004 meeting of the American Association of State Highway & Transportation Officials (AASHTO) Research Advisory Committee (RAC). The responses were subsequently used to develop recommended useful life values for research product evaluations in a tool developed in a National Cooperative Highway Research Program (NCHRP) project (Krugler *et al*. 2006) described later in this article. The highlighted cells in Table 1 indicate which useful life estimate range (column) contains the median. The entry for Standard Specifications refers to construction and maintenance.

Table 1. Useful Life Estimates of Research Product Categories (Adapted from Krugler)

Categories of Research Products (May be new approaches or improvements to existing ones)	**Useful Life Estimate (number of responses)**					
	< 3 Years	3 to 6 Years	7 to 10 Years	11 to 15 Years	16 to 20 Years	> 20 Years
Pavement Design Methods		1	5	3	3	2
Laboratory Test Methods		1	6	3	3	1
Field Test Methods for Pavements	1	1	7	3	1	1
Standard Specifications		6	5	1	2	
Quality Control / Assurance Methods	1	6	5	1		1
Construction Inspection Manuals	2	7				

Several recurring procedural challenges in evaluating benefits have been observed. Evaluators of research benefits in any sector must decide which significant aspect should be measured, how and when to measure and how to interpret results. Many benefits are difficult to characterize. Regardless of the approaches used, recurring challenges to the evaluation have been observed (Arjanovic *et al*. 2009),including:

- attributing impacts, which require linking outcomes to a specific research project and discerning previous research that influenced it as well as other projects that were influenced by it;
- setting boundaries, which require identifying the starting point of all contributing research (in retrospective studies) and identifying the timeframe and endpoint for analysis (in prospective studies, which are rarer);
- bias in the selection of research projects in case studies, e.g. projects with low payback vs. high payback;

- unclear descriptions of techniques in data collection and analyses, and
- non-uniform definition of terms and concepts, e.g. basic research vs. applied research.

One of the biggest procedural challenges is the attribution of impacts and benefits to a specific research project. A readily apparent linkage between research and identifiable benefits is more the exception than the rule. One review of international practice in assessing research impacts, characterized attribution of impacts as a main reason why such evaluations are usually considered "too hard" (Grant *et al*. 2009). Case studies that evaluate specific research have been successful in mitigating this challenge (Arjanovic *et al*. 2009, Grant *et al*. 2009). This approach focuses on evaluating a low number of benefits by examining a few where the pathway from research outcomes to benefits is evident.

Considering the emphasis of most publicly-funded transportation research on solving problems, the practical nature of research products that are implemented, and the long timeline needed to produce implementable research outputs that solve problems, it should not be surprising that evaluating benefits from research is a relatively low priority that remains an open field of study across all sectors (government, business and academia) worldwide.

Methods, measures and indicators

The term "benefits" as used in this article comprises the impacts (both positive and negative) from research outputs. Methods and measures for evaluating benefits generally are categorized as quantitative and qualitative. Quantitative measures have a numerical value, e.g. savings in dollars, travel times or lives, which can be viewed as objective. Qualitative measures are descriptive indicators without implicit numerical value and hierarchy, although numbers are sometimes assigned for analytical purposes. These do, however, reflect more subjective assessment such as satisfaction and quality.

Evaluation approaches offer advantages and disadvantages (depending on project-specific circumstances) that present various trade-offs. For example, a clear correlation of research outputs to benefits can be identified in a case study but may be difficult to determine using econometrics. However, broader assessment by econometric modelling can be difficult to do in a case study (RAND Europe, 2007). As mentioned above, selecting an appropriate method is context sensitive, reflecting specific needs such as determining whether a research approach should be terminated, whether results are within accepted quality standards and establishing whether there are direct economic benefits.

Descriptions about techniques to evaluate indirect benefits and larger-scale economic impacts of a project on the economy such as job creation, development of knowledge and understanding, are described in the literature and are not repeated here. Rather, this article focuses on direct benefits of APT in terms of qualitative and quantitative techniques.

Qualitative and Quantitative Techniques

US Congress, Office of Technology Assessment - OTA (1986)

Findings reported from the 1986 OTA study (US Congress, Office of Technology Assessment 1986) identified a variety of approaches used in industry and government at that time, which are still relevant. Qualitative evaluation was dominated by peer review. Quantitative evaluation consisted of a wider set of methods and measures as summarized in Table 2.

Table 2. Quantitative methods and measures to evaluate government and industry research funding (Adapted from OTA)

Category	**Methods and Measures**
Retrospective	
Economic (measures output in terms of productivity or dollars)	- Macroeconomic (production function) - Investment analysis ▪ Return on investment ▪ Cost-benefit analysis ▪ Rate of return ▪ Business opportunity - Consumer and producer surplus
Output (measures output in terms of published information)	- Bibliometrics ▪ Publication count ▪ Citations/co-citation analysis - Patent count and analysis - Science indicators (and others)
Prospective	
Project selection	- Economic models - Scoring models - Risk analysis and decision analysis - Portfolio analysis (constrained optimization)

National Academies Report (1999)

Several of the same approaches identified in the OTA study are prominent in a 1999 study by the National Academy of Sciences Committee on Science, Engineering and Public Policy (COSEPUP). COSEPUP conducted studies and workshops with Federal agencies, the research community, industry, states and agencies in other countries (National Academy of Science, 1999). The goal of this later effort was to identify and analyse the most effective ways in which to assess research results. The study identified advantages and disadvantages for each method. Like more recent studies, the COSEPUP findings suggest that a multi-faceted approach, which combines measures and indicators in a complementary manner, should enable analysis of outcomes and impacts from many types of research.

Transportation Research Board, NCHRP

Two years after the COSEPUP report, the Transportation Research Board's (TRB) National Cooperative Highway Research Program (NCHRP) Synthesis 300 reported on a review of approaches to measure the performance and effectiveness of transportation research and development (Sabol, 2001, Transportation Research Board Special Report 313, 2014). Results were reported from a survey of US state DOTs that found 25 percent used performance measures for projects after implementation. All measures relied on qualitative information, although respondents described them as quantitative. The report presents a synthesis from the literature and survey responses from state DOTs, private sector research programs and academia.

Results from this NCHRP study showed substantial differences in the approaches to and concerns about assessing benefits of research in public, private and academic sectors. The study found that most state DOTs were not satisfied with their cost-benefit techniques and that many issues associated with establishing benefits would need to be resolved in order to provide useful and reliable information over a long-term assessment period. Quantifying benefits via cost-benefit analysis in the private sector was also problematic, but aided by higher quality and more extensive cost data as well as a more customer-driven business environment. Less quantification of benefits of research in the academic sector was attributed to greater concern about quality and productivity. The study reported that neither college faculty nor administrators frequently discuss cost-benefits of academic research activities.

NCHRP Synthesis 300 found quantitative assessment to be a much sought after priority but one that faced many challenges, including the need for better quality cost data, more clearly identified benefits and clearer links of research outputs to benefits. The study also found peer review to be the standard for qualitative assessment (Sabol, 2001).

The synthesis also reported that state DOTs were not satisfied with their cost-benefit approaches, with the establishment of benefits being a main concern that must be resolved. Another high priority was the need for a measure of payoff from implementation. The study found that existing measures of payoff varied substantially and were neither rigorous nor robust. The report found a need for research to provide guidance on the use of cost-benefit analysis, that no existing method was clearly superior and that techniques of estimating benefits compatible with cost-benefit analysis should be emphasized.

Comparisons from Published Reviews

Several reviews of evaluation techniques have been published in the past two decades. Categories and findings from selected reviews (discussed below) are summarized in Table 3, which shows (with check marks) methods described in publications from Europe and the USA. The European publications, both by RAND Europe (an independent non-profit research institute), come mainly from literature reviews that include earlier US reports (RAND, 2007, RAND 2009).

Table 3. Summary of Evaluation Techniques (References appear in parentheses)

	Transportation		Non-Transportation	
	USA (FHWA)	USA (NCHRP)	Europe 2009 (RAND)	Europe 2007 (RAND)
Methods – Qualitative:				
Peer and Expert Review	✔		✔	✔
Survey	✔	✔		✔
Case Study – Descriptive			✔	
Training and Education		✔		
Tracing and Logic Modelling			✔	✔
Benchmarking				✔
Sociometric Analysis			✔	
Methods – Quantitative:				
Cost-benefit / Savings analysis	✔	✔	✔	✔
Bibliometrics	✔			✔
Safety (less crashes/fatalities)		✔		
Econometrics			✔	
Outputs (products and reports)		✔		
Performance	✔	✔		

The summary in Table 3 leads to several observations including the following:

- Qualitative and quantitative techniques are both well represented;
- Many techniques are cited in at least two publications;
- The most common methods are cost-benefit/savings analyses, peer reviews and surveys;
- These common methods are used in transportation research as well as non-transportation research.

Evidently, a wide variety of methods are in use. The choice of approach is driven by the purpose and conditions of the study as well as time, resources and other constraints. Each technique offers advantages and disadvantages. It is evident from the different observations and insights provided from various methods why analysts face challenging trade-offs in choosing an approach that is best suited for evaluating benefits in a specific study. Advantages and disadvantages associated with various techniques are outlined in Table 4 (Arjanovic *et al*. 2009).

Table 4. Advantages and Disadvantages of Various Techniques (Adapted from Arjanovic)

Method	Brief Description	Advantages	Limits
Survey	Asking multiple parties a uniform set of questions about activities, plans, relationships, accomplishments, value or other topics, which can be statistically analysed	> Provides an economical way to gather information about a programme and its participants that is not available through other sources > Accommodates the use of control and comparison groups or the collection of counterfactual information > Usually, diverse audiences can understand the approach and results	> Phone interviews work best when timeliness is important Mailed questionnaires often have low response rates > Does not provide the richness of individual project detail that stakeholders tend to find interesting > Responses are often subjective in nature and respondents may not be truthful
Case study – descriptive	Investigating in-depth a programme or project, technology or a facility, describing and explaining how and why developments of interest have occurred	> Many decision-makers read and process anecdotal cases more easily than quantitative studies > Provides richness of detail > Can be used to identify best practice experience	> The anecdotal evidence provided is generally considered less persuasive than quantitative evidence > The results of one or more individual cases may not apply to other cases > Can be difficult to aggregate findings
Case study – economic estimation	Adding to a descriptive case study quantification of economic effects, such as through cost–benefit analysis	> Focuses on ultimate outcomes and impacts rather than on outputs > Provides quantitative estimates of results > Uses financial methods	> The value of important benefits may not be estimated in monetary terms > Needs to be carried out a long time after the project has finished
Econometric and statistical analysis	Using tools of statistics, mathematical economics and econometrics to analyse functional relationships between economic and social phenomena and to forecast economic effects	> Produces quantitative results with detailed parameters > Demonstrates cause-and-effect relationships	> Difficult for non-specialists to understand, replicate and communicate > Not all the effects can be captured in these highly quantitative methods

Table 4. Advantages and Disadvantages of Various Techniques Continue

Method	Brief Description	Advantages	Limits
Sociometric and social network analysis	Identifying and studying the structure of relationships by direct observation, survey and statistical analysis of secondary databases to increase understanding of social or organisational behaviour and related economic outcomes	> Focuses on the understanding of the process of innovation > Requires relatively modest inputs which can be obtained through survey, interview or existing databases	> Remains largely unfamiliar to most economists and programme stakeholders > Results may not be very informative on a programme's performance
Bibliometric analysis	Use data on numbers and authors of scientific publications and on articles and the citations therein (and in patents) to measure the output of individuals or research teams, institutions and countries, to identify national and international networks and to map the development of new (multidisciplinary) fields of science and technology	> Widely applicable to evaluation of programmes with an emphasis on publishing or patenting > Can address a variety of evaluation topics, including research output, collaborative relationships and patterns and intensity of knowledge dissemination > Diverse audiences can understand the results > Can be applied to a programme with a relatively short time-lag > High degree of credibility	> Treats only publications and patents as programme outputs and ignores other outputs and long-term outcomes > Time must pass before extensive publication and/or patent citations can be observed > Counts indicate quantity and impact of output, not quality > Not all publications and patents are of equal importance > The propensities to publish and patent differ among scientific and technological fields
Historical tracing	Tracing forward from research to a future outcome or backward from an outcome to precursor contributing developments	> Produces interesting and credible studies documenting a chain of interrelated developments > Sheds light on process dynamics	> Chains of events tend to be highly complex with many organisations and researchers involved
Expert judgement	Using informed judgements to make assessments	> Provides a relatively quick, straightforward, feasible and widely-accepted approach to assessment > Offers the opportunity for an exchange of ideas which can lead to new perspectives	> Not much is known to the quality or accuracy of expert judgement as applied to R&D programme impact assessment

Tables 3 and 4 suggest that evaluations of research may be more representative if they combine qualitative and quantitative information. Some analysts suggest that relying on one indicator can mislead analysts and decision-makers (Ruegg and Jordan, 2007). In addition to using more than one technique, some investigators have recommended using many sources of information as well as several separate investigators to evaluate benefits in a technique referred to as "triangulation" (Arjanovic *et al*. 2009).

Approaches that enable characterisation of benefits using more than one measure have been evolving in recent years with the development of "toolbox" or "toolkit" frameworks, which consist of many measures and methods. Toolkits have been developed in the US and are in different stages of implementation for both transportation research (Krugler *et al.* 2006) and non-transportation federal research in energy at the Department of Energy (Ruegg and Jordan, 2007), and technology development at the National Institute of Standards and Technology (NIST) (Powell, 2006; Ruegg and Feller, 2006). A European review (summarised in Table 3 under "Europe") highlights the NIST toolkit and refers to it as "one of the most influential reference works, practical aids and planning guides for practitioners of research evaluation" (Arjanovic *et al*. 2009). The NIST toolkit was developed from extensive evaluations of 45 NIST research projects between 1990 and 2000.

Transportation Research

Federal Research Case Study

In assessing the federal investment in infrastructure research and development from 2006 to 2009, TRB Special Report 295 observed that evaluations of past Federal Highway Administration (FHWA) research in materials and structures found substantial savings and extension of service life far in excess of the cost of the research (Research and Technology Coordinating Committee, 2008). The observation is based largely on retrospective evaluations of benefits from FHWA-sponsored research.

The FHWA's 2003 report observed that estimating cost savings was "…the most demanding part of the assessment…". However, those projects for which data were obtained resulted in very high agency, road user and safety cost savings. The report estimated costs savings at a national level that was more than ten times the annual research funding (Federal Highway Administration, 2003).

Under the theme of "Performance Assessment", the Transportation Research Board Special Report 313 (2014) it is recognised that in the USA, Europe and Asia-Pacific there is a general trend towards evidence-based decision-making and little evidence is publicly available on the requirements' impacts, whether positive, negative or neutral. Although there are tools used in prioritising research and development activities such as rate of return on investment, cost-benefit analysis and bibliometrics (citation analysis, content analysis) it highlights the need for identifying the correct methods and measures used in setting research and development priorities, the time frames involved and to what extent performance assessment of prior investments influences decisions regarding future research investments.

NCHRP Project 20-63B Toolbox (NCHRP, 2016)

A relatively recent addition to the toolkit approach is the NCHRP Project 20-63 toolbox, which is summarised in the column under the heading "US" in Table 3. The main objectives of the NCHRP project were to define performance measures for transportation research projects and to assemble a useful and practical toolbox of performance measures (with examples) for use by state DOTs. After completing a literature review, surveys were conducted of state DOT staff and managers as well as federal and private sector research managers. Survey respondents rated their organisations' experience with each measure as well as the perceived value of each measure in their organisation. The study identified 30 performance measures to include in the toolbox.

Return on investment or benefit-cost ratio (BCR) tied for third rank (tied with agency cost savings), just following behind lives saved and reduction in crashes.

The toolbox contains 30 performance measures automatically programmed in the Research Performance Measurement (RPM) software available for state DOTs. The RPM software allows users to import other performance measures, enabling customisation of the toolbox to meet a DOT's specific needs. The performance measures, most of which are quantitative, are categorised under five major headings as follows:

- Outputs – products, research reports published and graduate students
- Outcomes – agency cost savings, lives saved and reduction in crashes
- Stakeholders – customer satisfaction and input
- Efficiency – BCR, percentages of projects on time, within budget, implemented, etc.
- Resource allocation – funding and contractor issues, quality of life and safety projects

NCHRP Project 20-63 was completed in 2010 and was followed by Phase II (designated NCHRP 20-63B), which began in July 2010. Having developed and established electronically formatted tools for evaluating research as described above in the initial project, Phase II focused on enhancements and refinements in functionality of the system as well as ongoing maintenance of a website. The ultimate goal was to expand access to the system for routine use by state DOTs. Phase II is scheduled to be completed in 2016 (National Cooperative Highway Research Program, 2016).

Southeast Transportation Consortium (2014)

A comprehensive literature review regarding the determination of the benefits of transportation research has been done through three fact-finding surveys in which twenty USA states as well as the FHWA and TRB participated (Southeast Transportation Consortium, 2014). Different methods are used by transportation agencies depending on the transportation focus area such as safety, traffic congestion reduction, engineering design improvement, materials and pavements, increased service life, etc. No single method stands out as the preferred one, but in the case of materials and pavement, the value of research is measured through parameters such as pavement reduced construction, lower operations and maintenance costs. Table 5 summarises these methods along with areas of benefits for which these methods have been utilised to determine the value of research. It can be seen that BCA and benefit (monetary value) analysis are widely used methods to determine the value of research across all focus areas.

The Connecticut Academy of Science and Engineering (2013)

A report released by The Connecticut Academy of Science and Engineering (2013) analysed the economic impacts of transportation projects. Eighteen different analytical tools for analysing economic impacts of transportation investments were reviewed. Although this study is not directly aligned with the topic of measuring the success of transportation-related research, there are significant similarities in the report findings in comparison with the other previous studies investigated. Two types of interrelated analyses are suggested: those for estimating economic impacts (prospective analysis) and those for evaluating economic impacts (retrospective analysis). BCA is suggested for evaluating transportation investments because it captures the costs and most direct benefits of a transportation investment to the society at large. It concludes that typically only direct economic impacts of transportation investments are included in the BCA while the indirect economic impacts are often ignored due to the difficulty associated with measuring these and uncertainty associated with realizing the impacts.

Table 5. Preferred benefit analysis methods for various focus areas in transportation (Southeast Transportation Consortium, 2014)

Transportation Focus Area	Benefit Analysis								Benefit ($) Analysis	Benefit ($) / Cost ($) Analysis	Life cycle Cost Analysis
	Before and After Study	Statistical Analysis	Simulation Analysis	Assumption based Estimation	Field Exp	Lab Exp	Surveys	Benefit in other areas			
Safety	■	■	■	■	■				■	■	
Environment Sustainability Safety	■		■	■	■	■			■		
Improved Productivity and Work Efficiency				■	■				■	■	
Traffic and Congestion Reduction	■		■		■				■	■	
Reduced Construction Operation and Maintenance Costs									■	■	■
Customer Satisfaction					■		■			■	
Engineering Design Improvements		■	■					■	■		
Increased Service Life		■			■	■			■		■
Reduced User Cost								■			
Reduced Administrative Cost										■	
Materials and Pavements								■			

Conclusions and Recommendations

The methods, measures and indicators discussed in this article show substantial variability in approaches used worldwide to evaluate benefits of research in- and outside of transportation research. No universal approach is recommended because there is no "one size fits all" technique. Despite the recurring observation that no country appears to have a totally satisfactory technique, many approaches have been proposed, applied and reported. Developments during the past decade appear especially promising.

In the case of APT related research, there are qualitative and quantitative, direct and indirect benefits. The growing global interest and awareness of efforts to quantify the economic benefits of APT research was the main theme at the 2008 International APT Conference in Madrid, Spain (Du Plessis and Prozzi, 2008). Conference discussions explicitly associated technical activities with their relative costs and benefits, which are suitable for BCA. In the case of calculating cost savings (better pavement designs, construction processes and materials due to APT results), BCA is the ideal method to measure the impacts and benefits of APT related research. The key component of this method is obviously market uptake and the acceptance of new technologies. Case studies are suggested to prove a concept and the real benefits can be measured only after implementation on a larger scale.

It is suggested that all measurable parameters mentioned in Table 3 should be captured during APT experiments. Retrospective analysis of both qualitative and quantitative benefits will only be possible if quality information was gathered and kept for each APT experiment including information on implementation projects. BCA and positive benefit-cost ratios are powerful convincing tools to justify expensive research programs (such as APT), while bibliometrics, number of PhDs, peer reviewed articles, patents, etc. highlight the importance of APT in academia and political circles.

The authors hope this article encourages further developments in a balanced approach that provides practical improvements, e.g. the use of evaluation methods for retrospective assessment as well as for prospective analysis to aid in research project selection when setting an APT research programme portfolio and in strategic planning for APT. The ultimate goal is to help better understand, demonstrate and communicate the benefits of APT.

Acknowledgements

The authors would like to acknowledge the financial support of California Department of Transportation (Caltrans), in particular, the Caltrans Library for assistance in acquiring references and Mr. W.A. Nokes from Caltrans Division of Research, Innovation and Systems for his input, guidance and support.

References

Arjanovic, S., Hanney, S., and Wooding, S., 2009. A Historical Reflection on Research Evaluation Studies, their Recurrent Themes and Challenges. Technical Report 789, *Project Retrosight, RAND Europe,* Cambridge, United Kingdom, 2009. www.rand.org/pubs/technical_reports/TR789/. [Accessed 29 February 2016].

Baron, J. and Gurmankin. A., 2013. *Cost-benefit analysis can increase trust in decision makers*. University of Pennsylvania. www.sas.upenn.edu/~baron/papers.htm/cba.html. [Accessed 29 February 2016].

Du Plessis, L., *et al*., 2011. A Case Study of Economic Benefits Assessment of APT Research in California. *Transportation Research Record, TRR225,* Washington, D.C.

Du Plessis, L., and J. Prozzi, 2008. A logical framework approach to the evaluation of benefits derived from Accelerated Pavement Testing (APT) studies. 3rd International Conference on Accelerated Pavement Testing – APT08, Madrid, Spain, Workshop #3, October 2008.

Elston, D., *et al.*, 2009. Transportation Research Program Administration in Europe and Asia, Final Report. FHWA-PL-09-015, *FHWA, U.S. Department of Transportation*, http://www.international.fhwa.dot.gov/pubs/pl09015/. [Accessed 27 January 2016].

European Co-operation in the Field of Scientific and Technical Research (COST), 2005. Work Package 5: Future use of ALT. Report on Improvements in Pavement Research with Accelerated Load Testing. *COST Work Package 5 Final Report.*

Federal Highway Administration (FHWA), 2003. Synthesis of R&D Benefits Case Studies. *Office of Research, Development and Technology, FHWA, U.S. Department of Transportation*.

Grant, J., *et al*., 2010. Capturing Research Impacts. A Review of International Practice. *RAND Europe, Documented Briefing DB-578-HEFCE*, Cambridge, United Kingdom, 2010. www.rand.org/pubs/documented_briefings/DB578. [Accessed 27 January 2016].

The Connecticut Academy of Science and Engineering (2013). Analysing the Economic Impacts of Transportation Projects. Report CT-2279-F-13-13. Connecticut Department of Transportation, Newington, USA

The International Benefits, Evaluation and Costs (IBEC) Working Group, 2014. https://sites.google.com/site/ibecits/home. [Accessed 19 February 2016].

Krugler, P., *et al*., 2006. Performance Measurement Tool Box and Reporting System for Research Programs and Projects. *Highway Cooperative Highway Research Program (NCRHP).* Web-Only Document 127, Contractor’s Final Report for NCHRP Project 20-

63. Transportation Research Board of the National Academies, Washington, D.C., http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w127.pdf. [Accessed 1 February 2016].

National Academy of Science, 1999. Evaluating Federal Research Programs: Research and the Government Performance and Results Act. Committee on Science, Engineering, and Public Policy. *National Academy of Sciences, National Academy of Engineering, Institute of Medicine,* Washington, D.C.

National Cooperative Highway Research Program (NCHRP), 2016. Performance Measurement Tool Box and Reporting System for Research Programs and Projects. NCHRP 20-63B (Phase II), *Transportation Research Board of the National Academies*, Washington, D.C. http://slidegur.com/doc/4156092/performance-measures-for-research-programs- [Accessed 1 February 2016], https://www.rpmweb.org/rpm/. [Accessed 1 February 2016].

Nokes, W.A., Du Plessis, L., Mahdavi, M., Burmas, N., 2011. Techniques and Case Studies for Evaluating Benefits of Accelerated Pavement Testing. *Transportation Research Record, TRR225,* Washington, D.C.

Powell, J., 2006. Towards a Standard Benefit-Cost Methodology for Publicly Funded Science and Technology Programs. Report NIST IR-7319, *National Institute of Science and Technology (NIST)*, Gaithersburg, Maryland, USA.

RAND Europe, 2007. Measuring the Benefits from Research. *Research Bulletin 9202,* Cambridge, United Kingdom. www.rand.org/pubs/research_briefs/2007/RAND_RB9202.pdf. [Accessed 26 January 2016].

RAND Europe, 2009. A historical reflection on research evaluation studies, their recurrent themes and challenges from Research. Technical Report TR-789-RS, Cambridge, United

Kingdom. http://www.rand.org/pubs/technical_reports/TR789.htm [Accessed 26 January 2016].

Transportation Research Board Special Report 295, 2008. The Federal Investment in Highway Research 2006-2009: Strengths and Weaknesses, Research and Technology Coordinating Committee *Transportation Research Board of the National Academies*, Washington, D.C.

Transportation Research Board Special Report 313, 2014. Framing Surface Transportation Research for the Nation's Future. National Research Frameworks application to Transportation Coordinating Committee. *Transportation Research Board of the National Academies*, Washington, D.C.

Ruegg, R. and Feller, J. 2006. A Toolkit for Evaluating Public R&D Investment, report NIST GCR 03-857, *National Institute of Science and Technology* (*NIST),* Gaithersburg, Maryland, USA, 2006.

Ruegg, R., and Jordan, G., 2007. Overview of Evaluation Methods for R&D Programs. *Report prepared by TIA Consulting, Inc. for Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy*, March 2007.

Sabol, S.A., 2001. Performance Measures for Research, Development and Technology Programs. *NCHRP Synthesis 300, Transportation Research Board of the National Academies*, Washington, DC.

Southeast Transportation Consortium, 2014. Synthesis of best practices for determining value of research results. Final report LTRC Project No. 12-3PFt., *Southeast Transportation Consortium (STC) and the Louisiana Transportation Research Center (LTRC)*, 2014.

US Congress, Office of Technology Assessment (OTA), 1986. Research Funding as an Investment: Can We Measure the Returns? – *A Technical Memorandum., OTA-TM-SET-36 (NTIS #PB86-218278),* Washington, D.C.

US Department of Transportation (US DOT), 2016. Research and Innovative Technology Administration (RITA). http://www.benefitcost.its.dot.gov/its/itsbcllwebpage.nsf/krhomepage. [Accessed 28 January 2016].