International Journal of Project Management 40 (2022) 703–714

Contents lists available at ScienceDirect

International Journal of Project Management

journal homepage: www.elsevier.com/locate/ijproman

Public project success? Measuring the nuances of success through ex post evaluation

Gro Holst Volden [a,*], Morten Welde [b]

[a] Norwegian University of Science and Technology (NTNU), Trondheim 7491, Norway
[b] Norwegian University of Science and Technology (NTNU), Norway

ARTICLE INFO

Keywords:
Public projects
Project success
Ex post evaluation
Stakeholders
Citizen and taxpayer view

ABSTRACT

Public projects are conducted on behalf of citizens and taxpayers, who may have different views of what success looks like. The authors argue that the definition of success needs to be broad and multifaceted, even more in public than private projects. A generic six-criteria model is suggested, which covers project success on three analytical levels, from various valuation perspectives, and intended and unintended impacts alike. The model is used to evaluate 34 projects some years into their operational phases. The findings suggest that public projects are often more successful than people think. For example, cost performance is largely acceptable, contrary to the impression presented by the media and some academic studies. We also demonstrate how projects can be successful in some respects yet unsuccessful in others. We argue that the media, the most important source of information for members of the public, has a narrow definition of success and is negatively biased. It seems that ex post evaluation applying a standardized and multifaceted framework, provides a good basis for learning and improvement, to enhance the success of future projects on all levels.

1. Introduction

The public sector accounts for a large proportion of economic activity in most developed countries, and public investment levels are historically high. Public investments are often recognized as powerful catalysts for economic recovery and renewal, and for implementing long-term policies, such as investment in green energy infrastructure to support action on climate change. In 2019, public investment spending averaged 3.3% of GDP (gross domestic product) across the member countries of the Organisation for Economic Co-operation and Development (OECD). In 2020, public investment relative to GDP increased in 25 of the 26 OECD countries, mostly as a response to the COVID-19 pandemic (OECD, 2021).

Most of these investments are organized as projects. A project is a temporary endeavour to deliver specific objectives based on final deliverables that are constrained to a finite timescale and budget (Murray-Webster & Dalcher, 2019). For the objectives to be met, the projects must be delivered well (Meggs, 2018). The purpose of projects may be specific, but in most public projects the intention is to deliver benefits both in the short and long term and to a range of stakeholders. In such cases, the benefits of the project must be considered from a broader societal perspective (Samset & Volden, 2016) which calls for a social appraisal that incorporates all citizens' interests.

There are fundamental differences between private and public projects. Although both types of projects are means for change and may have a wide range of internal and external effects, the success of private projects is normally evaluated by using a narrow set of metrics, usually related to financial profitability, and the most important stakeholders are the owners/shareholders and employees of companies. These are kept up to date through quarterly and annual reports. Public projects are implemented on behalf of society and all taxpayers as means of implementing *policies*. However, there may be different views on what the purpose of these policies is. Furthermore, the planned benefits may be fuzzy or difficult to quantify, and a large proportion of benefits are not traded in markets. Also, it is more difficult to ensure accountability in public projects, and the governance of such projects may be less transparent. The terms of funding (by the government) may in itself reduce stakeholders' incentives to opt for the most socially beneficial or cost-effective solutions (see Volden, 2018a, on perverse incentives). Poor choices rarely have personal consequences for those responsible, and public organizations do not risk bankruptcy even if they regularly fail in their investment strategies. The ultimate owners of public projects can

* Corresponding author.
E-mail address: gro.holst.volden@ntnu.no (G.H. Volden).

https://doi.org/10.1016/j.ijproman.2022.06.006
Received 1 February 2022; Received in revised form 26 May 2022; Accepted 29 June 2022
Available online 15 July 2022

only carry out their version of corporate governance indirectly through elections.

The project management literature has long been concerned with developing a better understanding of the nature of project success (Pinto et al., 2021). During the last two decades, attention has turned from project delivery to the goal and purpose of projects and how they create benefits for various parties (Breese et al., 2015; Hjelmbrekke et al., 2017; Ika, 2009; Morris, 2009; Shenhar & Dvir, 2007; Zwikael & Meredith, 2021; Zwikael & Smyrk, 2012). Shenhar and Dvir (2007) offered a compound definition of project success: (1) meeting time, budget, and other requirements (efficiency), (2) impact on team, (3) impact on the customer, (4) benefit to the performing organization, and (5) preparation for the future. Similarly, Davis (2014) suggested categorizing stakeholders into three main groups – senior management, project core team, and project recipient – based on the groups' perceptions of project success. However, no universal measure of success exists, and the term is still not fully understood. The definitions of overlapping concepts, such as project benefits, value, and value creation, may also be unclear and multifaceted (Laursen & Svejvig, 2016). The confusion is even greater for public projects when the funding party, and thus the owner, is the whole society. The cost-benefit analysis has been much used to assess the value of public projects, but is based on a somewhat narrow definition of strategic success, and is not always trusted by political decision-makers (Volden, 2019). We need a wider set of criteria that covers more aspects that public project owners care about. Generally, little attention has been paid in the project management literature to public policy aspects and the role of the owner (Sanderson & Winch, 2017).

Another gap in the literature, which is related to the one discussed above, concerns how little we know about actual public project success in developed countries. Many studies have demonstrated public project success as measured by partial success factors such as costs (Cantarelli et al., 2010; Love et al., 2012), demand (van Wee, 2007) and even benefit-cost-efficiency (normally limited to the transport sector) (Kelly et al., 2015), but it is unclear whether the basis for comparison is the same, and measurement of the full spectrum of project success is rare. We simply do not know whether or not citizens and taxpayers have reason to view public projects as successful.

Certainly, multidimensional evaluation models exist, such as the Five Case Model in the UK (HM Treasury, 2013), three-dimensional sustainability impact assessment (OECD, 2010), and other multi-criteria decision models. However, such broad models are almost exclusively being used ex ante, based on forecasts and assumptions, and not ex post based on data and experiences. We argue that a good definition of project success should be accompanied by a framework for determining whether the projects have been successful. This, in turn, requires use of the same success criteria ex ante and ex post. Ideally, the model should also be generic enough to allow for comparison across projects. As far as we know, such models are rare or non-existent in use.

On the other hand, there is a lot of reporting on public projects in the media as well as some academic literature. Based on this, the performance of public projects may seem rather dismal. Headlines that report cost overruns and extreme delays are not uncommon, and public investigations are launched in the worst cases. Professor Bent Flyvbjerg, known for studying public project performance (especially cost performance) across the world, proposed the 'iron law of megaproject management' where he claimed that megaprojects are 'over budget, over time, over and over again' (Flyvbjerg, 2011). In subsequent publications, the author has repeatedly alleged that estimates of benefits and costs in public investments are highly inaccurate and biased (e.g. Flyvbjerg & Bester, 2021).

This paper builds on the two interweaved discussions mentioned above, namely the discussion of the definition of project success, and whether public projects are successful or not.

We draw partly on literature reviews and desk studies, and partly on empirical data from longitudinal project evaluation work carried out by the Concept Research Programme in Norway. The programme was initiated in the year 2002 and has since then followed more than 200 projects under the Ministry of Finance's public project governance scheme throughout their appraisal and planning processes, and eventually, many of them are completed and have reached their operational phases. A main ambition for the researchers has been to get to grips with the choice of concept (hence the name Concept Research Programme) and secure the strategic performance (i.e., impact on society) of projects. This is a complex matter and an interdisciplinary field. In this paper, we build on all of this material, and in particular on work done to design a broad evaluation framework and systematically use it to evaluate projects after they are completed and some years into their operational phases.

In the context of assessing public project success using ex post evaluations of Norwegian projects, we defined three research questions:

1. What is public project success?
2. How successful are public projects?
3. Why does the public think public projects always fail?

Project success has been discussed in the professional literature for decades and the public discourse on projects' contribution to matters such as economic efficiency and social and environmental sustainability suggests that we are far from consensus. Our contribution to this area lies in using empirical data to illustrate the multi-faceted nature of project success to bridge the gap between theory and practice and to improve the ex ante and ex post assessment of project success. Increased knowledge in this area should be beneficial to both project owners and to other stakeholders such as the wider public whose impressions of project performance may be influenced by one-sided media reporting.

We argue that there is a need for a wide set of success criteria, and present a goal-orientated evaluation model aimed at measuring success on different analytical levels, from different valuation perspectives and including intended and unintended impacts alike. Based on the framework, we then illustrate how projects compare. To answer the second research question, we build on and extend the contributions of Volden (2018b) by using more data and presenting more robust results. We argue that generally, public projects are often more successful than people think. We conclude that there is a need for a more nuanced view of project success than what can be measured through partial analyses. In the final section of the paper, we discuss why the public, who is also, ultimately, the funding party, apparently thinks public projects 'always' fail. We discuss issues such as selection and publication bias and argue that both the academic community and the press should strive to disseminate both successes and failures and not over-focus on the latter. In the end, we provide some concluding comments.

2. What is public project success?

Historically, in the project management community, the tendency was to focus narrowly on the iron triangle of cost, schedule, and quality. Samset and Volden (2016) referred to this 'narrow' perspective as one of the big paradoxes of project management and governance, since projects' impact on users and society should be far more important. Today, there is consensus in the literature that project success is multidimensional (Pinto et al., 2021). However, what does this mean for public investment projects?

2.1. *Why governments carry out big projects*

According to standard economic theory (e.g. Varian, 2010) governments provide services that are not being provided by free markets due to market failure (e.g. externalities, public goods, collective action problems). Further, governments may have ambitious goals for the development of society, for example in terms of economic development, or becoming a global leader in some areas. Governments are also

concerned with the distribution of goods – they may favour specific stakeholders or geographic regions, and most governments have a special concern for poor and vulnerable groups. Hence, profit motives are of secondary importance.

We suggest there are three main categories of public investment purposes:

- Security and other essential needs. Governments invest to meet the basic needs of the population. Economists define *public goods* as goods that are underproduced by markets, as they are commonly available to all (i.e. nobody *can* be excluded from using them) and the cost of providing them to an additional user is zero (i.e. nobody *should* be excluded from using them). Military defence and flood protection are examples of public goods. A related term is *merit goods.* These are goods that governments think people should have access to, based on some concept of necessity, rather than the ability to pay. This includes basic health services, education, and social and welfare services, and facilities that support these services are referred to as social infrastructure (Love et al., 2019). Both public and merit goods are mostly administered by governments and largely paid for through taxation.
- Investments to facilitate economic development (also referred to as economic infrastructure). Some types of infrastructure, such as transport, communication, and energy supply systems, are vital for the functioning of the economy, in that they increase and connect markets, and reduce transaction costs. Hence, such investments facilitate economic growth and improve the nation's competitiveness. Likewise, public investment in research and development is a way to promote innovation and future growth. It is sometimes argued that the level of economic investment is too low in many countries and regions (McKinsey Global Institute, 2016; Zachariadis, 2018). In a rapidly changing world, new trends and policies, such as climate change and the energy shift, digitalization, and urbanization, call for new types of investments. Further, in the short run, the need to stimulate the economy and create jobs is also sometimes used to justify big public projects.
- National pride and prestige. There are numerous examples of investments made to demonstrate greatness and strength or to honour a leader. How many military and space programs have not been implemented as part of a competition amongst nations? Even today, in lists of the world's largest construction projects (Rodriguez, 2019), many of the projects will break world records (e.g. largest, tallest, fastest). Flyvbjerg (2014) refers to the four sublimes – *political, technological, economic*, and *aesthetic*, and Frey (2016) adds *community pride*, which may explain the increased size and frequency of megaprojects initiated by governments.

The purpose of the investment is crucial – it is the reason why a project is established and the basis for defining more specific goals and targets. In addition, governments will also emphasize economic and financial considerations such as value for money and fundability.

The public sector is organized differently in different countries. Being the funding party does not imply that the government needs to be the implementing party. Normally, a sectoral ministry is an owner, while project delivery is entrusted to a public agency, or a private enterprise through some form of Public-Private Partnership arrangement. Public projects normally face special requirements related to the tendering process.

2.2. Multiple analytical levels of project success

The distinction between *project management success* and *project success* (Baccarini, 1999) was a significant advance in the project management community's understanding of project success. The former is concerned with the efficient delivery of an agreed output whereas the latter is concerned with the wider outcome generated by the project. The two dimensions may be related, and some see project management success as a precondition for project success. But whereas project management success depends heavily on the project manager, project success depends on project funders, owners and others involved in project selection and design. *Ceteris paribus*, project management success is neither a necessary nor a satisfactory condition for project success, as noted by Ika (2009).

Some have suggested dividing project success into more than two levels. Zwikael and Smyrk (2012) suggested a triple-test performance measurement framework for project success, comprising (1) project management success, (2) project ownership success (achievement of the business case), and (3) project investment success (return on investment), where the first is judged by the project owner and the second two by the funder.

While most authors either discuss private sector projects or suggest models that are implicitly best suited for private projects, Samset (2003) suggested a three-level framework for public projects, including the operational, tactical and strategic level of project success, as illustrated in Fig. 1. The three levels concern (1) project delivery, (2) achievement of agreed goals, normally related to user benefits, and (3) societal effects, respectively. As illustrated, higher levels of success are associated with higher levels of uncertainty and a longer timescale.

Volden and Andersen (2018) discussed how project governance arrangements could be established at different levels in the public sector hierarchy. A typical infrastructure project under the auspices of a government agency will have project owners on three levels: the agency, the responsible ministry, and the cabinet. Each owner conducts planning and introduces governance arrangements to ensure that projects are aligned with their goals and strategies. Owners at higher levels have wider goals and strategies than those on lower levels: for example, whereas the goal of the Public Roads Agency is to build highways efficiently, the Ministry of Transport is concerned with mobility for users, whereas the Cabinet should balance the need for mobility against other considerations when planning for the desired societal development.

2.3. Multiple valuation perspectives

Public project success can be viewed not just at different analytical levels; we argue that it can be assessed also from different valuation perspectives. Some people emphasize economic goals for society, whereas others value environmental or social concerns higher. Elkington (1999) suggested operationalising these conflicting interests through the triple bottom line which seeks to consider and balance economic, environmental and social concerns. Haavaldsen et al. (2014) used the term three pillars of sustainability and argued that these three perspectives should be assessed separately and that they could not be summarised in a single metric. The same authors also argued that the three perspectives should be sorted by analytical level, as '*we need to acknowledge the difference between doing the projects more sustainably and choosing more sustainable projects*' (Haavaldsen, 2014, p. 5).

We argue that it is particularly at the highest analytical level (i.e., the strategic level of success) that different valuation perspectives are needed, of which value for money, sustainability (all three pillars), and the decision-maker perspective (goals defined by the incumbent government) are all relevant.

In a private-sector project, the ultimate owners are the shareholders of the company. Their strategic goal is 'simple' in principle: to maximize the return on investment in the long run. And admittedly, this may involve strategies to take social responsibility and be "environmentally friendly", but only to the extent that such measures can help the company become more competitive, and thus increase its profits, in the long run.

In a public project, the funding party is multi-headed, in that each citizen and taxpayer may have a different view on what makes a good investment (Klakegg & Volden, 2017). Citizens on the left side of the political scale may, for example, have a more positive attitude to

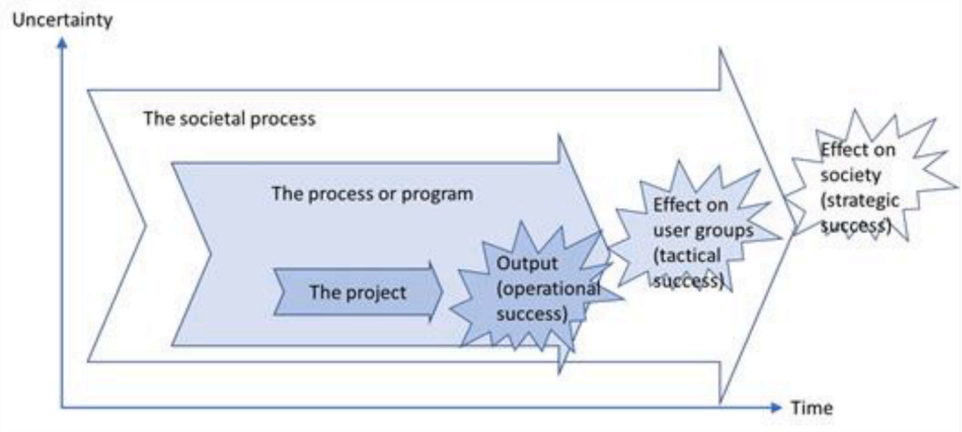

Fig. 1. Three levels of project success, inspired by Samset (2003).

investments with social and environmental purposes, while citizens on the right may prefer investments in economic infrastructure, and some would probably want to avoid government-funded projects altogether. The challenge of measuring and aggregating citizens' preferences for public projects is considerable (unless a referendum is held on each project).

One possible approach would be to define strategic success by the incumbent government's stated goals and strategies. In a democracy, the government is elected by the people, hence the government's goals and strategies should be supported by the majority of the population. However, public projects tend to last for years and decades, while governments come and go. Moreover, there will always be some groups that are not covered by these goals.

Another, much used, approach is to apply cost-benefit analysis to aggregate across individual preferences. Here, a monetary value is attributed to all costs and benefits, with benefits being interpreted in terms of people's willingness to pay for them. Various techniques have been developed to elicit the willingness to pay for non-market goods (Boardman, Greenberg, Vining, & Weimer, 2011). CBA is thus a tool to determine whether the project's benefits exceed the costs and can be used to rank projects unambiguously.

However, some fundamental problems limit CBA's usefulness in political decision-making (Volden, 2019). As an analytical tool, it recognizes people's preferences only in their role as consumers, and not, inter alia, their political goals and strategies (Mackie et al., 2014). It has been shown that people's preferences for what public money should be used for, may differ from their preferences as consumers (Mouter & Chorus, 2016). Furthermore, it is an aggregated success criterion that disregards how costs and benefits are distributed (Nyborg, 2014). Thus, CBA is of little help in cases in which governments have policy objectives related to the distributions of goods, which is normally the case for merit goods (as mentioned in section 2.1). A related critique is that CBA systematically downplays the welfare of future generations, not least due to the use of a discount rate (Pearce et al., 2006).

What is needed is a broad and holistic approach to evaluating strategic public project success that comprises current political goals and strategies, including goals related to equity, value for money, and possibly other valuation perspectives where relevant. Any conflicts between these perspectives ought to be made visible.

2.4. The success of projects is defined in their front-end

The front-end is the phase from when the project concept is conceived until decision-makers finally commit to the financing of a project. The phase includes problem identification, concept identification, and preparation and appraisal, but not detailed planning (Williams & Samset, 2010). The importance of the front-end phase in the development of projects has been increasingly recognized (Samset & Volden, 2016) and several studies have shown that projects that fail do so because of critical decisions taken during this phase (Williams et al., 2019). Furthermore, it is during this phase that success is defined, the choice of the project concept is made, and the framework conditions for efficient project delivery are set out. It is against these criteria for success that projects must be evaluated ex post.

The increasing recognition of projects' front-end is why several countries have introduced governance schemes for improving the success of major public projects (Volden & Samset, 2017b). Norway introduced a scheme that initially was aimed at improving cost performance in 2000 and then from 2005 to ensure the right choice of concept. The Norwegian scheme is a stage-gate model that requires that the plans, appraisals, and estimates of government agencies are scrutinized by external consultants, who are trained especially for the task before funding can be approved. The so-called Quality Assurance (QA) scheme includes two external reviews:

- QA1 – Quality assurance of choice of concept before Cabinet's decision to start a pre-project
- QA2 – Quality assurance of the management base and cost estimates before the project is submitted to Parliament for approval and funding.

Together, the two reviews constitute a broad evaluation of the project ex ante: QA1 looks for consistency with needs and strategy, and it assesses the project's value for money (i.e. to ensure tactical and strategic success), whereas QA2 checks that the project is well-planned and can be delivered within budget (i.e. to ensure operational success).

All large government projects (i.e. with an estimated cost above NOK 1000 million, some USD 110 million) within transport, defence, and building construction, as well as major information and communications technology (ICT) projects are subjected to the scheme. The formal decision to build cannot be taken before QA2 is completed. The Ministry of Finance has issued a suite of documents to guide the responsible agencies in their project development and to ensure that projects are sufficiently mature before they move from one project stage to another. The formulation of goals through a logical framework approach is mandatory, and both ensure consistency between projects and provide an institutional arrangement that promotes quality-at-entry to improve project performance.

2.5. A holistic model

Front-end loading through the arrangements described in the preceding section consolidates a baseline against which the outcome of projects can be mapped. Since 2012, a sample of projects (2–4 per year)

that have been through external QA is selected for ex post evaluation. The chosen framework for evaluation is inspired by a broad, goal-orientated model endorsed by the OECD Development Assistance Committee (DAC) (OECD, 2002). The model has been used for decades to evaluate different kinds of projects in international development, by the World Bank and other funding agencies, but has rarely been discussed in the project management literature (exceptions are Ika, 2018, 2012). As noted by Picciotto (2013), development projects are not much different from projects in developed countries, and the OECD's goal-orientated model's criteria reflect hard-won lessons of experience. The original model's five criteria – *efficiency, effectiveness, impact, relevance*, and *sustainability*,[1] together cover the operational, tactical, and strategic levels of success, and which cover intended and unintended impacts alike. We have made a few adjustments to the model to ensure alignment with the evaluation criteria used ex ante in the Norwegian QA scheme. Most importantly, we have introduced *value for money* (benefit-cost efficiency) as a sixth criterion on the strategic level, which allows for a narrow, operational definition of the original efficiency criterion (i.e. cost- and time-efficient delivery of the project). We have also used *other impacts* instead of the wider term 'impact', to make it clear that this criterion only comprises side-effects and does not overlap with effectiveness. Finally, to facilitate comparison across projects, we have introduced the use of a score between 1 and 6 to summarize the assessment per criterion.

Together, this gives four criteria on the strategic level, which should cover various valuation perspectives. Definitions of the six criteria are presented in Table 1.

We do not claim that this model is the only possible or best evaluation framework, but it covers the crucial aspects of public project success identified earlier, and it corresponds well with the QA scheme used for ex ante appraisal in Norway. A variant of the model is widely used, mainly with positive experiences, in the evaluation of development projects. A revision in 2019 concluded that, by and large, the five OECD criteria worked well and had led to standardization and consistency in the evaluation profession (OECD, 2019). It should be noted, however, that a good evaluation model is not enough to ensure successful projects. Ika et al. (2012) studied World Bank experiences in light of the high project failure rate for African projects and identified *design* and *monitoring* to be the most prominent critical success factors.

The model is generic and applicable to all kinds of projects. However, each of the criteria needs to be operationalized with more specific indicators adapted to the project in question. When collecting and analysing data, evaluators need to combine a wide range of sources and methods, and they should ensure evaluation quality and validity by performing triangulation. Unavoidably, a certain degree of subjective assessment must be accepted from the evaluator's side, especially related to the score-setting. The scores cannot be considered an absolute truth, and caution should be exercised when comparing scores across very different projects. Still, the use of scores to provide a picture of the evaluators' assessment has proven very useful following an increasing sample of evaluated projects. The main advantage of such models is that they ensure that key aspects of public project success are taken into consideration. For a more detailed presentation of the chosen evaluation framework, and some pros and cons and experiences from using it in a Norwegian context, see Volden (2018b).

Table 1
Definitions of the six evaluation criteria.

Level of success	Evaluation criterion	Definition
Operational	Efficiency	Project implementation and performance in terms of cost, time, and quality (the iron triangle)
Tactical	Effectiveness	Whether the agreed goals (typically related to user needs) have been obtained and to what extent the project has contributed to this outcome.
Strategic	Other impacts	This includes all consequences *beyond* the agreed outcome that can be attributed as the result of the project, positive and negative, short-term and long-term, for different stakeholders.
	Relevance	A project is relevant if there is a *need* for what the project delivers. Project relevance is assessed in relation to national political priorities, but also stakeholders' preferences.
	Sustainability	A project is sustainable if its benefits are likely to persist throughout its lifetime. The total impacts (financial, environmental, and social) ought to be acceptable in the long run.
	Benefit-cost efficiency	Total willingness to pay for what the project delivers in relation to cost, as measured by the Cost-Benefit Analysis.

[1] After we started applying the model, the OECD DAC revisited its evaluation criteria in 2019 (OECD, 2019). Overall the conclusion was that the criteria worked well, but there were some challenges with the way they were applied in practice, and the revision led to some clarifications and improvements in the evaluation guidelines. The OECD DAC added a sixth criterion, 'coherence' to capture in a better way linkages, systems thinking, partnership dynamics, and complexity. We considered that in the context of projects in developed countries, there was no need for this separate criterion, since 'coherence' was already considered as part of 'relevance'.

3. How successful are public projects?

3.1. Four cases

Below we discuss the success of four Norwegian projects using the evaluation model presented above. The studied projects were all part of the Ministry of Finance's QA scheme (see section 2.4), which implies that they were all large in terms of cost, and all were relatively complex. However, they differed in terms of sector, purpose, stakeholders involved, and particular features. The evaluations are conducted under the auspices of the Concept Research Programme, but each project is evaluated by a separate evaluation team. When applying the model, the team summarized their conclusions by awarding a score between 1 and 6 for each criterion, where 1 was a failure and 6 was highly successful. An overall guideline for score-settling was prepared in advance to assist the evaluation teams.

3.1.1. Oslo opera house

The Opera House in Oslo opened in 2008 and was evaluated in 2016. The investment cost was approximately €500 million. The project was heavily debated. The investment cost was considerable, especially considering the limited demand for opera nationally, and several alternative locations were considered before it was decided to locate the new opera in the largely abandoned Oslo dockside.

The evaluation concluded that the project performed acceptably both operationally and tactically. On the strategic level, the opera has proven to be both an overwhelming success and a financial burden.

The final cost was below budget. The opera opened six months ahead of schedule, and the quality was as planned.

The new opera has led to an increase in the demand for tickets, in line with the project's goal. The tickets are heavily subsidized, to make opera accessible to all. More people, from different parts of the country and different social classes, visit the opera, compared with when the old opera was located within an office block in the city centre. The opera is also being used for musicals and other types of performing arts. Therefore, the evaluation considered the project a tactical success.

The most positive result of this project was its impact on city regeneration, and perhaps surprisingly, as a tourist attraction. No formal

goals were defined for such impacts, hence they were considered positive side-effects. This remarkable building is the first thing people see when they arrive in Oslo by sea. The number of people who visit the opera to attend a performance is small compared with the number of people who come to see the building and take a walk on its' roof. The opera house has also had a positive effect on subsequent urban development. The Oslo dockside in the city's east end was largely abandoned and dominated by motorways, storage buildings and drug consumption. Today, the motorway is tunnelled, and the opera sits within one of the highest-valued commercial and residential areas of the capital. The new district has 5000 homes and some 20,000 jobs, all within walking distance of the central station – the largest hub for public transport in the country. The opera is considered the cornerstone for the subsequent development of the area. Public investment attracted private investment manyfold.

However, the evaluation questioned the opera's long-term financial sustainability and value for money. The opera relies on increasing government subsidies and receives a very large share of the national budget allocated to culture. Future pension commitments for ballet dancers and opera singers are an increasing burden. Please see Fig. 2.

3.1.2. *Double-track railway from Stavanger to Sandnes*

The 14.5 km double-track railway from Stavanger (Norway's fourth-largest city, pop. 130,000) to neighbouring Sandnes (pop. 72,000) opened in 2009, more than doubling the capacity of the then single-track railway. The project was evaluated in 2015.

The evaluation considered the operational and tactical success of the project to be poor and yet found its strategic relevance to be good.

The project experienced a large cost overrun, explained mainly by an unexpected boom in the regional construction market. On the other hand, the project was delivered within schedule, with no quality problems, and the construction costs per kilometre were lower than comparable railway projects at the time.

The tactical goals of the project were related to increased ridership, travel time savings and modal shift from the private car. The frequency of trains on the line increased as intended, partly due to new rolling stock. Passenger numbers increased, but much less than forecasts. The modal share of the train remains the same. The evaluation noted that the goal to obtain a modal shift was very ambitious and would normally require efforts beyond the building of new tracks.

In strategic terms, the project is well aligned with government strategies for a green shift in transport and thus considered relevant. Restrictions on car travel are increasingly on the agenda in Stavanger, as in other urban areas. If motorists are priced-off from the roads or left without a place to park, a high-capacity, zero-emission, fast and comfortable mode of transport may be in line with future needs. On the other hand, with passenger numbers lower than forecasted, and high construction costs, the project's value for money is considered to be poor. The strategic success of the project is therefore mixed. Please see Fig. 3.

3.1.3. *The Atlantic tunnel*

This project consisted of a sub-sea road tunnel (5800 m) and access roads (4500 m) on both sides. It connected the city of Kristiansund (pop. 24,000) to the south. It opened in 2009 and was evaluated ten years later. The evaluation considered the project's operational success to be poor, but its tactical and strategic success to be good.

The project experienced a cost overrun and the tunnel opened one year later than planned. Final costs were 12% over budget. However, compared with other subsea tunnels, the planned construction time and estimated unit construction costs (NOK per metre) were underestimated.

Tactical goals were related to time savings for motorists. The tunnel replaced an infrequent ferry service and therefore, travel time savings have been significant (around 30 min). The flexibility of travel has improved, but for HGVs with heavy loads, the tunnel gradient of 10% can be challenging.

The project is in line with strategic goals for bringing firms and labour markets closer to each other, especially in rural areas. Most of the Norwegian export values are created along the coast and various governments have worked to promote further growth through infrastructure investments. The tunnel has led to increased "commuting", and, "traffic" today is three times higher than before it opened. The net benefit to cost ratio is estimated to be above 1, which is much higher than most other Norwegian road projects. The strategic case for the tunnel thus remains high. Please see Fig. 4.

3.1.4. *The Norwegian public safety network*

The Norwegian Public Safety Network is a public safety network system based on Terrestrial Trunked Radio (TETRA). The purpose of the large ICT project was to provide a common digital mobile network adapted to the needs of the emergency services and with good coverage throughout the country. The old systems were analogue and the separate systems for the three emergency services did not communicate. The project was evaluated in 2021, six years after its opening.

The evaluation considered the project's operational performance to be mixed. There were some minor cost increases and considerable

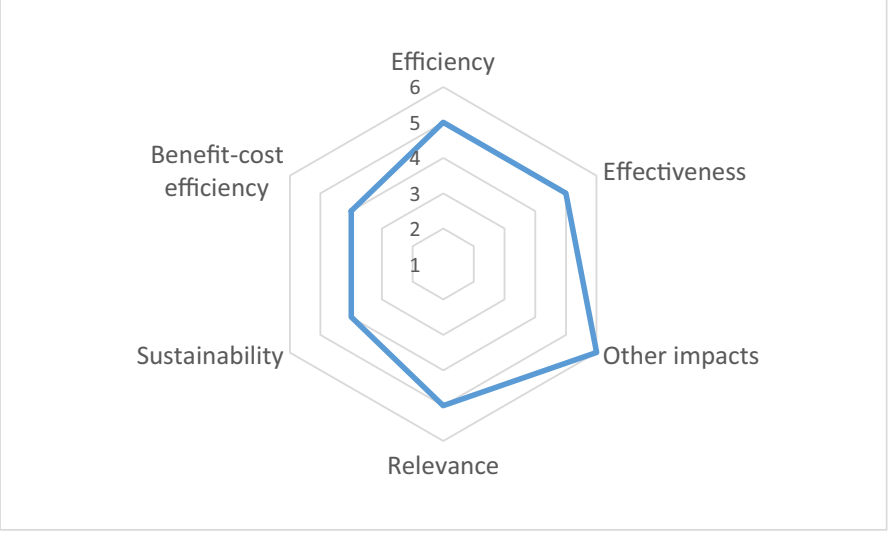

Fig. 2. Evaluation scores Oslo Opera House.

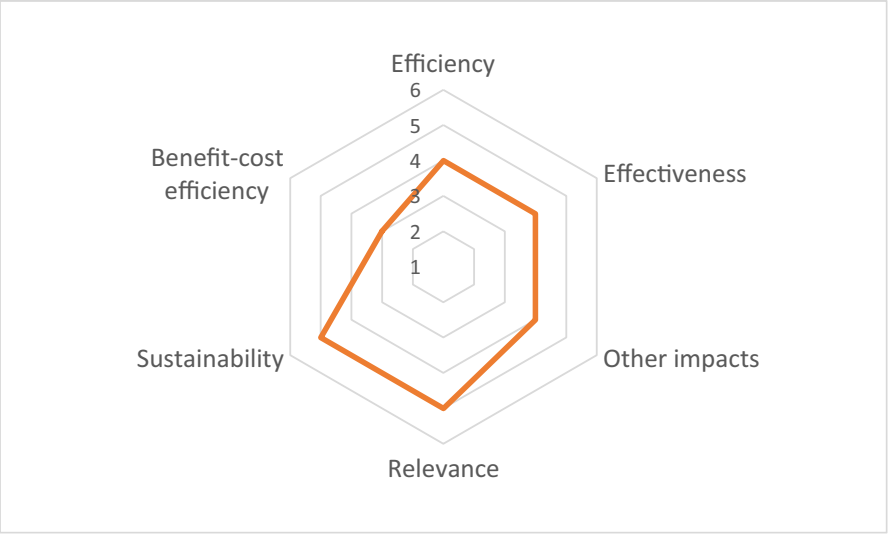

Fig. 3. Evaluation scores Stavanger-Sandnes.

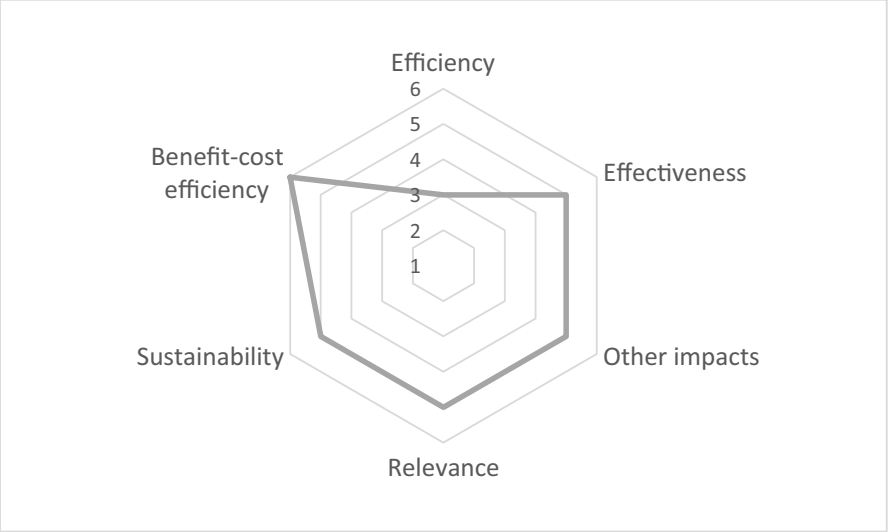

Fig. 4. Evaluation scores of the Atlantic Tunnel.

delays, but the quality was as expected. For an extremely large and complex project as this was, it was better than most people expected.

In tactical terms, the results were acceptable as well. The emergency services have all increased their efficiency and coordination. However, in the longer term, users will increasingly demand more capacity for data transfer which the current system cannot provide.

The project's strategic performance was less impressive. Admittedly, the project's relevance was good, in that it was well aligned with government strategies to improve safety levels through a better emergency network. However, the investment cost for the selected solution was extremely high compared with the benefit flows, which were postponed due to the delays. Hence, value for money was poor. The project is also awarded a low score on sustainability since its lifespan will be relatively short and the system will be outdated in a few years. The next generation of emergency networks is now being planned and will be based on commercial networks and not a separate system built for the emergency services. Please see Fig. 5.

3.2. The need for systematic evaluation and meta-evaluations

As much as there is to learn from broad single-project evaluations, there is more to learn from meta-evaluations, especially when the included evaluations are conducted according to the same evaluation framework. A meta-evaluation is defined as a synthesis of several related evaluations (HM Treasury, 2011) and is also often referred to as cross-case synthesis (Yin, 2013). The purpose is to provide an aggregation of the findings from more than one evaluation, to estimate the average or combined effect.

In the following, we present the aggregated results from the evaluation of 34 public projects in Norway, including the four mentioned above. The evaluations were conducted in the period 2012–2022. The evaluated projects represent different sectors but have in common that they are large, government-funded projects that are now completed and have been in operation for some years. They constitute about half of all quality-assured projects which have so far been completed and been some years (at least 3–5 years) into their operational phases.

Ideally, the selection of projects to be evaluated should be free from bias. The sampled projects should reflect the characteristics of the

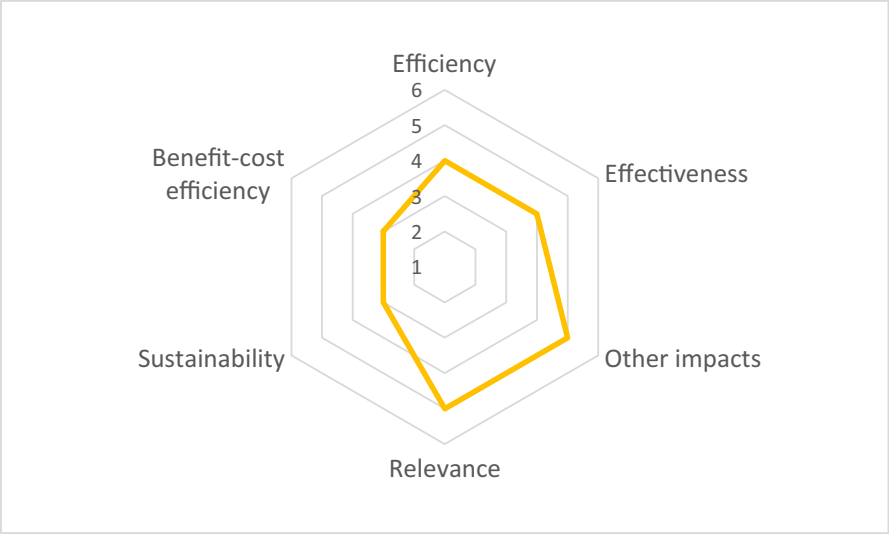

Fig. 5. Evaluation scores Norwegian Public Safety Network.

population (in this case, projects subjected to external QA, see section 2.4). This could be achieved through random sampling, but for different reasons, this may not be fully achievable. Some projects are selected because they exhibit specific features that warrant further examination through evaluation, such as particularly large cost overruns or underruns. Other projects are selected for evaluation because they represent a typical project in a particular sector. The sample of projects should include projects from all sectors subject to the ex ante QA scheme. We would argue that as the sample of evaluated projects increases, the risk of sampling bias decreases.

A main feature of the evaluations is that they all applied the same broad six-criteria framework presented in section 2.5. Another is that all evaluation teams summarized their conclusions by awarding a score between 1 and 6 for each criterion. Although reservations should be made concerning the accuracy of the scores, they are particularly useful when conducting a meta-evaluation because they allow the calculation of averages and provide an overview of how projects and sectors compare.

As shown in Table 2, the overall picture of performance and achievements is quite positive, with mean scores between 4 and 5 for all six criteria. The evaluators concluded that most of the projects were successful in more than one aspect.

The projects' operational success (efficiency) was mostly acceptable, contrary to the impression in the media and some international studies. 24 out of 34 projects (71%) were completed within budget. These encouraging results are supported by Welde and Klakegg (2022). They found that in a sample of 96 government projects in Norway the average final cost was 4% below budget and that 75% of the projects experienced cost underruns. The authors suggested that the Ministry of Finance's QA scheme had contributed to the successful cost performance.

In terms of tactical success too, the results were acceptable. All except for one project delivered the requested product or service, they put it to use and generated outcomes for various stakeholders. However, not all projects had realized the full potential in terms of benefits for users. Evaluators noted that very few projects had benefit realization plans.

Most projects were relevant and in line with strategic goals. They were also sustainable and with wider impacts that were more positive than negative. In most projects (except roads), benefit-cost efficiency received the lowest score. This indicates that although most projects achieved a positive outcome, the benefits did not necessarily outweigh the cost. In some cases, the low benefit-cost performance was known in advance, but in many others, there was no CBA in advance or it was too optimistic. Some evaluators noted that the selected project was overly excessive and that the government could have selected a simpler solution (e.g., an off-the-shelf ICT system instead of developing a complex system of their own).

The results also indicate some interesting sectoral differences. Building construction and ICT projects stand out positively on most criteria, efficiency in particular. Most of the road projects (especially those in urban areas) performed excellently in terms of effectiveness and benefit-cost efficiency, but somewhat poorer on some other strategic criteria due to environmental issues. The railway projects had an evaluation profile in the opposite direction regarding those three criteria. The defence projects in our sample performed unsatisfactorily on most criteria, but it should be noted that they only include three projects.

There appears to be some correlation between the scores for the various criteria. This is not surprising, since a well-thought-out and carefully planned project will normally be successful in several respects. However, there may also be conflicts, for example when some of the projects scored high on relevance and sustainability, and lower on benefit-cost efficiency, or vice versa. Generally, few projects scored high on *all* of the six criteria. Similarly, none of the projects were total disasters.

Table 2
Summary of evaluation results from 34 public projects in Norway.

Evaluation score	Operational success	Tactical success	Strategic success			
	Efficiency	Effectiveness	Other impacts	Relevance	Sustainability	Benefit-cost efficiency
Total (34), Mean	4,3	4,4	4,3	4,5	4,5	4,3
Building (8), Mean	4,6	4,3	4,6	4,6	4,6	4,0
Defence (3), Mean	3,3	3,7	3,7	4,0	3,0	2,7
ICT (5), Mean	4,6	4,6	4,2	4,6	5,0	4,2
Railway (6), Mean	4,2	3,8	4,5	4,5	5,0	3,7
Road (12), Mean	4,2	5,0	4,2	4,6	4,4	5,3

3.3. Why is ex post evaluation so rare?

To our knowledge, broad ex post evaluation of public investment projects is rarely conducted systematically as part of public organizations' project governance arrangements in developed countries. To the extent that some countries do require ex post evaluation, it is normally restricted to the transport sector, with the main focus on economic aspects (cost performance and benefit-cost efficiency).

In the UK, Highways England produces so-called POPE reports (post-opening project evaluation) on major highway projects 1 and 5 years after project opening. They apply four success criteria that resemble the evaluation model discussed above: cost performance, achievement of objectives, value for money, and 'other impacts' (environment). The results are encouraging, based on a sample of 60 highway schemes. Cost performance is generally good and has improved over time. Further, 90% of the projects' objectives are achieved (normally related to reductions in travel time and safety improvements). In terms of value for money, the combined benefits constitute approximately £3 for every £1 spent, which implies they fall in the high value for money category. Only the environmental impacts are negative on average, but not worse than expected (Highways England, 2019). Hence, the results are not far from those seen for road projects in the Norwegian sample.

In France, several permanent observatories have been established to follow and evaluate major transport investments over time. The main purpose is to inform the public about the status of the operation and to render accounts for the use of public funds by evaluating the economic, social, and environmental efficiency of the investments made (ITF, 2017). Here too, the general conclusion is that the economic returns are acceptable but sometimes lower than estimated. Cost overruns are not uncommon but are often offset by higher-than-expected traffic flows. Road projects perform better economically than rail projects (ITF, 2017).

However, ex post evaluation is the exception rather than the rule, even for transport projects. Worsley (2014) referred to ex post evaluations as 'the weak link' in the assessment process for transport projects in OECD countries. This contrasts the large amount of estimation and planning activities that take place ex ante. Ex post, the tendency is to investigate projects only in cases where there is reason to expect severe problems. Evaluations for learning, on the other hand, are rare.

Part of the explanation may be the belief that each project is so unique that there is little to learn across projects and sectors. This would be a misconception. The field of project management is largely generic, and seemingly different projects face surprisingly similar problems. Other explanations for not conducting project evaluations systematically are related to the scarcity of time and resources (evaluation for learning purposes is seldom urgent, but rather an activity that generates long-term benefits), and to political decision-makers' lack of interest. Politicians may view evaluation (ex ante and ex post) as part of a technocratic system that threatens their freedom of choice. The non-use of evaluation by decision-makers is a widely discussed topic within the evaluation literature (Dahler-Larsen, 2012).

4. The myth of public projects that 'always' fail

Project success is not an easy issue, not even in the case of private projects, but one that has been a central concern for project management researchers for decades. Despite the many studies, the academic and professional communities have yet to reach a consensus on the definition of success, and how it could be measured (Ika, 2009). Based on this, it should come as no surprise that the media, the most important source of information for members of the public, struggle to present an unbiased account of public project performance. Judging by anecdotal examples from projects that have experienced large overruns and delays such as those of Crossrail in England, the Scottish Parliament, Berlin Brandenburg Airport, the Big Dig in Boston and others, one could easily get the impression that public projects always fail. However, the evaluations discussed above show more nuanced results, suggesting that media coverage is rarely an unbiased source of information.

Firstly, the media often refers to available public investigations or evaluations of the project in question. Such investigations, initiated by, inter alia, the National Audit Office, are carried out when there is already cause for concern. As discussed above, evaluation of a larger sample of projects for learning purposes, which would have provided a more balanced picture of project success, is rarer.

Secondly, the dominance of negative coverage of public projects may be a psychological phenomenon. Stafford (2014) argued that people are drawn to depressing stories. Large public projects may have both winners and losers, but there is a tendency to focus on the latter group. The readers or viewers have trained journalists to focus on these things. We often chose stories with a negative tone, such as involving corruption, hypocrisy, cost overruns, and disadvantages for neighbours and the local community, rather than neutral or positive stories. This contrasts with what we say we prefer, positive news. In a world where reader interest can be measured accurately by the number of clicks a story gets, this mechanism becomes self-reinforcing and encourages more negative reporting. This is an example of negativity bias which is the notion that things of a more negative nature have a greater effect on behaviour and the mental state than neutral or positive things (Rozin & Royzman, 2001). Bad news sells and has an impact on behaviour. That is the reason why many of the successful presidential campaigns in the U.S. over the last decades have focused on negative elements of the opposing candidate, rather than positive messages.

A third reason for the prevalence of negative reporting is that bad news and disasters are more compelling and measurable than good news. And in the case of projects, immediate results, related to cost and time, are easier to observe and measure, as compared with long-term success at the strategic level. The current state of western societies, or indeed the world, is based on slow and incremental change. Neither today's welfare systems, transport systems nor education systems came about because of sudden shocks justifying big headlines. When things go according to plan, journalists will struggle to find a newsworthy angle.

The bad news bias of the popular press has a parallel term in the academic literature. In empirical research, the goal is normally to draw a sample that accurately reflects the characteristics of the wider population. However, sometimes the published results only reflect one side of the story. Publication bias occurs when the published results in an area of study are systematically unrepresentative of the population of all research in that area (Rothstein, 2008). This can lead to research to err both on the positive and the negative side and threatens the validity of the research. Flyvbjerg et al.'s (2002) study of cost overruns sparked a massive increase in the subject of cost overruns and project delivery failures. Academics from around the world framed their research questions and conceptual arguments around this study (Siemiatycki, 2018) not just on the issue of overruns, but on deliberate underestimation, while perhaps ignoring many examples of good practice. Seek and you shall find (Matthew 7:7). Love and Ahia Dagbui (2018) suggested that much of this research was based on 'fake' news and raised serious questions regarding methodologies, data, and conclusions. Media and decision-makers, they argued, are drawn to sensationalist and overly negative rhetoric which will further add to the delusion of the public regarding the success of public projects. Flyvbjerg et al. (2018) later strongly rebutted the allegations and maintained that failure concerning cost performance is the norm rather than the exception.

On the other side of the equation, Holmgren and Merkel (2017) argued that the literature on the relationship between infrastructure and economic growth has produced an unrepresentative high amount of large and positive results. Since the early 2000s, there has been increasing interest in the impact of transport projects on agglomeration and firm-level productivity (Welde & Tveter, 2021). Most studies have been ex ante and focused on *potential* additional benefits (hence the term wider economic benefits) rather than actual net impacts - positive and negative.

There is no doubt that public projects face several challenges. Many such projects are the outcome of a political tug-of-war between stakeholders in society, whose needs and priorities will concur or conflict to varying degrees, and the outcomes of such processes are not always predictable. This is clearly shown in Miller and Lessard's study of 60 international projects (Miller & Lessard, 2000). Also, the public sector has some internal challenges, such as a weakness in establishing strategic visions, a lack of professional skills and coordination amongst levels and actors, as noted by OECD (2015). Mega-projects have particular characteristics that make them difficult, such as scale, complexity, risk, and the need for social responsibility (Williams et al., 2019). But most countries have well-developed democratic processes to ensure that solutions are agreed, and governance arrangements to promote efficiency and effectiveness.

Based on the above, it is uncertain if the media or the academic literature presents an unbiased account of public projects. On the question of public project success, the jury is still out and there may be some time before a verdict can be reached, if ever. In that context, it is good news for governments and decision-makers in well-developed countries that public trust in government is high. In the latest OECD survey, around 80% of the citizens of the Netherlands, Finland, Norway, and Switzerland reported having confidence in the national government (OECD, 2021). Would results like that be possible if public projects 'always fail'?

5. Conclusions

This paper has its origins in two discussions, namely the discussion of the definition of project success, and whether public projects are successful or not. It contributes to the same two discussions by suggesting a holistic evaluation framework and by using it to demonstrate the multifaceted nature of project success. Our intention has been to bridge the gap between theory and practice and to improve the ex ante and ex post assessment of project success.

5.1. Research question 1: what is public project success?

Our findings, based on the literature and empirical data, indicate that public investments are made for different purposes (normally other than profit motives) and have multiple stakeholders that may have different views of what constitutes success. A definition of public project success needs to be broad enough to include different analytical levels of success and different valuation perspectives. One of these valuation perspectives should be value for money as measured by CBA, as public finance is a scarce resource that should be allocated to those investments that pay off the most. On the other hand, CBA as the sole success criterion would normally be too narrow, as governments may have goals and strategies beyond those that can be measured by the narrow metrics of CBA.

We have suggested a six-criteria evaluation model. It is a generic framework that has been demonstrated to work well for various types of public projects (Volden, 2018b), and a variant of the model has been used for decades for the evaluation of development projects. In this paper, we have used it to demonstrate how projects can be successful in some respects yet unsuccessful in others, which is precisely the purpose of a holistic model with various criteria. Many academic studies have discussed various aspects of project success, and departments and agencies in most countries spend a lot of resources on improving methodologies for estimating user effects and social impacts ex ante. However, actual evaluation of project success ex post remains rare. We suggest that studies of project success should be based on empirical data. Comparison of different frameworks for ex post project evaluation and the results thereof should be an avenue for future research.

5.2. Research question 2: how successful are public projects?

Our findings from evaluations of public projects in Norway indicate that in general, the results are not as disastrous as the impression often presented by the media. Their operational performance (efficiency, i.e. the iron triangle) was mostly acceptable in the studied projects. In terms of tactical and strategic success, the projects mostly delivered the planned user effects and are largely considered to be aligned with current and future government strategies. These results are encouraging but should come as no surprise, as these are the government's highest-priority projects that are largely well-planned. All of the studied projects underwent the Quality Assurance scheme before the formal decision to build. The results may also be mixed – a project can experience a cost overrun and yet be successful at a strategic level, and vice versa.

The evaluations are part of a subjective nature due to the use of scores to summarize the findings by various evaluators. Therefore, caution should be made when interpreting and comparing results across projects. We expect this weakness to become less significant on an aggregated level as the number of evaluated projects increases.

The Norwegian results cannot be generalized to all other countries. An area for future research could be to apply the evaluation framework to public projects in other developed countries, to compare results at the general level as well as by sector. There is no evidence to suggest that public projects are generally unsuccessful, especially not if they are well planned. As noted by Love et al. (2015), different empirical studies on cost performance come to different conclusions, depending, inter alia, on the point of reference from which a cost overrun is measured. The 34 Norwegian projects in this study stand out as relatively successful, which can be explained, at least to some extent, by the governance scheme they underwent ex ante to ensure that they were thoroughly planned and reviewed before being submitted to Parliament for approval and funding. However, this is not an evaluation of the scheme as such, which would have required a comparison with results from projects that had not been subject to QA. Our results, based on a sample of projects that have been organized within a standardized institutional framework for both ex ante appraisal and ex post evaluation, suggest that the results of Flyvbjerg and associates may not fully reflect the performance of all public projects. The cost performance of projects relies on front-end activities such as proper cost estimation and risk analysis, paired with external quality assurance, thus combining the inside view with the outside view in the preparation of projects before funding approval, which is in line with the recommendation of Flyvbjerg (2013).

5.3. Research question 3: why does the public think public projects always fail?

People base their impressions of public projects mainly on what the media conveys to them. Our findings indicate that the media is negatively biased and tends to report on anecdotal experiences from large, visible, and very expensive projects that fail. Furthermore, in practice, the media is mainly concerned with cost performance, which is easy to observe and measure but is only one aspect of project success. Given the lack of consensus on the definition of success in academia and the lack of systematic evaluation that can tell us whether projects were successful, this should not come as a surprise. On the other hand, many people see through the narrow media coverage and still have confidence in their government.

5.4. Ex post evaluation – a key tool for learning and improvement

We believe that a key step towards a broader and more realistic picture of public project success is to conduct a systematic and transparent evaluation of projects after they are completed and have entered their operational phases.

Ex post evaluation is still a weak link in the assessment of public projects. In Norway, ministries and agencies have become quite good at ex ante evaluation of their largest projects. As discussed in Volden and Samset (2017a), the Ministry of Finance's governance scheme for public projects has provided decision-makers with a better basis for selecting

projects and implementing them efficiently. However, to learn and improve, we need knowledge of the actual results of projects.

Knowledge about projects' operational success should be used by project managers to improve the planning and implementation of future projects. The building of a database with actual results from completed projects might contribute to more realistic cost figures. The more similar the projects from the past, the more relevant are their experiences.

Similarly, knowledge about tactical results should be used by project owners to strengthen their strategies for benefit and cost management. For example, some of the most recent ICT projects in our sample had done good work to introduce benefits management arrangements where the project was seen as an integrated part of the organization's change process. Their experiences should be useful to other projects and sectors where benefits require more than the infrastructure, such as railway investments to promote a green shift.

On the other hand, strategic success is, more than anything else, dependant on the initial choice of concept, and may be difficult to improve once the choice is made. For example, evaluations may demonstrate that railway as means of transport in areas with a small and scattered population, is poor value for money. There is little a project manager or owner can do to change this, and the lesson would be to opt for a different solution next time, such as investments in the road or non-rail public transport. However, some projects scored high on relevance and sustainability, but low on benefit-cost efficiency, and vice versa. This type of deviance needs to be communicated to the government, various stakeholders, and citizens in general, who might have conflicting views on the weighting of the criteria. The evaluations thus provide a basis for discussing how to find a better balance between different concerns.

Evaluation is not the only way to learn and improve in the area of public projects. Project managers, owners and other stakeholders should also learn in direct and informal ways, and the population can obtain information through, inter alia, public meetings and hearings (and, of course, the media). Different sources of learning should complement and reinforce each other.

Pinto et al. (2021) point to the role of context in explaining both project success and the relative importance of various success criteria. With an increasing number of evaluated projects of different types, with different stakeholders and other characteristics, there is a lot to learn about the role of context. Our findings indicate that the inclusion of more than one evaluation criterion at the strategic level is especially useful in public projects. A topic for future research could be to determine the relative importance of these criteria, as perceived by specific stakeholders and by the public. We expect this would vary across different types of public projects, as much as between public and private projects.

Funding

The work was supported by the Concept Research Programme at the Norwegian University of Science and Technology, which in turn is funded by the Norwegian Ministry of Finance.

Acknowledgments

The authors would like to thank their colleague Ole Jonny Klakegg for interesting discussions and useful comments to an earlier draft.

References

Baccarini, D. (1999). The logical framework method for defining project success. *Project Management Journal, 30*(4), 25–32.
Boardman, A., Greenberg, D., Vining, A., & Weimer, D. (2011). *Cost-benefit analysis* (4th edition). Pearson.
Breese, R., Jenner, S., Serra, C. E. M., & Thorp, J. (2015). Benefits management: Lost and found in translation. *International Journal of Project Management, 33*, 1438–1451.
Cantarelli, C. C., Flyvbjerg, B., Molin, E. J. E., & van Wee, B. (2010). Cost overruns in large-scale transport infrastructure projects: Explanations and their theoretical embeddedness. *European Journal of Transport Infrastructure Research, 10*(1), 5–18.
Dahler-Larsen, P. (2012). *The evaluation society*. Stanford Business Books.
Davis, K. (2014). Different stakeholder groups and their perceptions of project success. *International Journal of Project Management, 32*(2), 189–201.
Elkington, J. (1999). *Cannibals with forks: The triple bottom line of 21st century business*. New York: John Wiley & Sons.
Flyvbjerg, B., & Bester, D. W. (2021). The cost-benefit fallacy: Why cost-benefit analysis is broken and how to fix it. *Journal of Benefit-Cost Analysis, 12*(3), 395–419.
Flyvbjerg, B., Holm, M. S., & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? *Journal of the American Planning Association, 68*(3), 279–295.
Flyvbjerg, B., Ansar, A., Budzier, A., Buhl, S., Cantarelli, C., Garbuio, M., et al. (2018). Five things you should know about cost overrun. *Transportation Research Part A: Policy and Practice, 118*, 174–190.
Flyvbjerg, B. (2011). Over budget, over time, over and over again: Managing major projects. In P. W. G. Morris, J. Pinto, & J. Söderlund (Eds.), *The oxford handbook of project management*.
Flyvbjerg, B. (2013). Quality control and due diligence in project management: Getting decisions right by taking the outside view. *International Journal of Project Management, 31*(5), 760–774.
Flyvbjerg, B. (2014). What you should know about megaprojects and why: An overview. *Project Management Journal, 45*(2), 6–19.
Frey, T. (2016). Megaprojects set to explode to 24% of global GDP within a decade. https://futuristspeaker.com/business-trends/megaprojects-set-to-explode-to-24-of-global-gdp-within-a-decade/(retrieved April 5, 2022).
Haavaldsen, T., Lædre, O., Volden, G. H., & Lohne, J. (2014). On the concept of sustainability. *International Journal of Sustainable Engineering, 7*(1), 2–12.
Highways England. (2019). *Evaluation insight paper. Post opening project evaluation of major schemes.*
Hjelmbrekke, H., Klakegg, O. J., & Lohne, J. (2017). Governing value creation in construction project: A new model. *International Journal of Managing Projects in Business, 10*(1), 60–83.
HM Treasury. (2011). *The magenta book: Guidance for evaluation*. HM Treasury.
HM Treasury. (2013). *Public sector business cases using the five case model: Green book supplementary guidance on developing public value from spending proposals*. HM Treasury.
Holmgren, J., & Merkel, A. (2017). Much ado about nothing? A meta-analysis of the relationship between infrastructure and economic growth. *Research in Transportation Economics, 63*, 13–26.
Ika, L., Diallo, A., & Thuillier, D. (2012). Critical success factors for World Bank projects: An empirical investigation. *International Journal of Project Management, 30*(1), 105–116.
Ika, L. A. (2009). Project success as a topic in project management journals. *Project Management, Journal, 37*(4), 31–41.
Ika, L. A. (2018). Beneficial or detrimental ignorance: The straw man fallacy of Flyvbjerg's test of Hirschman's Hiding Hand. *World Development, 103*(1), 369–382.
ITF. (2017). *Ex-post assessment of transport investments and policy interventions*, ITF roundtable reports. OECD Publishing.
Kelly, C., Laird, J., Constantini, S., Richards, P., Carbajo, J., & Nellthorp, J. (2015). Ex post appraisal: What lessons can be learnt from EU Cohesion funded transport projects. *Transport Policy, 37*, 83–91.
Klakegg, O.J., & Volden, G.H. (2017))(. Governance in public projects: The Norwegian case. In R. Müller (Ed.), *Governance and governmentality for projects: Enablers, practices and consequences* (pp. 129–156). Routledge.
Laursen, M., & Svejvig, P. (2016). Taking stock of project value creation: A structured literature review with future directions for research and practice. *International Journal of Project Management, 34*(4), 736–747.
Love, P. E. D., & Ahiaga-Dagbui, D. (2018). De-bunking 'fake news' in a post truth era: The plausible untruths of cost underestimation in transport infrastructure projects. *Transportation Research Part A: Policy and Practice, 113*, 357–368.
Love, P., Sing, C.-. P., Wang, X., Irani, Z., & Thwala, D. W. (2012). Overruns in transportation infrastructure projects. *Structure and Infrastructure Engineering, 10*(2), 141–159.
Love, P. E. D., Smith, J., Simpson, I., Regan, M., & Olatunji, O. (2015). Understanding the landscape of overruns in transportation infrastructure projects. *Environment and Planning B: Planning and Design, 42*(3), 490–509.
Love, P. E. D., Ika, L. A., & Sing, M. C. P. (2019). Does the planning fallacy prevail in social infrastructure projects? Empirical evidence and competing explanations. *IEEE Transactions on Engineering Management*. https://doi.org/10.1109/TEM.2019.2944161
Mackie, P., Worsley, T., & Eliasson, J. (2014). Transport appraisal revisited. *Research in Transport Economics, 47*, 3–18.
McKinsey Global Institute. (2016). *Bridging global infrastructure gaps*, June 2016.
Meggs, T. (2018). *Delivering government major projects in a modern age*. Paris: Project Management Institute Global Executive Council. Speech delivered on 21 March 2018.
Miller, R., & Lessard, D. (2000). *The strategic management of large engineering projects: Shaping institutions, risk and governance*. MIT Press.
Morris, P.W.G. (2009). Implementing strategy through project management: The importance of managing the project front-end. In T. W. Williams, Samset, K., & K. J. Sunnevaag (Eds.), *Making essential choices with scant information front-end decision making in major projects* (pp. 39–67). Palgrave Macmillan.
Mouter, N., & Chorus, C. (2016). Value of time – a citizen perspective. *Transportation Research Part A: Policy and Practice, 91*, 317–329.

Murray-Webster, R., & Dalcher, D. (2019). (Eds.).*APM body of knowledge*. Association for Project Management.
Nyborg, K. (2014). Project evaluation with democratic decision-making: What does cost-benefit analysis really measure? *Ecological Economics, 196*, 124–131.
OECD. (2002). *Evaluation and aid effectiveness: Glossary of key terms in evaluation and results based management*. OECD Development Assistance Committee (DAC).
OECD. (2010). *Guidance on sustainability impact assessment*. OECD.
OECD. (2015). *Towards a framework for the governance of infrastructure*. OECD.
OECD. (2019). *Better criteria for better evaluation. Revised evaluation criteria definitions and principles for use.*OECD/DAC network on development evaluation. OECD.
OECD (2021). *Government at a glance 2021*. https://doi.org/10.1787/1c258f55-en.
Pearce, D., Atkinson, G., & Mourato, P. (2006). *Cost-benefit analysis and the environment: Recent developments*. OECD.
Picciotto, R. (2013). The logic of development effectiveness: Is it time for the broader evaluation community to take notice? *Evaluation, 19*(2), 155–170.
Pinto, J., Davis, K., Ika, L., Jugdev, K., & Zwikael, O. (2021). Call for papers for special issue on project success. *International Journal of Project Management, 39*, 213–215.
Rodriguez, J. (2019). 10 of the largest construction projects in the world. The Balance. https://www.thebalancesmb.com/top-ten-largest-construction-projects-844370.
Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology, 4*(1), 61–81.
Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*(4), 296–320.
Samset, K., & Volden, G. H. (2016). Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. *International Journal of Project Management, 34*, 297–313.
Samset, K. (2003). *Project evaluation*. Tapir.
Sanderson, J., & Winch, G. (2017). Public policy and projects. *International Journal of Project Management, 35*, 221–223.
Shenhar, A.J., & .& Dvir, D. (2007). *Reinventing project management: The diamond approach to successful growth and innovations.*Harvard Business School Press.
Siemiatycki, M. (2018). The making and impacts of a classic text in megaproject management: The case of cost overrun research. *International Journal of Public Management, 36*(2), 362–371.
Stafford, T. (2014). Psychology: Why bad news dominates the headlines [online]. Available from: Https://www.bbc.com/future/article/20140728-why-is-all-the-news-bad.
van Wee, B. (2007). Large infrastructure projects: A review of the quality of demand forecasts and cost estimations. *Environment and Planning B: Urban Analytics and City Science, 34*, 611–625.
Varian, H. (2010). *Intermediate microeconomics: A modern approach*, eighth edition, W.W. Norton & Company.
Volden, G. H., & Andersen, B. (2018). The hierarchy of public project governance frameworks: An empirical study of principles and practices in Norwegian ministries and agencies. *International Journal of Managing Projects in Business, 11*(1), 174–197.
Volden, G.H., & Samset, K. (2017a). Quality assurance in megaproject management, the Norwegian way. In B. Flyvbjerg. (Ed.), *The oxford handbook of megaproject management* (pp. 406–428). Oxford University Press.
Volden, G. H., & Samset, K. (2017b). Governance of major public investment projects: Principles and practices in six countries. *Project Management Journal, 48*(3), 90–108.
Volden, G. H. (2018a). Public funding, perverse incentives, and counterproductive outcomes. *International Journal of Managing Projects in Business, 12*(2), 466–486.
Volden, G. H. (2018b). Public project success as seen in a broad perspective: Lessons from a meta-evaluation of 20 infrastructure projects in Norway. *Evaluation and Program Planning, 69*, 109–117.
Volden, G. H. (2019). Assessing public projects' value for money: An empirical study of the usefulness of cost–benefit analyses in decision-making. *International Journal of Project Management, 37*, 549–564.
Welde, M., & Klakegg, O. J. (2022). Avoiding cost overrun through stochastic cost estimation and external quality assurance. *IEEE Transactions on Engineering Management*. https://doi.org/10.1109/TEM.2022.3173175
Welde, M., & Tveter, E. (2021). The wider local impacts of new roads: A case study of 10 projects. *Transport Policy, 115*, 164–180.
Williams, T., & Samset, K. (2010). Issues in front-end decision making on projects. *Project Management Journal, 41*(2), 38–49.
Williams, T., Vo, H., Samset, K., & Edkins, A. (2019). The front-end of projects: A systematic literature review and structuring. *Production Planning and Control, 30*(14), 1137–1169.
Worsley, T. (2014). *Ex post assessment of transport investments and policy interventions: Prerequisites for ex-post assessments and methodological challenges*. ITS Roundtable summary and conclusions. International Transport Foundation Discussion Paper 2014-19.
Yin, R. (2013). *Case study research: Design and methods*. SAGE.
Zachariadis, I. (2018). *Investment in infrastructure in the EU: Gaps, challenges, and opportunities*. EPRS: European Parliamentary Research Service.
Zwikael, O., & Meredith, J. (2021). Evaluating the success of projects and the performance of its leaders. *IEEE Transactions on Engineering Management, 68*(6), 1745–1757.
Zwikael, O., & Smyrk, J. (2012). A general framework for gauging the performance of initiatives to enhance organizational value. *British Journal of Management, 23*, 6–22.