LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
School of Business
Master's Degree in Strategy, Innovation and Sustainability

Heidi Korhonen

CRITICAL SUCCESS FACTORS FOR IMPROVING SOCIETAL EFFECTIVENESS OF TRANSPORT SECTOR'S R&D

Case: Finnish Transport Agency

1st Supervisor: Professor Kaisu Puumalainen
2nd Supervisor: Professor Ari Jantunen
Lappeenranta 2015

ABSTRACT

Lappeenranta University of Technology
School of Business
Strategy, Innovation and Sustainability

Heidi-Maria Korhonen
Critical success factors for improving societal effectiveness of transport sector's R&D
November 2015
78 pages, 13 figures, 3 tables and 1 appendix
1st Supervisor: Kaisu Puumalainen
2nd Supervisor: Ari Jantunen

This study investigates societal effectiveness of transport sector's Research & Development (R&D) operations. In this study effectiveness refers to organization's capability to produce the intended and desired impacts through its operations. The aim of this study is to identify the motives for evaluating societal effectiveness and recognize the critical success factors for improving effectiveness. The theoretical framework focuses first in the policy context of effectiveness evaluation in public sector and secondly the framework introduces the concept and process of effectiveness evaluation. The empirical part is carried out as a case study, which investigates societal effectiveness of Finnish Transport Agency's (FTA's) R&D. The aim is to recognize FTA's critical success factors for improving R&D operations' societal effectiveness. Based on these factors, the organization is able to define indicators for measuring effectiveness in the future operations. In this study societal effectiveness is investigated from R&D purchasers' and R&D end-users' points of views according to Purchaser-Provider-model. The results indicate that societal effectiveness evaluation is important part of R&D operations, but the implementation of the evaluation as part of daily operations is challenging. Because of limited resources, the organization is forced to strong priorization and therefore R&D tasks are secondary after the operational tasks. Based on the results the critical success factors can be recognized as resources and priorization, clear strategy and objectives, internal communications, cooperation between public and private sector and R&D implementation and dissemination.

Keywords: Societal effectiveness, Societal impact, Effectiveness evaluation, Research & Development

Avainsanat: Yhteiskunnallinen vaikuttavuus, Yhteiskunnallinen vaikutus, Vaikuttavuuden arviointi, Tutkimus & Kehittäminen

TIIVISTELMÄ

Lappeenrannan Teknillinen Yliopisto
Kauppakorkeakoulu
Strategy, Innovation and Sustainability

Heidi Korhonen
Kriittiset menestystekijät liikennesektorin T&K:n yhteiskunnallisen vaikuttavuuden parantamisessa
Marraskuu 2015
78 sivua, 13 kuviota, 3 taulukkoa, 1 liite
1. Tarkastaja: Kaisu Puumalainen
2. Tarkastaja: Ari Jantunen

Tässä tutkimuksessa tarkastellaan liikennesektorin Tutkimus ja Kehittämis- (T&K-) toimintojen yhteiskunnallista vaikuttavuutta. Vaikuttavuudella tarkoitetaan organisaation kykyä tuottaa tarkoitettuja ja toivottuja vaikutuksia. Tutkimustyön päätavoitteena on tunnistaa motiivit yhteiskunnallisen vaikuttavuuden arviointiin sekä kriittiset menestystekijät, joiden avulla yhteiskunnallista vaikuttavuutta voidaan parantaa. Tutkimuksen teoreettisessa viitekehyksessä keskitytään ensin yhteiskunnallisen vaikuttavuuden arvioinnin poliittiseen kontekstiin julkisella sektorilla, jonka jälkeen käsitellään yhteiskunnallisen vaikuttavuuden arvioinnin osa-alueita. Tutkimuksen empiirisessä osuudessa tarkastellaan tapaustutkimuksena Liikenneviraston T&K toimintojen yhteiskunnallista vaikuttavuutta ja samalla tunnistetaan organisaation kriittiset menestystekijät, joihin pohjautuen voidaan tulevaisuudessa rakentaa vaikuttavuuden mittarit. Tässä tutkimuksessa yhteiskunnallista vaikuttavuutta tarkastellaan T&K tilaajan ja T&K loppukäyttäjän näkökulmista julkisen sektorin Tilaaja-Tuottaja-mallin mukaisesti. Tutkimuksen tulokset osoittavat että yhteiskunnallisen vaikuttavuuden arviointi on merkittävä osa-alue T&K:ssa, mutta arvioinnin käyttöönotto koetaan haastavaksi. Niukkojen resurssien vuoksi organisaatiossa joudutaan priorisoimaan toimintoja, minkä vuoksi T&K tehtävät toteutetaan toissijaisena operatiivisten tehtävien jälkeen. Kriittisiksi menestystekijöiksi yhteiskunnallisen vaikuttavuuden parantamiseksi tunnistettiin riittävät resurssit ja priorisointi, selkeä strategia ja tavoitteet, sisäinen viestintä, julkisen ja yksityisen sektorin välinen yhteistyö sekä T&K tiedon ja tulosten käyttöönotto ja jakaminen.

Avainsanat: Yhteiskunnallinen vaikuttavuus, Yhteiskunnallinen vaikutus, Vaikuttavuuden arviointi, Tutkimus & Kehittäminen

Keywords: Societal effectiveness, Societal impact, Effectiveness evaluation, Research & Development, Purchaser-Provider model

ACKNOWLEDGEMENTS

I would like to thank all the people who helped and supported me during my research process. Without the guidance and support, I would not have been able to finish my research in time. My special thanks goes to following people:

First of all I would like to express my very great appreciation to my first supervisor, LUT-professor *Kaisu Puumalainen* for the support and guidance during my research process. The time that Kaisu dedicated for helping me to proceed in my research was been priceless.

Secondly I would like to thank *Tiina Jauhiainen* and *Sinikka Hartonen* from the Finnish Transport Agency for the cooperation and support during my whole research process. I would also like to thank you for the opportunity to familiarize myself with the agency's R&D in the practical level.

Thirdly I would like to thank all the people who dedicated their time for participating to my research. I truly appreciate that these people were honest and brave to express their personal opinions. Without these people, this research would have not been possible.

And finally I would like to thank my family and friends for encouraging and giving me positive thoughts especially in times of frusturation.

Gratefully yours,

Heidi Korhonen

November 2015

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

BIM	Building Information Model
EU	European Union
FTA	Finnish Transport Agency
ITS	Intelligent Transport Systems
MTC	Ministry of Transport and Communications
PM	Performance Management
PPP	Public-Private Partnership
R&D	Research & Development
RUSE	Research Unit for the Sociology of Education
Trafi	Finnish Transport Safety Agency
VTT	Technical Research Centre of Finland

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background

Millions of euros are annually invested in research for developing the transport system. The question is, how to ensure that these investments are effective and produce the impacts that was initially aimed for? Although measuring the impact of research and its relevancy in society is challenging to undertake, it is issue that the public sector is keen to embrace (Lyall *et al.* 2003).

While the public sector's resources are decreasing, but requirements for better efficiency and productivity are crowing, *effectiveness* – *i.e* capability to produce intended impacts – has become the keyword. Effectiveness evaluation not only supports the development of the organization, but it also provides evidence for policy makers and civil society on public organization's performance and furthermore, whether the organization has achieved the desired outcomes. The pressure for engaging to societal effectiveness evaluation can be seen especially on public sector, which has been criticized for lacking transparency, accountability and alignment to respond customers' needs (Steffens & Matthews 2004).

Public R&D managers have commonly focused on measuring inputs and immediate outputs rather than assessing whether the intended goals have been achieved for improving well being (Gertler *et al.* 2011). In other words the evaluation focus was mainly on organization's performance measures and performance management. The growing global trend of evidence-based policy shifted the focus from the inputs to evaluating outcomes and results by utilizing the data obtained from performance evaluation. However, there is an interesting relation between the performance and actual impacts. Although organization's performance is good, it does not always correlate as high impacts. This is why both types of evaluation are important for an organization to improve its transparency and accountability.

In evidence-based policy monitoring and evaluation are key issues when improving quality, efficiency and effectiveness of the operations. The obtained evaluation data not only informs about the program's performance but also promotes a strong foundation for transparency and accountability (Gertler *et al.* 2011). According to Paasio (2003) the true challenge is implementing and engaging effectiveness evaluation into organization's everyday operations. First each actor in the organization needs to understand what evaluation and effectiveness concretely means in one's own professional practice. Secondly the whole work community needs to be able to implement those professional practises in which the evaluation is part of everyday operations, which are effective and whose effectiveness is recognized (Paasio 2003). At the end, evaluation is more than just measuring; it is a wide paradigm and organization culture that is not possible to create without reliable and systematic data production (Paasio 2003).

1.2 Research gap

Although the performance of policy interventions has been evaluated for ages, these evaluations have often failed to tackle the societal impacts and cognitive interactions (Lähteenmäki-Smith *et al.* 2006). The focus in the public sector's evaluations has been mainly in program inputs and processes until the 1980s reforms, which led to a change in emphasising program outputs and outcomes. These reforms were directed at improving the responsiveness of the public sector to the needs of its stakeholders (Guthrie & English 1997). At the end of 1990's the scope of research expanded and societal impact of research was highlighted as a crucial part in impact assessment.

Societal impacts have been studied by several authors (van der Meulen & Rip 2000, Lähteenmäki-Smith, Hyytinen, Kutinlahti & Konttinen 2006, Bozeman & Sarewitz 2011) and these studies concentrated in assessing social, cultural, environmental and economic impacts of publicly funded research. The results of these studies indicate that although many organizations are willing to assess

societal impacts, it is not clear how societal quality of research can be evaluated. Paasio (2003) investigated societal effectiveness on public health care sector and the main finding of this study is that societal mission cannot be implemented in the public sector without having effectiveness evaluation as part of professional practice. Societal effectiveness of R&D operations has been investigated very little. The Ministry of Environment evaluated societal effectiveness of its R&D operations in 2013 and the results basically indicate that the role of the Ministry in R&D operations was weakly known. The results emphasized the importance of communications, priorization, open data and monitoring as the key issues in societal effectiveness. Technical Research Centre of Finland (VTT) made a research on societal effectiveness of R&D operations of Finnish polytechnics' in 2006. The results emphasized that R&D should be part of organization's strategy and that widening networks and prioritaziation of resources are the key issues for creating effectiveness. These results are similar as the results of a research on universities societal effectiveness made by Research Unit for the Sociology of Education (RUSE) in 2015.

The investments on developing and maintaining the transport sector are enourmous and these investments also include the citizens' (i.e tax payers') money. To prove that this money is utilized effectively and the operations are creating the intended societal impacts, the transport sector needs to be able to show it. Effectiveness evaluation as a strategic tool can improve organization's transparency and accountability towards the citizens and stakeholders. Lähteenmäki-Smith *et al.* (2006) categorized the motives for evaluating societal effectiveness as internal and external motives. The most crucial internal motives of evaluating societal effectiveness are connected to improved management and organizational learning. However, the first thing is to ensure that effectiveness and its evaluation methods are understood in the same way inside the organization. Therefore this study focuses in investigating how societal effectiveness and societal impacts are defined in the transport sector and how societal effectiveness can be improved in the future.

1.3 Objectives and research questions

This case study investigates societal effectiveness of transport sector's R&D. The aim is to identify the motives for evaluating societal effectiveness and the critical success factors for improving effectiveness in the future. The case organization, Finnish Transport Agency (FTA), is a public agency that is responsible for purchasing such of R&D that has positive impacts on the society. R&D purchases in FTA are made by following *Purchaser-Provider-model*, in which FTA is the purchaser of R&D and provider is an external party. In this study societal effectiveness is investigated from R&D purchasers' and R&D end-users' points of views and therefore the R&D provider aspect is left out of the research scope. The purchasers of this study are FTA's R&D experts who are the key people in R&D operations. Each of these experts present different R&D field which are *Pavements, Environment, Building Information Model* and *Intelligent Transport Systems*. The R&D end-users of this study were chosen from different organizations and companies, which are involved in FTA's R&D operations.

This research is seeking answer for the following main research question:

> *RQ1: What are the critical success factors for improving societal effectiveness in transport sector's R&D?*

Based on the main research question, 4 sub questions were made:

> *SQ1: What are societal impacts in the case of Finnish Transport Agency?*
> *SQ2: What are the motives to evaluate societal effectiveness?*
> *SQ3: What are stages of effectiveness evaluation?*
> *SQ4: What are the key challenges in societal effectiveness evaluation?*

In this study *impact* indicates the ultimate, long-term outcomes. Impacts are always generated by a specific intervention and they can be anticipated or

unanticipated, positive or negative, intentional or harmful (Dahler-Larsen 2005, 7).

In this study *effectiveness* indicates the capability of producing the desired impacts. When *efficiency* stands for “doing things rights”, *effectiveness* stands for “doing right things”. Paasio (2003) has recognized 4 levels of effectiveness, which are case-by-case effectiveness, service effectiveness, societal effectiveness and exploitation of effectiveness knowledge. Societal effectiveness refers to organization’s capability to produce societal impacts.

In this study *effectiveness evaluation* indicates evaluating to what extent the desired impacts have been achieved. The indicators for measuring effectiveness can be based on organization’s critical success factors.

Critical success factors are the key variables or conditions that have a impact on how successfully organization meets its mission and objectives (Rouse 2014). In this study critical success factors are those variables or conditions in which organization must be excellent and which should be evaluated for achieving societal effectiveness.

1.4 Theoretical framework and structure

The theoretical framework of the research is illustrated in Figure 1. The first part of the theoretical framework deals with policy rationales and transport sector’s policy. The aim of this part is to describe the context of societal effectiveness evaluation and clarifying the motives for societal impact evaluation. And finally the theory of evaluation introduces the main concepts of effectiveness evaluation process before continuing to the empirical part of the study.

Figure 1. Framework of research

This paper has 8 main chapters. The theoretical framework is introduced in

chapters 2 and 3 in the order illustrated in Figure 1. Chapter 2 describes the policy context of societal impact assessment in relation to core rationales of innovation policy. The focus of chapter 3 is in effectiveness evaluation and its key concepts. In chapter 4 the methods and data of the research are introduced. Chapter 5 introduces briefly the case organization and its R&D operations. In chapter 6 the research results are introduced and analyzed. In chapter 7 the results are further interpreted by providing answers to the research questions. Final conclusions and further research topics are presented in chapter 8.

2. POLICY CONTEXT OF EVALUATION

2.1 Role of evidence-based policy

Millions of euros are annually invested in research for developing the transport system. The question is, how to ensure that these investments produce the impacts that was initially aimed for? As Chiesa and Masella (1996) noted the belief of "*the higher the R&D expenses are, the more effective the outputs are*" has proved to be false in many cases. In the transport sector the investments decisions have been traditionally made from economic aspect and the focus was especially on cost-benefit issues. However, today a broader range of evidence is required as result of growing interest towards the social implications of transport sector decisions (Steinbach 2013). While seeking to justify innovation policy interventions, the policy-makers often tend to evoke notions of market and system failures. The existence of imperfect markets has led to traditional market failure legitimacy in relation to national policy-making and public R&D funding (Lähteenmäki-Smith *et al.* 2006). The role of public R&D is to compensate the wider market failures.

Lähteenmäki-Smith *et al.* (2006) proposed an adoption of evidence-based decision-making and horizontal-driven innovation approaches as the basis for rationalizing societal impact analysis. Both of these approaches are relevant when reforming government organizations and improving the exploitation of knowledge produced in public research organizations. However, Lähteenmäki-Smith *et al.* (2006) argued that a paradigm shift might occur, which would place societal impact of R&D in more systematic context. In this context the actors would be allowed to place their own strategic choices by utilizing impact assessment as tool in strategic policy-making.

The growing global trend of evidence-based policy shifted the focus from analyzing the inputs to evaluating outcomes and results. Evidence-based policy indicates political decision making, in which the decisions are rationalized by

objective evidence. In evidence-based policy monitoring and evaluation are key issues when improving quality, efficiency and effectiveness of the operations. The obtained evaluation data not only informs about the program's performance but also promotes a strong foundation for transparency and accountability (Gertler *et al.* 2011). Evidence-based policy emphasizes especially the cooperation and interaction between the policy makers and scientific community (Hyytinen & Toivanen 2010).

Some academics have argued that in the transport sector the use of other forms of "evidence" has been limited as due of governmental principles for commissioning research (Terry 2000). According to Sanderson (2002) two forms of evidence is required in order to improve governmental effectiveness. The first form is such of evidence that promotes accountability in terms of results. This type of evidence is to show that the government is working effectively and it is primarily in the form of information on attributes of performance. The second form is evidence in promoting improvement through more effective policies and programmes. This evidence shows how well the policies and programs work in different circumstances, i.e how the policy interventions achieve change in social systems. Conventionally it can be assumed, that reliable knowledge provides a basis for effective action; it is explanatory, theoretical and provides an understanding of how the policies work (Sanderson 2002). Nutley & Davies & Walter (2002) recognized 4 key issues for improving evidence use in policy and practice, which are:

1. Nature of evidence
2. A strategic approach to knowledge creation
3. Effective dissemination and wide access
4. Increasing the uptake of evidence

The nature of evidence stands for the agreement as to what counts as evidence in different circumstances, where research is one source of evidence. Strategic approach in creation of evidence gathers evidence in the form of robust knowledge. The third requirement for improving evidence use is to enable effective

dissemination and wide access to the knowledge. However, effective dissemination has its own limits. Only providing knowledge is not enough, the knowledge need to be also pulled from the potential end-users (Nutley & Davies & Walter 2002). And final requirement is that the implementation and utilization of the evidence in practice needs to be ensured and encouraged. Uptake needs to be defined widely, because there are many ways to utilize the evidence appropriately (Nutley & Davies & Walter 2002). When preparing and implementing evidence-based policy, the exploitation of the existing data and ex-post evaluation of the policy are the key issues (Hyytinen & Toivanen 2010). Research should be therefore commissioned to ensure that the impacts of the policy are understood in the best possible way.

2.2 Transport sector's innovation policy

R&D can be seen as the fundamental input into the innovation process and through innovations the organizations are able to increase productivity and competitiveness. Besides the private companies, this causal relationship applies also to public R&D organizations, which constantly seek to improve their performance and create positive impacts on private sector (Bozeman & Melkers 1993). Although investments on R&D can have significant impacts on economic growth and competitiveness, the key challenge is to prove the exact relationship between the investment and economy. The outcomes of the R&D programs are often described, but scientific measurement for investigating whether the investment was efficient is often lacking.

Lähteenmäki-Smith *et al.* (2006) emphasized in their study that there is a need to identify the ways, in which societal impacts could be developed in relation to improving R&D environments. The domestic governance of innovation development has maintained its range of instruments from vertical to horizontal levels despite the wide internationalisation of R&D. While science and technology policy is realized at sectoral level, innovation policy is assumed to be horizontal and realized at cross-sectoral level (Lähteenmäki-Smith *et al.* 2006). The

boundaries between these two policies are conceptually and functionally shifting.

The main phases of innovation policy-making are illustrated in Figure 2. Rationale and strategy revision focuses in why public policy is needed and what policies are implemented. In the first phase the rationales of public intervention for innovation are defined. In the second phase the policy strategy i.e the action plan is formulated for the implementation of the policy, which is the third phase. At this phase the efficiency and effectiveness are analyzed for investigating how well the policy was implemented. The final phase of innovation policy-making is the assessment of socio-economic impacts of the policies and policy measures. At this phase impact assessment provides feedback in form of new information for developing the phases of policy-making.

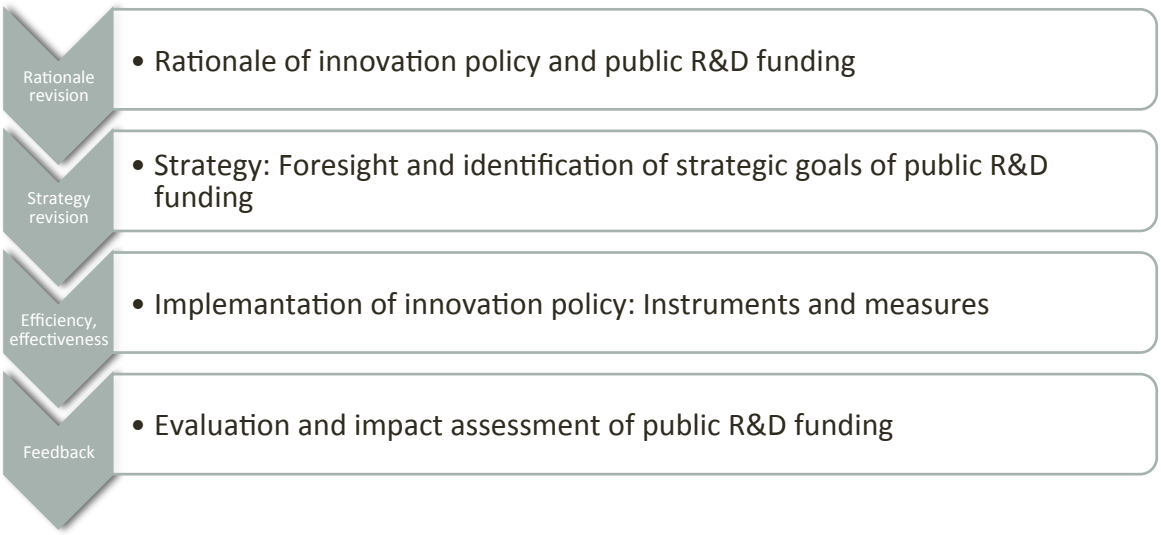

Figure 2. Innovation policy-making framework (Adopting Loikkanen & Kutinlahti 2005).

The development towards service society has changed the operational structure and administration in the public sector. This change reflects in the transport sector as transport's development towards a comprehensive service, where the needs of customers and functionality are the key issues. In Finland, Ministry of Transport and Communications (MTC) is responsible for the transport policy and transport system. According to the Ministry (2015), the new Finnish transport policy aims to improve productivity and effectiveness and promote sustainable growth,

competitiveness and well-being. The role of MTC and the administration is to act as a facilitator in creating new business opportunities and enable these businesses to improve competitiveness.

2.3 Purchaser-Provider approach in R&D

The attention for public administration's financing opportunities and targets has been increasing during the last decades. At the same time the public hierarchical production system has been criticised for being too heavy and inefficient. This criticism led to change in the role of the public sector as provider of services towards public sector as the enabler and organizer of services (Valkama 2004). This popular public sector reform is referred as the *Purchaser-Provider* approach, which is one version of Public-Private-Partnership (PPP). According to PPP private sector participates to public sector's projects as the co-financer.

Purchaser-provider model was adopted in the public sector by the end of 1980's. At the same time a new practical format of quasi-market was also established. Quasi markets differ from other markets especially from the demand point of view. In quasi markets the demand is not defined straight from the consumers' but it is directed through boundaries that the public investor has set *(*Figueras *et al*. 2005*)*. The purpose of the quasi-market is to increase the political decision defining in defining the services to be ordered and to open the service delivery for the public sector's internal and external markets (Huuhtanen *et al.* 2009). After the mid 2000's the researchers started to emphasize Purchaser-Provider as an *approach* rather than model because of its several variations (Huuhtanen *et al.* 2009).

On theoretical level, Purchaser-Provider model separates the purchaser and provider functions from each other to achieve tighter targeting of clients and outcomes, improved accountability and transparency, and improved efficiency and effectiveness in the delivery (Steffens & Matthews 2004). Huuhtanen *et al.* (2009) added that when these two actors are separated from each other, their ability to focus in developing their own core missions would also improve. Purchaser-Provider-model can be further defined as the relationship between any public

purchaser and public or private provider, in which these roles are differentiated based on a subscription agreement. The main objective of the model is to change the traditional hierarchical steering into more contract-based approach. In FTA purchaser-provider approach has led in adoption of new professional practises and wider scale of operation development as a result of wider networks.

There are 5 main actors in the Purchaser-Provider model (Figure 3). Tdhese are 1) Principals, who define the objectives and allocates the resources, 2) Purchasers, who prepare the orders and contracts and monitors the process, 3) Providers, who provide the service based on the contract, 4) End-users of the services and 5) The operation regulators (Rantala 2008).

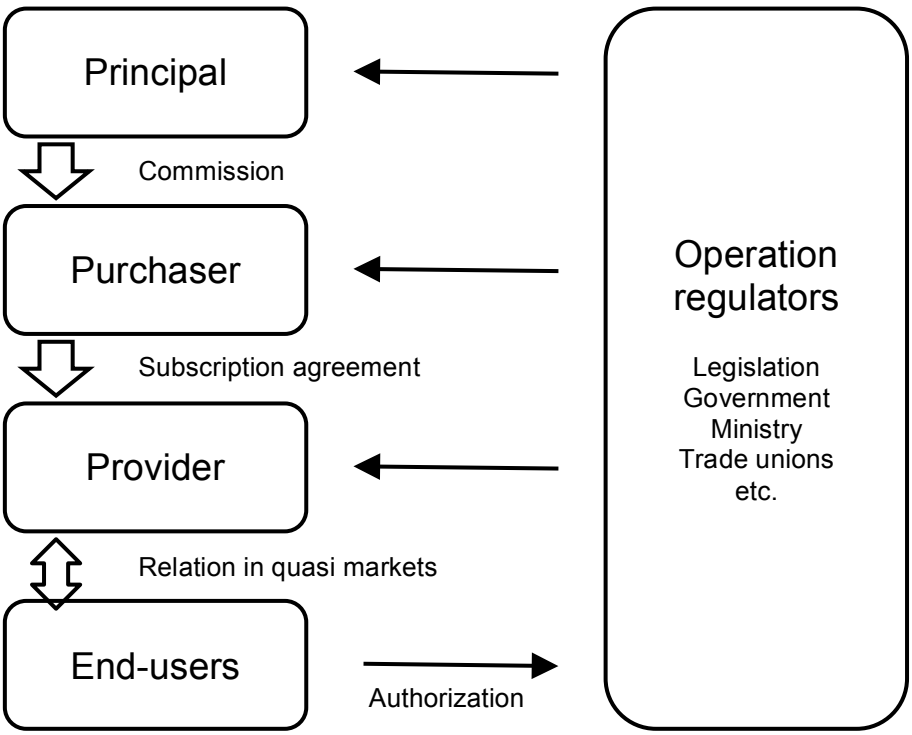

Figure 3. Main actors and elements of the Purchaser-Provider model (Rantala 2008).

From the purchaser point of view the approach requires strong expertise on recognizing the needs of services and also ability to prioritize these needs. At the same time the providers need to be able to control their business approach for remaining competitive in the markets (Kuopila *et al.* 2007). The role of end-users is more limited in government-funded services than in privately funded. In

government-funded services the end-users are not retained as customers, because the relationship is mostly based in exchange (Rantala 2008). In purchaser-provider approach the end-users have two roles: they are the users of the services and the ones who vote for organizing the services as taxpayers. In FTA's case, FTA is the purchaser of R&D and provider is usually an external consultant, university or research institute. The range of FTA's R&D end-users is wide as the end-user can be for example road user, contractor or company.

Based on market theory, the inefficiency of public service production is due to the reason that the providers are in monopoly-position and therefore lack the incentives to improve the production efficiency (Savas 2000). According to Humphries & Wilding 2004 monopolistic (only one provider) and monopsonistic (only one purchaser) markets are the most significant barriers when aiming at added value for public money. Rantala (2008) noted when the markets operate transparently; the outcomes of purchaser-provider approach are positive, the providers have adequate decision-making power, the end-users have sufficient discretion and the contracts' uncertainty is controlled.

2.4 Motives for evaluating societal impacts of R&D

When it comes to making investment decisions on research, *effectiveness* is the keyword (Bornmann 2013). The concept is often mixed with term of *efficiency*, which compares the obtained outcomes to what could have been achieved with the same amount of resources. In turn, effectiveness is the degree to what extent the organization achieved the desired impacts. Hyytinen & Konttinen (2006) noted that by undertaking effectiveness evaluations, organization can strenghten its transparency from wider societal aspect and support the organizational development when transferring the evaluation data through all organizational levels. The evaluation data can be therefore considered as tool for internal learning in understanding which operations need more effort and which should be eliminated.

Then what drives organizations to evaluate societal effectiveness? Although the performance of policy interventions has been evaluated for ages, these approaches have often failed to tackle the societal impacts and cognitive interactions (Lähteenmäki-Smith *et al.* 2006). According to Hyytinen & Konttinen (2006) the motives for effectiveness evaluation are connected to pressure of change. The change can occur in administrative culture, it can be need for reconciliation of political objectives or need for organizational development. When the public sector adopted performance management and performance accountability, it directed strongly the change in administrative culture. According to Lähteenmäki-Smith *et al.* (2006) the motives can be classified as external or internal motives. External motives are based on factors such as result-driving planning or general shift to more service-oriented culture, while internal motives stand for the need to develop activities, programs and organizations.

The most important external motives are the demand of results and increased effectiveness, accountability and transparency. In addition to reasons of accountability and transparency, evaluations are often made to justify those policies and actions that have already taken. (Lähteenmäki-Smith *et al.* 2006). Therefore evaluation data is needed to prove that policy measures have been correct or to indicate that policies need to be changed. But the external motives are not only derived from the requirement for proving effectiveness and accountability. R&D organizations are under inspection due to other trends as well, such as the emergence of service-oriented culture (Lähteenmäki-Smith *et al.* 2006).

The internal motives are related to the needs of developing R&D operations and organizations. The most central internal motivation is connected to the need to develop the organization, its expertise-base and its learning capacity in line with the new challenges in the working environment. This model is build on a continuous cycle, with self-evaluation and monitoring, thus contributing to the systematic identification of needs within the organization and its environment (Lähteenmäki-Smith *et al.* 2006). Evaluation and impact assessment are

instruments for developing practices of accountable management and good governance, and for creating societal benefits for a broader set of stakeholders.

3. EFFECTIVENESS EVALUATION

3.1 Concept of evaluation

As traditionally in earlier program evaluations, the R&D measures have also concentrated in program's inputs rather than outputs. However, during the last few years the public sector's focus has changed towards measuring the results and effectiveness, which is partly as result of the criticism that public organizations lack transparency and alignment to respond the society's needs (Steffens & Matthews 2004). The evaluation of government R&D program is proved to be more challenging than private R&D programs, because governmental programs have externalities and multiple objectives. The objectives of governmental programs are often stated in non-financial terms and the conventional financial reporting mechanisms do not always capture the performance measurement (Guthrie & English 1997). Therefore other types of reporting methods are required for evaluating whether the programs have achieved what was intended. The government R&D programs are also highly constrained when compared to private R&D programs as the abilities of the R&D managers are limited due of several governmental constraints. The R&D manager needs to allocate the funds as effectively as possible, while taking the priorities of bureaucratic superiors, federal budget controllers, political institutions and all the researcher stakeholder groups into account. It is no wonder why the word 'strategic' is not often linked to government R&D management (Bozeman & Rogers 2001). Another significant limitation is the diversity among the objectives of the different government agencies. According to Bozeman & Rogers (2001) the efforts to develop hierarchical rationalization related agencies' objectives have shown the discontinuities among the various bureaucratic units of government.

Any type of evaluation is used for answering a specific question related to design, implementation or results. In general matter, evaluations are about measuring the achievement of the objectives, which need to be realistic, clear and acceptable by the personnel and customer (Hyytinen & Konttinen 2006). In contrast to continuous monitoring, evaluations are implemented at certain moments or project

or program cycle. Rossi & Freeman & Lipsey (2004) defined evaluation as:

Utilizing the research methods of social sciences,
For systematic clarification of effectiveness,
So that the new information is utilized in political and organizational environment,
In order to improve social conditions.

The design, method and cost of the evaluation depend on the type of question that the evaluation is trying to answer (Gertler *et al.* 2011). According to Paasio (2003) evaluations have 5 basic questions, which are also the main parts of any evaluation. For an intervention to have value, these following questions must have reliable answers:

1. What is the problem's or phenomenon's character that is the target of the intervention?
2. What is the theory or logic of the intervention's operation?
3. How the intervention has been implemented?
4. What are the intervention's impacts?
5. What is the cost-effectiveness of the intervention?

The first step in effectiveness evaluation is to recognize and define the problem or challenge. In transport sector the main challenge can be for example how to improve the functionality of the transport system or how to decrease the time spend in the traffic. Regardless of the organization, the needs of the stakeholders are the key starting point. This first step of effectiveness evaluation can be defined as *need assessment,* which is utilized to describe the societal problems and people's needs. Need assessment needs to be made in every effectiveness evaluation for the effectiveness to occur in the organization's operations (Paasio 2003).

Introducement and evaluation of the intervention's theory is one key part of effectiveness evaluation. The theory of the intervention's operation is regarded as

those assumptions, which exist between the operations and the goals of achieving welfare (Paasio 2003). Basically these assumptions answer to question *"Why these operations reduce societal problems or increase welfare?".* According to Paasio (2003) one of the most common theories of intervention is that the positive development of customer's life situation is based on consumption of services and by producing services this positive development can occur.

The third stage of effectiveness evaluation is evaluating the implementation of the operations. According to Rossi, Freeman & Lipsey (1999) 3 questions should be considered in evaluating the implementation: 1) Does the service reach its target group?, 2) Does the service operations and supportive functions meet the criterias set in the strategy? and 3) Is positive change occurring in customers' welfare?. Evaluation of implementation is crucial because it provides information about how the future operations could be developed and improved. At this stage organization is also able to recognize those functions that are not effective and should be eliminated from the operations.

All the first 3 stages are aiming for enabling the fourth stage of the effectiveness evaluation, which is impact assessment. In literature impact assessment is often confused with effectiveness evaluation. Impact assessment concentrates in monitoring what kind of impacts the organization produces, while effectiveness evaluation monitors to what extent the desired impacts have been produced. The key challenge in impact assessment is to identify the causal relationship between the project or program and the outcomes of interest (Gertler *et al.* 2011).

Impact assessment can be divided into two categories: ex-ante evaluations, which are made before a program or project is launched and ex-post evaluations, which concentrate on goal achievement and effectiveness after project or program has ended. In theory these evaluations are also referred as *Prospective* and *Retrospective* evaluation. According to Gertler *et al.* (2011) prospective impact evaluation produces strong and credible evaluation results for three reasons. First, the baseline data can be collected in order to establish pre-program measures of

outcomes of interest. Baseline data can also be used to evaluate effectiveness, i.e whether the program reached its intended impacts. Secondly, when defining the measures of program's success in the planning stage, the focus remains on the evaluation and intended results. The impact evaluation design helps in clarifying program objectives, because it requires establishing well-defined measures. And finally, in prospective evaluation the comparison groups are identified before the program implementation. In retrospective evaluations it is more challenging to analyse whether the program was implemented successfully, because of limited information. One reason for limited data is that many programs do not collect baseline data unless the evaluation was built in from the very beginning (Gertler *et al.* 2011). Retrospective evaluations are necessary when assessing programs that were set in the past. The feasibility of retrospective evaluation depends on the context and is therefore never guaranteed (Gertler *et al.* 2011).

In the final stage the evaluation data is systematically utilized for improving the organization's operations and services. Effectiveness evaluation supports at its best the development processes of the organization, in which the learned experiences and new knowledge is transferred crosscutting to whole organization's operations. These evaluations have an important role in instilling the organization's strategic choices and motivating the personnel for the development. At the end, evaluation is culture that is based on scientific professional practices and professional ethics, and to learn this culture might take several generations (Paasio 2003).

Organization's management has the main responsibility in implementing evaluations. The true question is whether the organization wants to improve its effectiveness. From this perspective the operational objectives, plans and overall performance are crucial. When the systematic information between the performance and impact is lacking, it normally indicates problems in improving the productivity (Paasio 2003).

3.2 Logic model of impact

Performance management (e.g *efficiency*) and impact assessment (e.g *effectiveness*) form the knowledge base for impact. According to Kahn and McGourty (2009) effective performance management focuses in achieving optimal value from the resources that are allocated to achieving its objectives. Performance management therefore enables organization to assess its processes towards the objectives and helps in decision-making on future initiatives with the goal of improving the organizational performance (Amaratunga & Baldry 2002).

Evaluation process starts by defining the evaluation question in other words why this evaluation is to be made. Impact evaluation is a type of evaluation that seeks to answer to cause-and-effect questions (Gertler *et al.* 2011). According to Getler *et al.* a basic impact evaluation question can be for example ”*What is the impact or causal effect of the program on an outcome of interest?”.* After the evaluation question is defined, the second step is the development of a chain of result (Figure 4). This logic model describes the causal logic of how and why a specific program or project will achieve the intended outcomes. Logic model should be developed at the beginning of the process together with stakeholders for achieving a common vision of the program and its goals (Gertler *et al.* 2011).

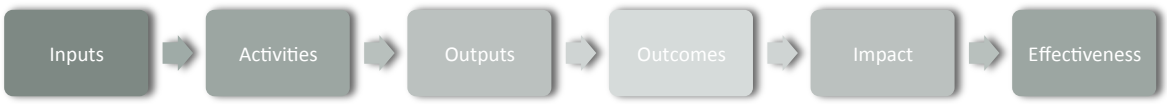

Figure 4. Logic model of a result chain (Adopting Weiss 1995)

The idea of the logic model is simple, but the challenges appeal when moving from impact to effectiveness, as the difference between impact and effectiveness is often difficult to recognize. Therefore it is crucial that the evaluation is carried out through all the stages of the logic model for the effectiveness to exist. If the effectiveness is unknown, the evaluation of the earlier stages is impossible and therefore it is difficult to analyse whether the work done for the program or project

was useful in general matter. Result chains are useful because they allow policy makers and managers to make program goals explicit, thus helping them to understand the causal logic and sequence of events behind a program (Gertler *et al.* 2011). Organizational performance metrics are necessary in everyday operations, but in measuring the progress towards impact depends on the relationship of the outputs and outcomes.

3.2.1 Inputs

Inputs are the resources dedicated to the program or project, such as the employees, budget and time. The focus should be on those resources, which are crucial for achieving the strategic goals. The most important resources are personnel, material and economic resources. According to Kettner & Martin (1998) central inputs are customers' and citizens' needs and problems, which is the key starting point in the process towards effectiveness. This also indicates that the resources should be in relation with the customers' and citizens' needs (Paasio 2003). In the case of FTA, the citizens' needs are the key starting point of the operations. FTA utilizes its expertise in regocnizing the needs of the society, and to respond to the demands, sufficiently personnel, money and time are required.

3.2.2 Activities

Activities are the operations that are needed to implement a program or project. In other words activities convert inputs to outputs. Organizations can achieve its objectives and fulfill its mission through activities, which can be processes, techniques, tools, technology and actions. Some common activities are product development, engagement in policy advocacy, building infrastructure and providing services. In the case of FTA, the main activities can be divided into two categories: operational tasks and R&D tasks. The operational tasks concentrate on maintaining the Finnish transport system, while R&D tasks concentrate on developing new methods for ensuring that the transport system remains effective and safe.

3.2.3 Outputs

Outputs are the immediate results of the activities that affect directly the stakeholders. They are the measurable, tangible and direct products or results of activities. Outputs will lead to desired outcomes, but they are not changes by themselves (Kellogg Foundation 2004). In the study of Walter *et al.* (2007) the degree of engagement with the outputs signifies the level of stakeholders' involvement in the process. This engagement should appeal also in the public R&D operations and stakeholders need to be able to involve in the R&D operations from the beginning of project. In the case of FTA, an example output is publication or report of R&D.

3.2.4 Outcomes

The terms of outcomes and impacts are often challenging to recognize from each other. Outcomes are the achieved, measurable, short-term benefits or changes of program or project. These changes can reflect in organization's knowledge, skills and level of functioning. Short-term outcomes are necessary when aiming for intermediate outcomes or long-term outcomes, which are impacts.

According to Brennan & Levy (2006) outcomes should:

- Represent the results that occur because of program activities
- Be in the scope of program's control or sphere and in the chosen timeframe
- Be generally accepted by stakeholders
- Be phrased in terms of change
- Be measurable

The progress from short-term outcomes to long-term outcomes should reflect as impacts within 7 to 10 years (Kellogg Foundation 2004). In the case of FTA, examples of short-term outcome are R&D cooperation and networking. The long-term outcomes in this paper are considered as impacts.

3.2.5 Impacts

Impact can be conceptualised as the long-term effect of an outcome (Harding 2014). While outcome is the direct result of research activity, impact is the effect that this outcome has on society (Godin & Dore 2005). Impacts are always generated by a specific intervention and they can be anticipated or unanticipated, positive or negative, intentional or harmful (Dahler-Larsen 2005, 7). The intended and desired impacts of any project or program should be defined at the same time as the project/program objectives are defined. For an organization to generate the intended impacts, these objectives must be realistic and clear for everyone involved in the process.

Impacts can be categorized as (1) social, (2) cultural, (3) environmental and (4) economic impacts (Lähteenmäki-Smith *et al.* 2006). Societal impacts of research are referred as societal impacts in this study. These impacts indicate the research's contribution to the social capital of the nation, for example as improved policymaking (Donovan 2008). Societal impacts are long-term societal benefits, for example societal quality, usefulness, public values, knowledge transfer and societal relevance (Bornmann 2012). Previous analyses have showed that the key to generate societal impacts is based on the interaction between science (scientists) and society (stakeholders). This refers that societal impact occurs on the basis of interactive processes between the researchers and stakeholders (Bornmann 2012). Cultural impacts are additions to the nation's cultural capital, for example understanding the relations between other cultures (Donovan 2008). Environmental impacts can be seen for example as in waste reducement or pollution. The environmental benefits are additions to the nation's natural capital. And finally the economic impacts, such as improvements in productivity, are additions to the nation's economic capital (Bornmann 2012, 218).

Impacts can be further categorized depending on their complexity and variety. According to Tassey (2003) impacts can be immediate, intermediate or ultimate (Figure 5). Immediate and intermediate impacts can be considered same as short-

term and medium-term outcomes, while ultimate impacts are the long-term societal outcomes.

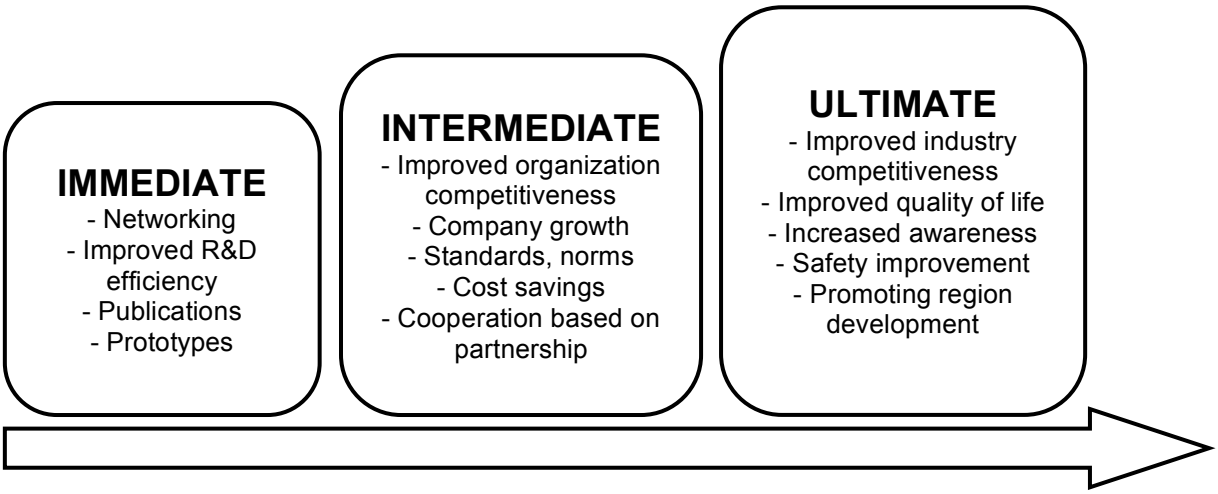

Figure 5. Timeline of impacts according to their intensity (Adopting Tassey 2003)

Immediate impacts are the direct outcomes, which can be seen as the organization's publications, prototypes or patents. In R&D operations immediate impact can occur also as improved efficiency. Intermediate impacts are those mid-term outcomes, which can be seen as company growth, improved company competitiveness and cost savings. In public R&D one important intermediate impact is the level of cooperation based on partnership, for example with universities or private companies. The final category is the ultimate long-term impact such as improved quality of life and increased awareness. The ultimate impacts are visible after several years and this is why they are the most challenging to evaluate (Tassey 2003). This study aims to recognize the key impacts of FTA's R&D operations, and especially those long-term impacts that affect to society.

3.2.6 Effectiveness

Effectiveness is a multidimensional concept and therefore often mixed with term of *efficiency.* While effectiveness indicates organization's capability to produce the intended impacts, efficiency is about using the lowest amount of inputs to create the greatest amount of outputs. According to Hyytinen & Konttinen (2006)

effectiveness is:

- Change
- Doing the right things
- Goal achievement and meeting the needs
- Both positive and negative impacts
- Both proactive and unpredictable impacts
- Societal effectiveness and ability of customer service

When discussing about effectiveness, it is important to recognize what is the context. Effectiveness can be investigated from many different levels and the relation between these levels and organization's functions depend on the organization. The levels of effective functions and effectiveness are illustrated as a pyramid with 4 steps (Figure 6). Paasio (2003) applied this model in his research on "*Effectiveness evaluation in the Human Services".* The basic idea of the pyramid is that the knowledge received from lower level is utilized in the next level.

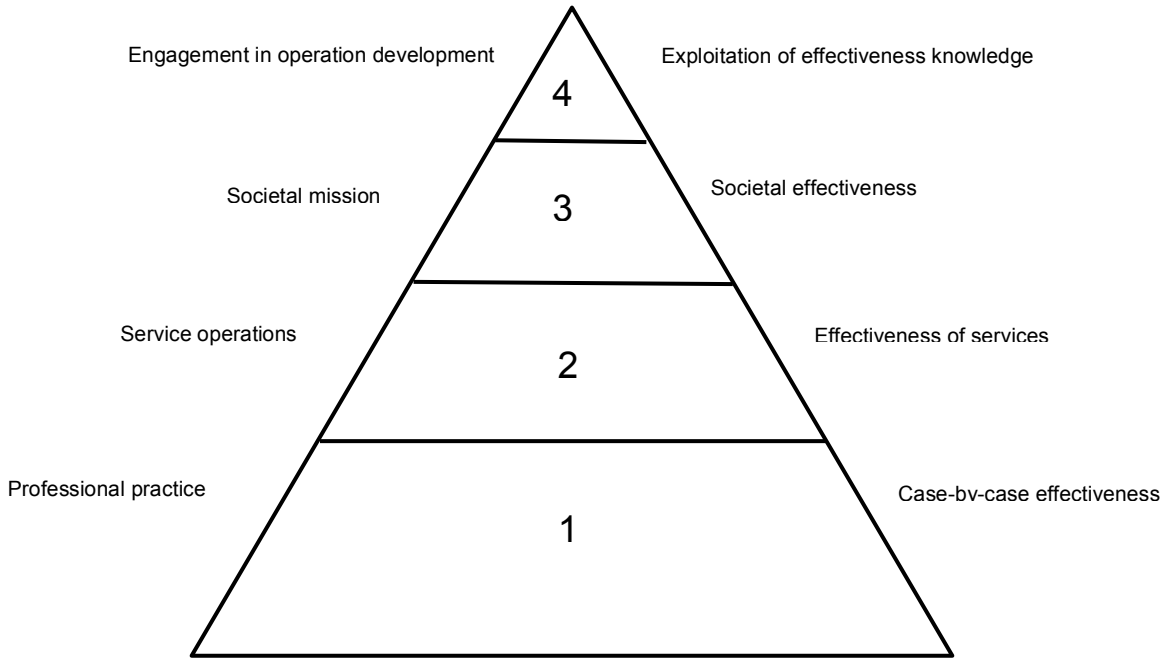

Figure 6. The levels of effective functioning and effectiveness in organization (Adapting Paasio 2003)

The basis of the pyramid is the level of professional practice, which is same as case-by-case effectiveness. Case-by-case level indicates the process between the

customer and the employee and if no effectiveness is born within this level, there will not be effectiveness in other levels either (Paasio 2003). In this study case-by-case effectiveness is the effectiveness between FTA and the R&D users. At this first level, FTA utilizes its expertise on identifying the needs of the citizens.
The second level is organizing service operations, where effectiveness is considered as the effectiveness of the services. At this level the organization's initial mission and tasks are defined to fulfill the objectives. In this study service effectiveness is challenging to recognize as FTA does not directly provide services, but the organization utilizes its expertise in recognizing the needs of their end-users for maintaining and developing the transport systems. In other words the second level indicates FTA's initial mission and tasks. The first level of the pyramid provides the necessary tools for defining the desired societal impacts of FTA's R&D.

In the third level of societal mission, effectiveness is considered as the societal effectiveness. At this level the main purpose is to answer to the question, whether the organization succeeded to implement what was the initial mission or task, which was defined at the second level (Paasio 2003). The focus is in issues such as society and social challenges. The current level of FTA's R&D societal effectiveness is investigated in this study. The social challenges and main focus areas that FTA should respond in the near future are defined in the new Government Programme 2015.

Besides the basis of the pyramid, another necessary level in effectiveness evaluation is the top of the pyramid. This level of engagement in developing the operations is about exploiting the knowledge of effectiveness that was obtained from previous levels. How this stage is understood in the organization, effects to the whole structure of the pyramid. At the top level, the professional practice, research, tacit knowledge and learning are united (Paasio 2003). In FTA's case exploiting effectiveness knowledge is not possible before the concept of societal effectiveness, i.e the third level of the pyramid is defined and identified in the whole organization. All these 4 levels and their relation with each other should be

taken into consideration when developing the effectiveness evaluation. It is also important that the actors of each pyramid level are engaged in the development process.

3.3 Selection of indicators

A clear logic chain can provide a useful map when selecting the indicators. It is crucial that the indicators illustrate both the achievement of objectives and the demands that emerge from the environment (Lähteenmäki-Smith *et al.* 2006). The indicators for evaluating effectiveness should be created based on the critical success factors of the organization. These factors indicate those issues in which the organization should succeed in order to achieve effectiveness. The challenge is to recognize critical success factors before the indicators can be selected and defined.

Because the concept of effectiveness is multidimensional, the indicators should also be variable and comparative. Each objective should have own indicator and it can be either qualitative or quantitative, but most crucially it needs to be transparent and verifiable (Lähteenmäki-Smith et al. 2006). In terms of stakeholders' expectations, effectiveness can be understood in many different ways. Therefore effectiveness becomes a socially constructed concept and indicators should have ability to reflect this social construct (Taysir & Taysir 2012). In this sense it is important to engage the stakeholders in the process of selecting indicators. In a non-profit organization, where the economic benefits are not the most relevant indicators, the social indicators become more significant. Rooney (2011) stated that those non-protfit organizations, which regularly evaluate their work and utilize the obtained data to improve effectiveness, are in better position to engage partners and other stakeholders. And most importantly, these organizations are able to serve their stakeholders better and increase impacts.

When defining an indicator, *a SMART criterion is* a good starting point. According to SMART criterion, a good indicator is Specific, Measurable, Attainable, Reliable

and Time-bound. In other words, the indicator must be specific to the objective, measurable either in qualitative or quantitative methods, realistic to attain in relation to objectives and costs, relevant to the information needs of the decision-makers and timed for understanding when the target is to be achieved.

A consortium of Finnish public R&D organizations developed indicators for analyzing R&D's socioeconomic impacts (Lähteenmäki-Smith *et al.* 2006). All together five dimensions of socioeconomic impacts were identified and the example indicators are presented in Table 1 below.

Table 1. Dimensions of impacts and indicators (Adopting Lähteenmäki-Smith *et al*. 2006)

Dimensions of impacts	**Examples of indicators**
Impacts on economy, technology and commercialisation	•Improved competitiveness •Cost-savings •Improved R&D efficiency
Impacts on knowledge, expertise, human capital and management	•Strengthened expertise •Improved research methods
Impacts on networking and social capital	•Improved networking between research partners, firms etc. •Domestic networks and global networks •Organizational and social innovations
Impact on decision making and public discourse	•Support of decision making through expert consultancy and governmental advice •Participation in legislative and strategy planning
Impacts on social and physical environment	•Reduction in material/resources and energy consumption •Promotion of regional development and growth •Promotion of safety

According to Lähteenmäki-Smith *et al.* (2006) there are some special requirements for ideal indicators as outlined in the Finnish context. Based on these requirements indicators should:

- Relate to outcomes, results and effects
- Describe strategically prioritised themes
- Be related to the core activities of the organization
- Be selected such that target values can be set in a similar scale
- Be well enough established and stable in such way that they can be monitored on a long-term basis and therefore used as an indicator of change

In this study no indicators for measuring societal effectiveness are to be created, as the case organization's critical success factors need to be identified first. The most important issue in the case of FTA, is that the whole administration has uniform understanding of societal effectiveness and its evaluation methods, and that the chosen indicators are interpreted in the same way by everyone.

3.4 Challenges

There can be seen some contradictory opinions about evaluating and measuring effectiveness between the scientists. Some of the scientists see evaluating societal effectiveness hazardous because there is a risk that the evaluation occurs only from the immediate recoverability point of view. This criticism is based on the fact that government - as the biggest financier of research - should be aware of the importance of research without any need to concretely prove it (Mustajoki 2005). According to Godin & Dore (2005) organizations must tackle 3 challenges before implementing effectiveness evaluation: first organization must distinguish the concepts of output and impact, secondly organization needs to identify the transfer mechanisms by which science translates into impact, and finally organization needs to develop appropriate and reliable indicators for evaluating effectiveness.

Previous studies have shown that the assessment of long-term impacts and the prioritisation of the R&D indicators are challenging tasks. Even if there is a high desire for developing a standardised indicators, they may not be compatible with the complex policy goals, which require a long-term perspective (Kuitunen & Hyytinen 2004). Besides the practical need for the developing the indicators for analysing societal impacts, there is also need to understand the role of impact assessment in relation to organizational development and competence-building.

Martin (2007) combiled 4 main challenges in evaluating societal impacts. First there is the causality problem, which indicates that it is not clear what impact can be attributed to what cause. Bornmann (2012) noted that the causality between a certain piece of research and a certain impact is difficult to identify in some cases.

Second challenge is attribution problem, which indicates that it is not clear what part of the impact should be attributed to a certain research. Thirdly there is the problem of internationality, which means that the attribution is difficult in practice as the R&D is intrinsically international. And finally there is timeframe problem as the measures may result in overemphasizing the short-term benefits of the evaluation.

According to Paasio (2003) the true challenge is to takeover the methods of evaluation and concept of effectiveness. By this Paasio means that a professional actor needs to be able to understand what evaluation and effectiveness concretely means in one's own professional practices. And secondly the whole work community needs to be able to implement those professional practises in which the evaluation is part of everyday operations (Paasio 2003). Effectiveness evaluation as part of daily operations, require specific culture, and changing the culture is always challenging. Organization culture is always based on understanding, values, ways of working and organizational structure. The crucial thing is to realize why something is done, when the change or challenge is great (Paasio 2003).

3.5 Synthesis of theoretical framework

Today a broader range of evidence is required as result of growing interest towards assessing the social impact of research (Steinbach 2013). Traditionally the focus of R&D evaluation has been on the inputs of the investments, until the growing global trend of evidence-based policy shifted the focus to evaluating the R&D outcomes and results. In evidence-based policy monitoring and evaluation are key issues when improving organization’s quality, efficiency and effectiveness.

The public hierarchical production system has been criticised for being too heavy and inefficient. The role of the public sector as the provider of services changed, when the public sector adopted the purchaser-provider model in their operations. According to this approach public sector is in the role of enabler and organizer of

services and the commission for these services is assigned from the principal. The main objective of the model is to change the traditional hierarchical steering into more contract-based approach. By adopting the purchaser-provider model, the public sector is able to improve the efficiency and effectiveness of the operations, which is today the keyword in public R&D.

The synthesis of the theoretical framework is illustrated in Figure 6 below. As introduced in the second chapter, he context for effectiveness is based on policy rationale and strategic planning. Strategic planning includes evaluating and combaring the current state of the organization to the ideal state of the future.

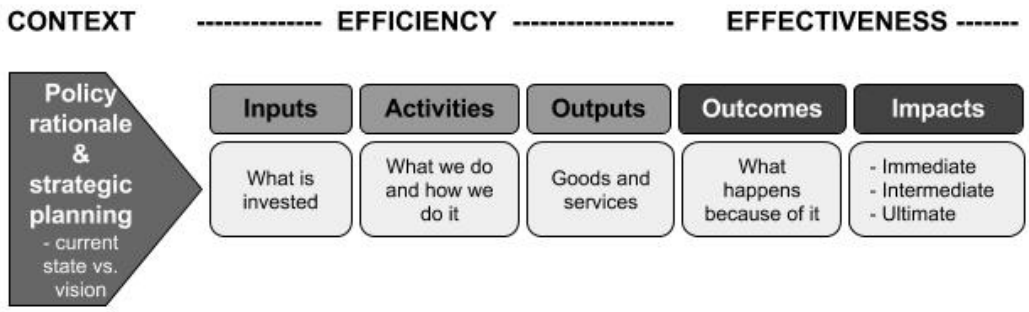

Figure 6. Synthesis based on theoretical framework

The path to impact starts by utilizing the performance measures to evaluate organization's operational efficiency. Efficiency can be evaluated through analyzing the relation of organization's inputs, activities and outputs. The level of organization's effectiveness can be analyzed by evaluating the outcomes and impacts. Effectiveness indicates whether the organization was able to produce those outcomes and impacts that operations were first aimed for. Impacts can be categorized as immediate, intermediate or ultimate depending on their timeline. In this study the focus is on ultimate impacts, which can be referred as long-term societal impacts. When all the elements of the logic model towards effectiveness are recognized, the organization is able to continue for defining and selecting the indicators to measure effectiveness.

Effectiveness evaluation and impact assessment are not only instruments for

developing practices of accountable management, but also tools for creating societal benefits for the stakeholders. Evaluation data informs about the program's performance and promotes a strong foundation for transparency and accountability (Gertler *et al.* 2011). According to Hyytinen & Konttinen (2006) the motives for effectiveness evaluation are connected to pressure of change. The change can occur in administrative culture, it can be need for reconciliation of political objectives or need for organizational development. According to Lähteenmäki-Smith *et al.* (2006) the motives can be classified as external or internal motives. External motives are based on factors such as result-driving planning or general shift to more service-oriented culture, while internal motives stand for the need to develop activities, programs and organizations. The most important external motives are the demand of results and increased effectiveness, accountability and transparency. In addition to reasons of accountability and transparency, evaluations are often made to justify those policies and actions that have already taken.

Organization's management has the main responsibility in implementing evaluations, but the true question is whether the organization wants to improve its effectiveness. From this perspective the operational objectives, plans and overall performance are crucial. When systematic information between the organization's performance and impact is lacking, the organization's productivity weakens. According to Paasio (2003) the true challenge in effectiveness evaluation is to takeover the methods of evaluation and concept of effectiveness. Effectiveness evaluation as part of daily operations, require specific culture, and changing the culture is always challenging. The crucial thing is to realize why something is done, when the change or challenge is great (Paasio 2003).

4. METHODS & DATA

4.1 Research approach

This research was chosen to carry out as qualitative research based on aim and the nature of of the study. When assessing social impacts, qualitative approach exposes the causal relation between events, processes and outcomes, and allows the indirect impacts to reveal (Wright 2003). Qualitative research aims to understand some aspect of social life, and its methods generate data as in words rather than numbers.

A qualitative research process is illustrated in Figure 7 below. The process starts by defining the research gap (i.e. why a research is needed) and determining the research questions to which the research aims to answer. Next the relevant subjects that support the research are chosen. Previous studies and theories are applied for understanding what has been founded so far related to the research topic. Next stage is collecting the data. The data can be primary or secondary data, primary refers to previous theories and studies and secondary is new data gathered for example through interviews or inquiries. In the next stage the gathered data is analysed and interpreted by the researcher. The research findings need to be reflected to the previous theories for investigating what new or different this specific research provided. And by the end of the process the final conclusions are to be made.

Figure 7. Qualitative research process (Adopting Bryman 2008)

This study was carried out as a case study. Case study is a research method that covers a wide selection of subjects such as public health, public policy, public administration and business. Although case studies have been widely used, they

have received the least guidance from all research approaches. Yin (2012) defined case study as an empirical inquiry that investigates a phenomenon within real-life context. Case study is an in-depth and detailed analysis of a specific subject and it is common method when investigating societal impacts. Although case studies do not permit generalizations, they can provide insight into processes, which have resulted in societal impact (Bornmann 2012).

According to Yin (2012) case study has two main strengths. First of all, a case study is useful when addressing questions in relation to how and why a specific phenomenon behaves. Secondly, case study is useful when exploring topics that do not have a strong theory. According to Bozeman and Kingsley (1997) a case study research is very forgiving from the researcher's point of view when learning about the observed phenomenon.

4.2 Case selection

This case study investigates societal effectiveness through 4 R&D themes which are *Pavements*, *Environment*, *Building Information Model (BIM*) and *Intelligent Transport Systems (ITS).* These themes were selected from FTA's R&D portfolio as the present different stages of maturity (Figure 8). Pavements and environment are more traditional research areas in FTA's R&D, while BIM and ITS are newer research fields.

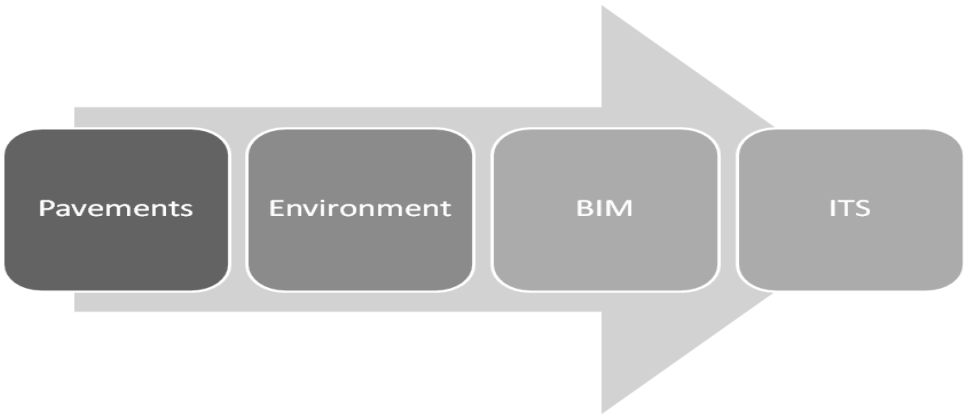

Figure 8. Maturity of FTA's R&D themes and selection of cases

A semi-structured interview was held with 8 people who operate in the above-mentioned R&D areas. From each R&D area one expert from FTA was first interviewed and these experts were asked to nominate one primary end-user from their own R&D area for an interview. The interviewees from FTA present the top experts on their own R&D themes and they are also centrally involved in R&D operations. The end-users are partners from other organizations and they are involved in FTA's R&D projects.

4.3 Data analysis

The data for the empirical part of the research was collected through interviews and observations. The goal of the interviews was to gather data of FTA's R&D in general level and the respondents' personal opinions about societal effectiveness of R&D. The sampling consisted of FTA's R&D experts and FTA's R&D end-users, which present different public organizations and companies. All together 8 people participated to the research from 4 different themes, which were pavements, environment, building information model and intelligent transport system. These themes were chosen together with FTA's R&D management.

From 8 respondents half participated to the research with face-to-face interviews and half answered to the interview questions in written form by e-mail. From each R&D themes one R&D expert and one R&D end-user participated to the research. The structure of the interview was divided into 2 themes: *FTA's R&D operations* and *Societal effectiveness of R&D*. The interview form for FTA's experts consisted of 13 open questions and the form for end-users consisted of 11 open questions (Appendix 2). Each interviewee was made aware that the aim of the research is to investigate the current state of FTA's R&D's societal effectiveness and find means for improving societal effectiveness in the future.

The interviews were held in Finnish and recorded with the respondents' permission. The interviews were transcribed later for further analysis and translated into English for the final report. The results of the interviews are analyzed in relation to the theoretical framework of the research. The findings are

illustrated anonymously by utilizing marks A and B, where A stands for FTA's R&D expert (respondent 1-4) and B stands for R&D end-user (respondent 1-4).

5. CASE: FINNISH TRANSPORT AGENCY

5.1 Overview of FTA

Finnish Transport Agency (FTA) started its operations in 2010 when Finnish Maritime Administration, Rail Administration and Central Administration of the Finnish Road Administration merged. FTA together with Finnish Transport Safety Agency (Trafi), Finnish Meteorological Institute and Finnish Communications Regulatory Authority is directed and financed by the Ministry of Transport and Communication (MTC). FTA with its key partners is responsible for developing the road, railway and waterway networks into efficient coherent transport system.

The strategy of FTA consists of common vision about the future and the goals that FTA has to achieve in terms of building the future vision (FTA 2015). FTA utilizes strategy as a tool of management in the sense that it steers FTA's long-term operations and provides guidelines for systematic development. The strategic goals of FTA are:

- Well-functioning transport routes enhance competitiveness
- Smooth and safe travel
- We act in a responsible, effective and innovative way
- The Finnish Transport Agency – an excellent workplace for professionals

FTA had an organizational reform, which was introduced in 1st of May 2015. The new organization chart can be seen in Figure 9 below. The units of Strategy and Communications and Sustainability operate under the director general. FTA's R&D operations are connected to organization's strategy and therefore the R&D coordination and management are located in the FTA's Strategy Unit.

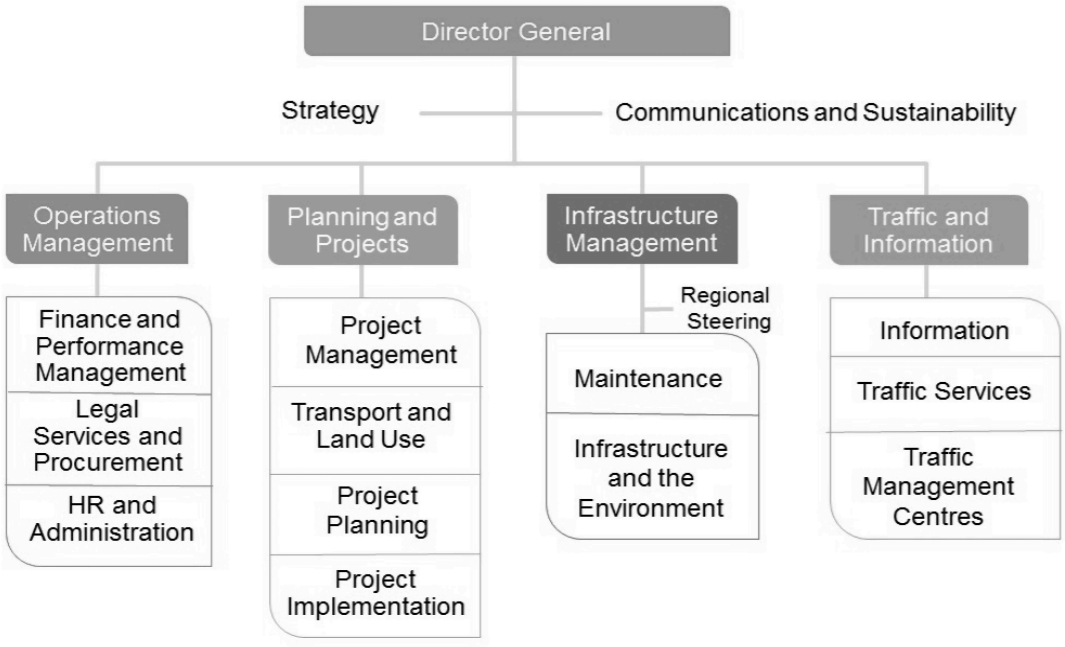

Figure 9. The organization of Finnish Transport Agency (FTA 2015)

5.2 Finnish Transport Agency's R&D

The purpose of FTA's R&D is to support the FTA's strategic objectives and the policies of the administrative sector. Ministry of Transport and Communications has delegated operative R&D to FTA and Trafi. FTA's R&D aims at creating positive impacts on society and welfare and the experts of R&D are responsible for identifying the R&D needs. FTA's R&D is currently developed towards more dynamic and proactive system aiming at responding more openly and agilely to the challanges of future transport and mobility. The current focus areas of FTA's R&D operations are illustrated in Figure 10 below.

Figure 10. FTA's current R&D focus areas (FTA 2015)

The first focus area is evolving databased mobility services, which includes testing new traffic innovations and evolving databased solutions and services. The second focus area of intelligent infrastructure and a well-functioning transport system consists of themes such as new ways of collecting and measuring data, allocating resources more precisely and adaptation of service levels to the needs of society. The third area is effective transport infrastructure management and lifecycle management. This includes creating innovative solutions and procedures, systematizing the life cycle approach and improving infrastructure network structures and safety. The fourth R&D focus area is the adjustment to climate change and its mitigation, which includes enhancing creation of energy-efficient and cleantech innovations and ability to anticipate changing weather and climate conditions. The R&D focus areas are flexible and changes are made when needed due to e.g. changes in the transport policy, strategies, operational environment and new discovered opportunities.

This research focuses on studying societal effectiveness of FTA's R&D through four R&D areas, which are Pavements, Environment, Building Information models and Intelligent Transport Systems.

6. RESULTS OF THE INTERVIEWS

6.1 R&D guidelines & objectives

The focus of the interview's first part was to clarify the case organization's current R&D guidelines and objectives from the FTA's experts' point of view. The R&D experts were asked to describe what is the current strategy of FTA's R&D and how well the R&D objectives have been emphasized in the strategy. Half of the FTA's experts considered that R&D objectives are brought out relatively well and that those R&D projects, which support organization's strategic objectives have been more likely taken forward. Other half of the respondents were not well aware of the current R&D objectives as one of the expert also noted:

"I actually do not know exactly what are the objectives of our R&D."

This comment indicates that there is a need to clarify the strategy and objectives inside the organization and especially among the people who are actively involved in R&D operations. As mentioned earlier, the first step in creating societal effectiveness is to clarify the strategy and objectives so that is clear for everyone in the organization. If there are any doubts of why or what for R&D is done, the intended impacts cannot be defined or achieved and most importantly no effectiveness will occur. Respondent A_3 commented that because FTA is a young organization, it has not yet concretized the R&D objectives in the organizational level.

All the respondents were also asked to describe how well the end-users have been notified in the R&D processes. According to FTA's experts about how well the end-user's point of view is emphasized in FTA's R&D depends on the project. In some cases the end-users are better involved in the R&D process and projects might even include specific parts that concern only the end-users. According to respondent A_1 the R&D end-user's point of view is not necessarily always considered in the best possible way and this issue needs to be improved. Respondent A_4 noted that when the end-users are involved in the R&D process,

they have a chance to direct the R&D into the way that serves them best.

> *"The concrete outcomes and effectiveness can be achieved when the end-user is kept in mind and in some level also involved in the project."*

Respondent A_3 criticized that good R&D is not about letting the "ordinary men" to involve in R&D operations. This response indicates that only the experts could do R&D and the more ordinary people should be left out of the processes. As mentioned in the theoretical part of this research, user-involvement is crucial in public R&D, which aims for improving society and welfare with the taxpayers' assets. Therefore the R&D end-users should be committed to the transport sector's R&D operations.

Respondent A_2 clarified that in the R&D projects there is a steering group of R&D purchaser, provider and end-user, which can be a consultant or contractor depending on the project. This steering group aims to ensure that R&D provides information that can be utilized by the end-users. According to respondent A_2 FTA's R&D aims to be more solution-oriented and in some cases the respondent has given more freedom to the researcher, which has led to more in-depth R&D.

> *"Now when everything is more hectic and more faster, everything should happen faster with smaller resources and the R&D is not always bringing out its best. More resources should be put in R&D so that there is a possibility to clarify the problem and observe something that might not be seen when operating faster and linear."*

In general level the end-users were satisfied with the cooperation with FTA and commented that their views have been quite well notified in the common R&D projects. Respondent B_4 has been involved in several PPP-projects and other cooperation projects with FTA and described that the communication has been only bilateral with FTA. Respondent continued that other type of communication has not occurred or then the information has not been reached. Respondent B_4 noted that at its best, the projects contribute both parties but if the goal is "only" to fulfill the need of the authorities the motivation for the cooperation comes straight

from the compensation paid by the authority. Respondent B1 wished that the research needs and expectations should be more mapped from the transport sector. The assembly of the steering group should be chosen carefully and ensure that everyone participate actively so that the research provider can not direct the research according to one's personal interests.

The R&D end-users were asked to evaluate on scale 1 to 10 what is the relevance of FTA's R&D as societal influence. The results are seen in Table 2 below. The answers variate between grades of 5 and 10. According to respondent B_2 the impacts that FTA's R&D has created are not particularly significant in societal context, therefore the respondent gave 5. Respondent B_4 was satisfied with the results and impacts that FTA's R&D has created, as the grade 10 also indicates.

Table 2. Relevance of FTA's R&D in the society

Respondent	FTA's R&D's influence in society
B1	**8**
B2	**5**
B3	**7**
B4	**10**

Some of the end-users noted that they have not been informed enough about FTA's R&D objectives or its operations in general. Cooperation between public and private sector is important when aiming for societal impacts and therefore the private sector must be aware and have a possibility to involve in public sector's R&D operations. Active communication between the actors of the field is extremely important also when avoiding ineffective duplication of work.

6.2 Motives for effectiveness evaluation

FTA's experts were asked to describe how relevance they consider societal effectiveness evaluation and what are the main motives for evaluation. Respondent A_1 noted that R&D is extremely important area and because the

financial resources are decreasing, it is even more crucial to clarify why R&D is done for. Therefore there is a clear need for developing the evaluation design with proper indicators for evaluating societal effectiveness.

The experts of FTA were asked to evaluate on scale 1 to 10 what is the relevance of societal effectiveness evaluation of R&D according to their personal opinion. The results are illustrated in Table 3. Most of the experts considered societal effectiveness evaluation of R&D important with grade of 8. According to respondent A_3 there are several other issues that should be evaluated also, and these issues do not necessarily have societal impacts. Therefore the respondent gave grade of 6.

Table 3. Relevance of societal effectiveness evaluation in FTA

Respondent	Relevance of societal effectiveness evaluation in FTA
A1	**8**
A2	**8**
A3	**6**
A4	**8**

6.3 Inputs

As discussed in the theoretical part, the logic model of impact starts from organization's inputs. Based on the results, the inputs of FTA are currently highly limited and therefore the R&D tasks are done when there is enough time. The operational tasks are the primary tasks that need to be done first. According to the FTA's experts the resources are narrow as there is not enough personnel and budget for the operations is too small. Because the organization needs to prioritize its R&D projects, it is challenging to choose the most important projects and sometimes areas that should be investigated have to be left out from the research scope. All the FTA's experts agreed that R&D is an important part of FTA's operations but it is challenging to find enough time or people for the tasks. This

has led to lack of knowledge of what else happens in the organization as one respondent commented:

"I haven't had much time to familiarize with what else is done in our R&D"

After the merger in 2010, the need for more extensive expertise and professional practise has increased in FTA, as the agency is now responsible for developing the whole public transport system of Finland. The limited resources raise the risk that the quality of R&D weakens and personnel become overloaded, which apperantly is already happening according to the FTA's experts.

6.4 R&D activities and research themes

As mentioned earlier, R&D activities are only one part of FTA's operations. To clarify the FTA's R&D process, the experts were asked to describe what is the starting point for R&D to be purchased. All respondents agreed that the R&D process starts from identified need. This need can be either a need to develop or a need to improve a specific issue, process or procedure. Respondent A_1 highlighted that R&D should not be done only in means of "just doing something", but the concreteness must always be realized which means committing to the everyday operations.

According to respondent A_3 the aspect of societal effectiveness needs to be included in the beginning of the process and even if the R&D results would not directly be utilized in the FTA they might help the private companies. Respondent A_4 noted that in the starting point of R&D project, one should always consider what new the organization can achieve as result of the project and what is the quality and relevance of that new information.

The experts of FTA were asked also to describe the new research ideas are brought up in FTA and how the R&D projects are chosen. The results indicate that FTA's the effort in identifying and highlighting new research themes has improved

during the last few years. Respondent A_1 described that FTA is now more active in finding new partners because the network has grown both nationally and internationally.

6.5 Outputs

One key output of FTA's R&D is implementation and dissemination of R&D data. The R&D experts and the R&D end-users were asked to describe how well the obtained R&D data and results are implemented and disseminated after R&D project has been ended. Most of the interviewed FTA's experts noted that the implementation requires more effort and still too many R&D reports are left in the shelfs. Piloting projects are currently implemented well, but the implementation of more abstract R&D projects require more effort. Respondent A_3 noted that implementation is considered as a challenge even if the results are good, because the organization is lacking a systematic implementation plan.

According to the respondents there is no clear channel for disseminating the R&D data and results and this is why the dissemination is ineffective. According to respondent A_1 only some of the project managers share information about FTA's operations and the key research results in different public forums. Some of the R&D end-users also agreed that FTA's R&D dissemination has not been effective enough. Most of FTA's experts commented that because people are so overloaded it is difficult to follow what is happening inside the organization, let alone outside the organization. This will, and already has, lead to duplication of work. Because the R&D actors have not been able to communicate enough with each other they might be doing partly equal R&D, which is ineffective and costly. Some of the experts commented that the R&D data is difficult to find even inside the organization. There is a need for improving the communication and define a clear channel or system where the data can reach everyone interested. Also respondents B_1 and B_3 noted that the main challenge is to find the right ways to disseminate the good practises quickly and easily for everyone after the project has ended. One of the end-users noted that the R&D results are well disseminated

and implemented in larger cities but not in whole municipal level.

The respondents were also asked to evaluate the quality and usability of FTA's R&D data and results. Most of the FTA's experts believed that FTA's R&D produces such data that the actors such as consultants, businesses and research institutes can utilize, but to what extent was considered as another case. Respondent A1 noted that this area still requires more effort and the right channels and tools need to be defined for the R&D dissemination, as mentioned earlier. These channels would increase FTA's visibility by connecting the experts of the field and informing the stakeholders about the R&D processes.

Respondent A_2 noted that the decision of what type of R&D should be made depends on the needs identified by the expert. By this the respondent means that the usability should be considered from expert's point of view: what kind of R&D data the expert needs in order to proceed in one's own projects. This refers that the obtained data should be usable for both the R&D purchasers and the R&D end-users and that FTA is only responsible for commissioning such R&D that serves FTA's objectives and operation development. The respondent also noted that FTA's experts need to be able to point out to the users of R&D where the relevant, usable, R&D data is available.

Respondent B_4 highlighted that from the commercial point of view such R&D that enables business operations is the most central issue. The private companies are aiming at effective business operations and higher returns, and therefore R&D is considered from different aspect than in the public sector. According to majority of the end-users the quality of FTA's R&D data and outputs is satisfactory but needs improvement. Respondent B_1 critized that the quality and usability vary a lot depending on the project, and that although R&D produce new knowledge this data is sometimes difficult to utilize.

The results related to visibility and availability of R&D data and results vary a lot. According to respondent B_3 the data and results are difficult to find because there

is no clear channel. On the other hand, respondent B_4 noted that data and results are visible in the FTA's web pages and as the results are sometimes shared in seminars the level of visibility is quite good. Respondent B_4 also highlighted that FTA has enabled the usability of data by following the open data principles. According to these principles data must be accessible in digital format without limitations based on user identity and free of restriction on use.

6.6 Impacts

All together 3 questions were asked to clarify the R&D's strategy in creating impacts: which party defines the intended impacts, how well the intended impacts are emphasized, and has FTA's R&D produced these intended impacts according to both the experts and end-users. Respondent A_1 described that the specifications for intended impacts come from many levels, namely strategic level and also from EU. FTA's mission is to identify the key issues from these specifications, and based on that form the research programs and objectives to respond expectations. One of the FTA's experts was not aware of which party defines the intended impacts and noted that this is due to weak level of interaction between the R&D experts and R&D management.

> *"It would be good to clarify for us who purchase R&D that what is our current strategy and what kind of projects we are going to support within the next five years. But there should also be a possibility for me, as involved in the substance, to explain what I see that is the R&D need. That kind of dialog we should have more between the R&D experts and R&D management."*

According to respondent A_4 the intended impacts should be set and directed from the board of directors because R&D reflects - or should reflect - how the management sees FTA's operations now and in the future. The management should identify the megatrends and the R&D operations directed towards these megatrends. Currently the intended impacts are set and defined by the R&D experts depending on what are the most important issues in one's own field of operation. This indicates that there is no uniform understanding or definition of the

desired impacts, towards the operations should be directed. As mentioned in the theoretical part, the intended impacts should be regonized in parallel with the start of the project. Effectiveness evaluation measures only whether these intended impacts have been achieved when the program or project has ended. Respondent A_4 proposed that the management could set the major policy and then the R&D experts would be responsible for defining the single themes, as they know their own fields best.

In general the results indicate that there is need to put more effort in defining and highlighting the impacts to which the operations should be aimed for. For example one of the respondents was not sure whether the R&D impacts have been highlighted in any particular way and what these impacts even are. As mentioned in the theoretical part of this study, the intended impacts should be clear in the beginning of any project or program and the operations should be focused to create these impacts and improve effectiveness. Also respondent A_3 noted that the intended R&D impacts are weakly emphasized and the problem is that the effectiveness of FTA's operations is rarely examined afterwards. Yet the effectiveness of R&D operations has not been evaluated because there is no systematic evaluation method or indicators for evaluating the societal impacts. One expert from FTA emphasized that:

> *"If we want to create indicators for producing societal effectiveness, it means we need to gather around and discuss about it together."*

It is important that the indicators are defined together in the organization, so that they are understood in same ways and that the methods are uniform in all organizational levels.

The question of whether FTA's R&D has produced the intended impacts was considered difficult according to many respondents. Respondent A_1 noted simply that more effort is required in order to achieve the intended impacts. The confrontation of operations and objectives is not always simple and this issue needs reinforcement so that the FTA would be able increase its effectiveness.

Respondent A_2 described that the project objectives have been in many cases achieved but about achieving the intended impacts the respondent was not able to comment. Based on the interviews the term *impact* was often mixed with *objective* and this is an issue that should be clarified in FTA.

One of the respondents described that the achieved impacts have reflected in issues such as exporting intellectual capital, meaning consultants disseminating the Finnish standards internationally. Also respondent A_4 agreed that R&D projects have generally created the desired impacts although the projects do not always proceed in desirable way and the risk of failure always exists:

> *"I think that in proper R&D risks need to be taken if one wants to achieve something new and innovative, and it includes the risk of failure. One good indicator of R&D innovativeness is to measure how much failures there has been, if everything succeeds, then it can be said that the program has been only a bit innovative."*

Respondent B_4 described that the desired and achieved impacts should be evaluated in project level but is not sure whether the projects have produced the intended impacts. Respondent B_3 noted that the projects objectives have been sometimes intentionally high and the cooperation should start more quickly. The respondent hoped improvement in implementing the new approaches.

6.7 Societal effectiveness

The respondents were asked to describe what societal effectiveness of R&D indicates and the answers are illustrated in Figure 11. According to respondent A_1, R&D's societal effectiveness starts from the project's objectives in other words there has to be a clear picture of what for the research or clarification is done when the project starts. This was also mentioned earlier in the theoretical part, that the path to effectiveness starts from defining clear and realistic objectives.

Respondent A_1 noted that because FTA is using the assets of taxpayers, these assets should be used as smartly as possible for bringing benefits for the society. This refers to societal impacts, which may occur after several years after a project

or program has ended. Societal impacts are those impacts that will ultimately affect to society for example by increasing competitiveness and improving employment issues and welfare.

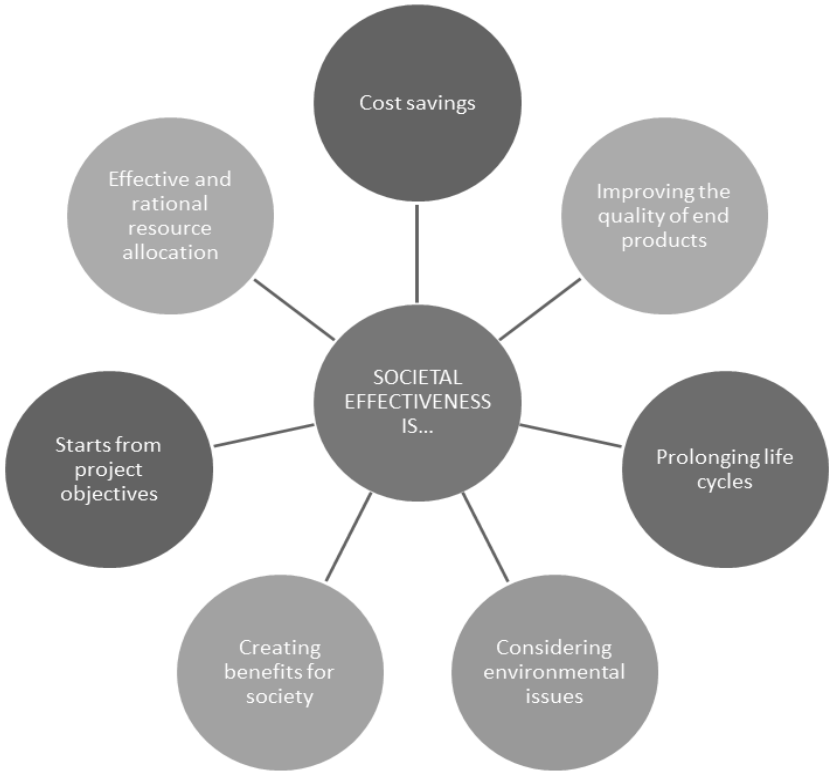

Figure 11. What is societal effectiveness?

Respondent A_2 analyzed societal effectiveness first from more economic aspect as cost savings, improved end products and longer life cycles. Secondly the respondent noted the environmental issues, that effectiveness can reflect in issues such as utilizing the resources more effective or decreasing the emissions. Respondent A_4 described that R&D produces always either new information, products or services. The obtained new information should be utilized to improve the future operations and decision-making.

Respondent B_3 described societal effectiveness of R&D as such operation that has impact in decision preparation and decision-making. Respondent added that societal effectiveness is also about the impacts to end-users, public economy and

private business and these impacts can be seen as improved quality of service and more efficient operations. As the results indicate, societal effectiveness can be viewed from economic, environmental and social aspects.

6.8 Challenges in societal effectiveness evaluation

The FTA's experts were asked to describe the key challenges in measuring societal effectiveness. According to respondent $A_{4,}$ the first challenge is to define and recognize the indicators for measuring societal effectiveness. In transport system there are already some indicators such as emissions and time spent in the traffic, which could be applied also in effectiveness evaluations. Currently there are discussions in FTA about moving into "level of service" thinking that is, FTA aiming to ensure a specific level of transport service. Respondent A_4 noted that these indicators applied in evaluating transport services could be also applied in R&D activities for measuring how effectively these levels of services have been achieved.

On the other hand, the key challenges that were mentioned by the R&D experts were mainly related to question of how to develop the common measures as a part of the everyday operations. According to the respondent A_1 one of the key challenges is how to allocate limited resources more efficiently for optimizing the costs and benefits. To implement societal effectiveness evaluations as part of daily operations is also challenging because of limited resources. As Paasio (2003) also stated, first the professional actor needs to understand what evaluation and effectiveness means in own professional practices and secondly the whole work community needs to be able to implement those professional practises in which the evaluation is part of everyday operations. Effectiveness evaluation therefore requires a specific culture, and changing the culture is always challenging. Before the societal effectiveness evaluation can be implemented, the organization - if not the whole administration - should gather for discussing about the common methods of evaluation and the proper indicators for societal impacts. As mentioned earlier, the indicators for evaluating effectiveness should be based on

the critical success factors of the organization. These factors indicate those issues in which the organization should succeed in order to achieve effectiveness. This study aims to recognize the success factors of FTA's R&D.

6.9 Improving societal effectiveness

The respondents were asked to express their personal opinions on how FTA can improve societal effectiveness in its R&D operations. Respondent A_1 highlighted communication as one of the key issues in increasing societal effectiveness. Also recognizing the challenges and focusing resources in the right issues is another crucial factor. These all are issues that can be recognized as the critical success factors of FTA's R&D. In other words, the organization needs to succeed in communication, identifying the challenges and resource allocation i.e. priorization in order to improve societal effectiveness. Respondent A_3 highlighted that one means of improving societal effectiness would be through producing more such R&D (data, results etc.) that would make comparisons in international level possible.

Respondent A_1 added that with the current resources finding the balance between the daily operations and R&D is extremely challenging. As most of the interviewed R&D experts noted earlier, the scarcity of resources forces to strong prioritazion in everyday operations. Therefore it is important that FTA clarifies its current R&D guidelines and invests in effectiveness evaluation to identify how well the organization has succeeded in R&D. According to respondent A_4, societal impacts could be one criterion when selecting R&D themes and projects. In this case the societal effectiveness could be evaluated case-by-case as illustrated in the pyramid of effectiveness in chapter 3.

Based on the interview results there are also unclarities with the terms of societal effectiveness. Although effectiveness is required on the half of many parties, the concept is not familiar with all the actors inside FTA. Also one of the FTA's experts noted that the concept of societal effectiveness needs to be clearly defined first

inside the organization before it is possible to improve it.

> *"I can't answer to that question of how to improve societal effectiveness because I don't know what it is at the moment in our case."*

One of the experts from FTA highlighted the importance of cooperation in R&D operations. According to this respondent, it is important to remember that FTA does not have all the competences and skills and that cooperation and interaction are key issues in the the future operations:

> *"We should have lot more dialogue with international R&D programs and see how our programs meet them and of course we should discuss with the companies as we are more and more going to the direction of service society in which we purchase more services from the private sector. Lot more interaction should be in both directions to understand what they think that should be done and what kind of R&D they are doing by themselves so that there wouldn't be duplication of work."*

The interaction between the actors of the field is crucial especially when trying to decrease duplication of work. Conducting R&D can be expensive and time consuming, therefore different parties should be able to actively follow what kind of research has been done or is currently going on. As respondent A_3 also noted, there could be an international forum for discussing about the current R&D trends and ongoing projects for being up-dated. Respondent B_4 also agreed that there could never be too much dialogue between the public and private sectors. The operations should be more predictable, transparent and stable, and this requires clarification of FTA's role in the transport sector. This can be seen also from the end-users responses as for example one of the end-users commented:

> *"In many issues there are still introvert thinking and things are developed only from one's own perspective and the entity is not seen although the holistic view in development would serve everyone."*

Also respondent B_4 noted that sometimes it is challenging to cooperate with FTA because the way of doing is too "official", which reflects especially as long processes and slow decision making.

7. DISCUSSION

The purpose of this research was to investigate societal effectiveness of transport sector's R&D operations. The research aimed to answer to following main research question:

> *RQ: What are the critical success factors for improving societal effectiveness of transport sector's R&D?*

And to following 4 sub questions:

> *SQ1: What are societal impacts in the case of Finnish Transport Agency?*
> *SQ2: What are the motives to evaluate societal effectiveness?*
> *SQ3: What are stages of effectiveness evaluation?*
> *SQ4: What are the key challenges in societal effectiveness evaluation?*

The answers to these research questions were gathered from the empirical part of the study, in other words based on the 8 interviews.

RQ: What are the critical success factors for improving societal effectivenss of transport sector's R&D?

Critical success factors are those variables in which the organization must succeed in order to realize the strategy and achieve the objectives, in other words success factors are the linkages between the strategy and indicators (Kaplan & Norton 2003). Organizations must perform the activities associated with critical success factors at the highest possible level (Rouse 2014). In this study the critical success factors can be drawn up from the findings of the interviews. In the case of FTA, the regocnized critical success factors are resources and priorization, clear strategy and objectives, internal communications, cooperation between public and private sector and R&D implementation and dissemination.

Resources and priorization

Based on the interviews, the resources of FTA's R&D operations are limited and especially the amount of personnel was considered too low in relation to the amount of duties. Because of scarce resources, the experts are not able to follow closely what kind of R&D is done at the organizational level, not to mention in national or international levels. Although the scarcity of resources was considered as weakness by the respondents, it can also be a strenght what comes to the level of organization's innovativeness. While the resources are limited, it forces to priorization and developing new, more effective ways to operate and manage the daily tasks.

Effective resource allocation is one of the key challenges in FTA's R&D. The organization needs to succeed in allocating the scarcity of resources in a way that besides the operative tasks, the R&D tasks could be also done in a way that the objectives can be achieved and the intended impacts can be produced. In priorization it is crucial, that the organization has common view of the future operations and guidelines. This means that the long-term strategy and R&D objectives need to be clear in all organizational levels.

Clear strategy and objectives

Based on the interviews, there is need for clarifying and strengthening the strategy and objectives of FTA's R&D. As the results indicated, some of the R&D experts were confused with the current R&D guidelines and R&D objectives. Especially some of the interviewed R&D experts were hoping for clarification of the R&D guidelines in long-term aspect. Strengthening and clarifying the guidelines and objectives of R&D could improve the long-term anticipation in the R&D operations. The current R&D guidelines are based on FTA's strategy, which in turn is formed in line with current transport policy. The R&D focus areas need to be therefore flexible, as changes are made when needed due to changes in the transport policy. Besided clarifying and strengthening the R&D guidelines and objectives,

the organization must also succeed in internal communication in order to effectively implement its mission and achieve the objectives.

Internal communications

Based on the interviews, one key challenge in FTA is internal communications. According to the interviewed R&D experts, organization is lacking a proper communication plan and therefore the communication is currently ineffective. The channels for effective communications were not clear according to the experts, and these channels should be defined together with the R&D management. Based on the interviews, the current level interaction between the R&D experts and R&D management should be improved. Also the interaction between the R&D experts requires effort, as it seemed that the experts are not well aware about what type of R&D the colleagues are currently doing. Therefore there is a high risk of duplication of work, or in worst case that the most central research areas are left out of the scope. As the fact is that the R&D resources are decreasing further, it is even more crucial that the amount of duplication of work is zero or at least at its minimum.

In the near future an internal communication plan with effective channels should be defined in order to improve FTA's R&D. If the communication is weak inside organization, it will most certainly reflect in the cooperation with the private sector. Especially while the resources are limited, the cooperation between the R&D actors of the field is vital.

Cooperation between public and private sector

The findings of this research support the earlier criticism that the cooperation between public and private sector is still limited and introvert. Some of the interviewed end-users critized that the operations of FTA are not visible enough and therefore improvement in the R&D cooperation and interaction is needed. Especially in the common R&D projects, the aspect and needs of R&D end-users

need to be emphasized better. For improving the level of cooperation and FTA's visibility, public R&D conferences could be organized annually for the stakeholders. This way the stakeholders could have a chance to involve in the FTA's R&D operations. Secondly a common channel for disseminating the R&D data and results between the R&D actors of the field could be developed. Through a common channel the R&D data could be available and accessible to everyone. As Rooney (2011) noted, sharing the results — good or bad — with stakeholders is a critical part of assessing organization's effectiveness. Therefore dissemination of R&D data and results is crucial not only for improving the level of cooperation, but also in improving organization's effectiveness.

R&D implementation and dissemination

Based on the interview findings, the implementation and dissemination of the R&D needs improvement in the future. Still some of the R&D reports are left in the shelfs and not implemented in order to improve R&D and future operations. Based on the interviews, implementation is considered challenging even if the R&D results are good, because lack of systematic implementation plan. A proper implementation plan could ensure also that the R&D data and results are not left in the shelfs. Also some of the end-users commented that the FTA's R&D results are well implemented in larger cities but not in municipal level. Further discussion is yet required for finding solutions to wider implementation of R&D data and results.

"*What? So what? Now what?"* are the 3 basic questions of any project or program evaluation. The purpose of these questions is to emphasize the stage, when a project or program has ended. According to the findings and observations in FTA, yet too often the projects are finished when a report or publication is produced. The obtained results and impacts need to be evaluated in order to investigate the current state of organization's effectiveness and how it can be improved in the future. Although the long-term societal impacts of R&D can be seen after several years, now it is the time for creating the qualifications to produce those impacts.

SQ1: What are societal impacts in the case of Finnish Transport Agency?

To answer what are the societal impacts of FTA, the path to effectiveness needs to be recognized first in the case of FTA's R&D. Based on the results of the interviews, an effectiveness model for FTA can be developed as illustrated in Figure 12.

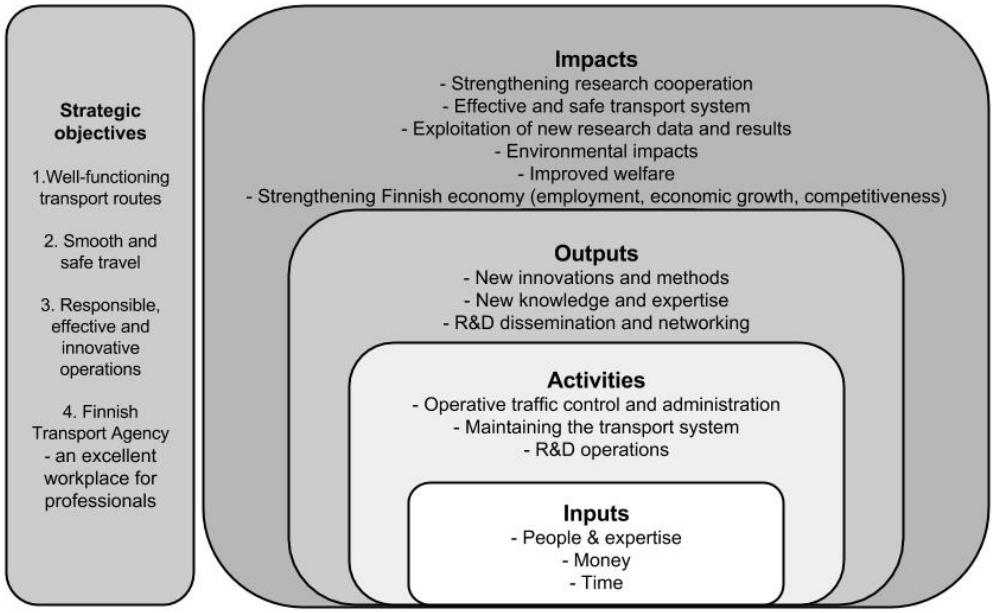

Figure 12. Proposal for FTA's effectiveness model (Adapting Hjelt et al. 2011)

This effectiveness model includes 5 main parts. First of all the FTA's strategic objectives are presented in the left side of the model. These strategic objectives are (1) Well-functioning transport routes, (2) Smooth and safe travel, (3) Responsible, effective and innovative operations and (4) Finnish Transport Agency – an excellent workplace fo professionals. The path to impact follows the logic model with its main concepts of inputs, activities, outputs and impacts. In the case of FTA the main inputs are the resources such as people (expertise), money and time. These resources are utilized in the organization's key activities, which are in FTA's case operative traffic control and administration, maintaining the transport system and R&D operations. The outputs obtained from these activities are new innovations and methods, new knowledge and expertise and results dissemination and networking. Based on the interviews, the societal impacts of FTA's R&D are

strengthened research cooperation, effective and safe transport system, exploitation of new research data and results, environmental impacts, improved welfare and strengthening the Finnish economy for example by increasing competitiveness and improving the employment situation. The societal effectiveness of FTA's R&D indicates the organization's ability to produce these previous mentioned societal impacts.

SQ2: What are the stages of effectiveness evaluation?

As the concept of societal effectiveness is multidimensional, it can be understood and defined in many ways. Therefore the most important thing is, that the organization has uniform understanding of the concept before the evaluation can be implemented. In the case of FTA, the results indicate that there is a need to strengthen the knowledge base of societal effectiveness and its evaluation. It is important that the means and indicators for evaluating societal effectiveness are commonly accepted in the whole administration under MTC.

Based on the findings of this research, there is not yet any clear way of defining the intended impacts of R&D. The interviewed R&D experts were not well aware of what the current desired impacts are, as the requirements are coming from several different directions. Therefore the FTA's board of directors should have a discussion of how to define the intended impacts of R&D in the future. As mentioned earlier by one of the respondents, the board of directors should define the impacts because R&D should reflect how the board of directors sees FTA's operations now and in the future.

Figure 13 illustrates the effectiveness evaluation cycle with 6 steps. This figure is based on combination of theoretical background and results of the empirical part and prepared for FTA's R&D. The first stage in effectiveness evaluation is need assessment to recognize the needs of the society. Also the objectives of the evaluation, i.e why the evaluation is to be made, should be defined in the very beginning. The second stage is choosing the appropriate evaluation method. The method can follow qualitative or quantitative approach or even both depending on

the case and context. The third stage is specifying the target's (i.e program, project, policy etc.) objectives that will be evaluated. Step four is identifying the impacts and impact mechanism of the target to be evaluated. At the fifth stage the implementation of the operations is evaluated. Evaluation of implementation is crucial because it provides information about how the future operations could be developed and improved. At this stage organization is also able to recognize those functions that are not effective and should be eliminated from the operations. And finally step 6 is the future aspect meaning what should be happening next. At this stage the previous mentioned 3 basic evaluation questions should be considered: What, What for, What next? This means that the evaluation data should not be left in the shelfs but utilized for future operations and improving effectiveness.

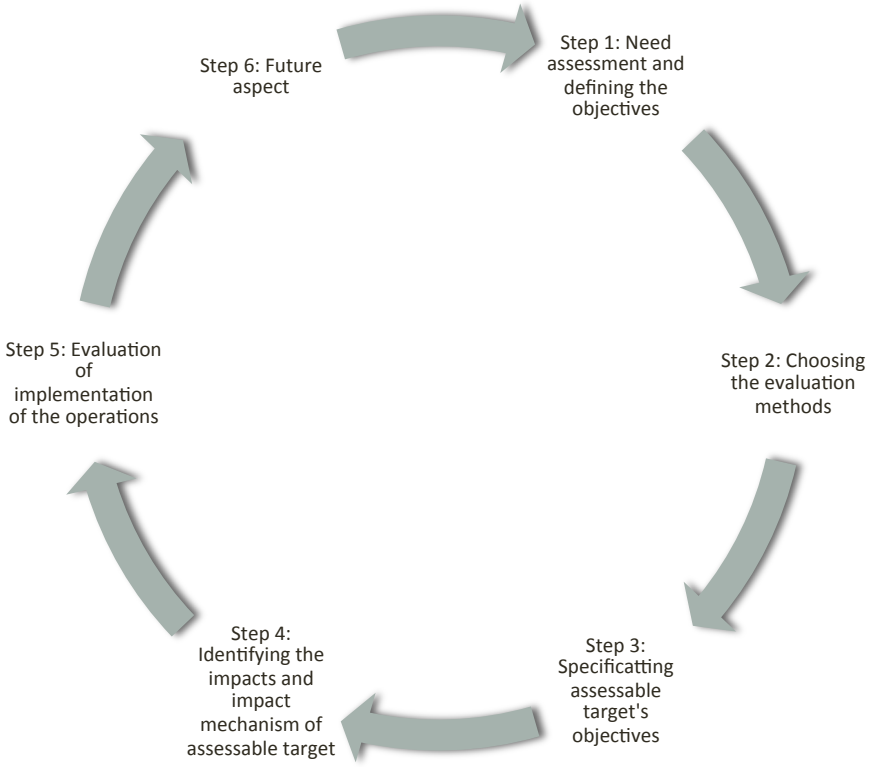

Figure 13. Stages of effectiveness evaluation (Adopting Hyytinen & Konttinen 2006)

SQ3: What are the motives to evaluate societal effectiveness?

In this research the relevance of effectiveness evaluation was considered from many aspects. With the evaluation data the organization could improve

organization's visibility, economy, welfare, safety and environmental issues within its operations. Lähteenmäki-Smith *et al.* (2006) classified motives to evaluate societal effectiveness as external or internal motives. External motives are based on factors such as result-driving planning or general shift to more service-oriented culture, while internal motives stand for the need to develop activities, programs and organizations.

Based on the interview results, the most important external motives in case of FTA are the demand of improved effectiveness and transparency. Improvement in effectiveness is required not only by the organization's stakeholders but from the principal as well. As the fact is that the resources are decreasing further, it is even more crucial to be able to justify why R&D is done. The central internal motives mentioned in the interviews were future development of R&D operations and internal learning. The evaluation data could be utilized in order to recognize and eliminate the ineffective parts of R&D for improving the long-term performance.

SQ4: What are the key challenges in societal effectiveness evaluation?

Based on the findings of this research, societal effectiveness is considered as important part of R&D, but the implementation of societal effectiveness evaluation was considered challenging for several reasons. First of all the limited resources such as people, money and time are forcing to strong priorization in the R&D operations. The scarcity of resources sets challenges in implementation of effectiveness evaluation and linking it as part of R&D operations. Therefore better resource allocation would be required in order to implement effectiveness evaluation. Secondly, societal impacts are visible after several years and due the long timeline the evaluation of these impacts is challenging. The challenge in ex-post evaluation is the time scale as the long-term impacts can be seen after several years. Therefore concrete case studies on impact chains and long-term impacts are needed. Analyses should be made from both long-term focus areas and to launch new areas. Thirdly, because there is no concensus on how the concept of societal effectiveness in understood, it is difficult to define the

measures. As Paasio (2003) also stated, first the professional actor needs to understand what evaluation and effectiveness means in own professional practices and secondly the whole work community needs to be able to implement those professional practises in which the evaluation is part of everyday operations. This raises the fourth challenge of how to engage the whole organization to applying effectiveness evaluations and how to win the possible resistance of change of the personnel. Especially when operating in hectic field with limited resources, the resistance towards learning and applying any new method is inevitable. This is the stage when the role and expertise of management is tested. But before the societal effectiveness evaluation is even possible to implement, the organization should gather to discuss about the common methods and indicators for measuring societal impacts of R&D. Based on the interviews the fifth challenge is the selection of proper indicators for measuring societal impacts. As mentioned earlier, these indicators should be based on the critical success factors of the organization. FTA's critical success factors were recognized based on the interviews as resources and priorization, clear strategy and objectives, internal communications, cooperation between public and private sector and R&D implementation and dissemination.

8. CONCLUSION

R&D is in its ideal position when it is utilized as a strategic tool in order to implement organization's mission and to achieve the strategic goals. However, the hectic environment and decreasing resources sets challenges for implementing R&D efficiently. First of all, R&D projects should be carried out as quickly and effectively as possible. The new obtained R&D information should be attached to organization's strategy and everyday tasks in order to improve and develop the future operations. At the same time the R&D data and results should be disseminated quickly to other R&D actors of the field. Behind all this the R&D financing is decreasing and human resources are limited, which leads that important research areas are left out of the scope due to strong priorization of R&D projects. The quality of R&D weakens easily, as the amount of duties is high and personnel overloaded. All of these issues lead to questionize the true effectiveness of R&D operations.

Based on this research, R&D is as an important part of transport sector's operations, but the challenge is to manage both operative tasks and R&D tasks effectively. It is clear, that through R&D effectiveness evaluations transport sector is able to improve its accountability and visibility in the society. In public R&D organizations, effectiveness evaluations support the organization's development and provide evidence for policy makers and society on organization's performance. However, the implementation of effectiveness evaluation and connecting the evaluations to organization's daily operations was concidered challenging. As Paasio (2003) noted, effectiveness evaluation as part of everyday operations requires specific culture, and changing the culture is always challenging.

This research provided the central tools and concepts for evaluating societal impact of research in transport sector's R&D. The findings of this research support that the knowledge base of societal impact needs to be strengthened in sectoral level. The path to societal impact may vary in organizational level, but the term of societal impact should be defined in a common way. Although the long-term

impacts of R&D can be observed after several years, now it is the moment when the organizations create the qualifications to produce these impacts.

8.1 Limitations

According to Ruegg & Jordan (2007) case study method has some general limitations. First of all, case study method is generally considered less persuasive than comprehensive statistical approaches. Secondly generalization is more challenging from descriptive case study results. However, it is possible to draw common themes from case studies if the representation of a population is provided through case study and through supporting quantitative results.

This study focused on investigating societal effectiveness from R&D purchaser's and end-user's point of view. Therefore the aspects of R&D principal and provider was left out from the research scope. However, these actors are also centrally involved in public R&D.

In this study societal effectiveness was investigated through 4 R&D themes, which were pavements, environment, building information model and intelligent transport systems. The field of R&D is however much wider, and due to limited timeframe this research focused only to these themes.

8.2 Reliability, validity & ethics

Reliability evaluates whether the research is producing reliable information of reality. In qualitative research reliability indicates the reliability of processing and analysing the data. In this research several research methods were utilized simultaneously to ensure the reliability of the research. The theoretical contribution of this research can be considered reliable because it utilized current and reliable professional literature. The material that was utilized in the theoretical part of the research is also consistent with the examined phenomenon.

The interview situations seemed exciting for some respondents, which resulted in

short responses. This affected naturally to the research results. In spite of the tension, confidence between the researcher and interviewees was detectable, which improves the reliability of the research. The empirical part was based on personal opinions of 8 people and these opinions obviously affected to the research results. Therefore any generalizations can not be done based on this research. The end-users were chosen by the R&D experts and the recognizing of the end-users turned out to be challenging. The spectrum of the end-users is extremely wide and therefore the findings should be analyzed critically.

Validity indicates whether the research succeeded to measure what was the initial objective. By setting clear objectives, the research focus is on the most relevant issues, which contributes the validity of the research. This research is valid, because answers to all the research questions were detected and therefore the aim of the research has been reached. Also the theoretical part of the research can be considered as valid, as the theoretical and conceptual definitions are in harmony with each other. This means that the theoretical contribution and the data are logic and uncontroversial.

In qualitative research the ethical aspect must be considered. According to ethical aspect, the researcher is ethically in response for the research result to correspond to data utilized in the research. The researcher must appreciate the anonymity of the interviewees and preserve their identity during the whole research process. In this research the results are analyzed in a way, that the the respondent is not identifiable. Participation to this research was optional and permission for participating to the research was asked from all the respondents.

8.3 Further research

This study raises up several new research topics that should be investigated before a comprehensive picture of societal effectiveness can be formed. First of all, a wider-scale research following the Purchaser-Provider-model should be done. This research focused only investigating societal effectiveness from purchasers and end-users points of views, and to gain comprehensive picture

about the issue the principals and providers should be included to the research scope. In FTA's case this would mean organizing interviews in the Ministry and with R&D providers i.e consultants or universities. Also a wider scale case-study could be made with the private sector's companies, which are centrally involved in FTA's R&D operations for evaluating the societal effectiveness from wider aspect.

One further research topic could be investigating how FTA's R&D perform in terms of societal effectiveness when compared to its peers abroad. FTA has wide international R&D cooperation network and societal effectiveness has been already discussed within other countries as well. This network should be utilized also in developing societal effectiveness and together the R&D actors could find proper solutions and even common methods for evaluating R&D from societal effectiveness point of view.

REFERENCES

Amaratunga, D. & Baldry, D. (2002). *Moving from performance measurement to performance management*, Facilities, Vol. 20, 217 - 223.

Bornmann, L. (2012). *What Is Societal Impact of Research and How Can It Be Assessed? A Literature Survey.* Journal of the American society for information science and technology, 217–233.

Bozeman, B. & Kingsley, G. (1997). Journal of Technology Transfer Vol. 22 (2): 33-42.

Bozeman, B. & Melkers, J. (1993). *Evaluating R&D impacts: methods and practice.*

Bozeman, B. & Rogers, J. (2001). *Strategic management of government-sponsored R&D portfolios*. Environment and Planning C: Government and Policy, Vol.19, 413-442.

Bozeman, B., & Sarewitz, D. (2011). *Public value mapping and science policy evaluation*. Minerva, 49(1), 1–23.

Brennan, K & Levy, D. (2006). *Measuring social impact.* Available in: http://skollworldforum.org/2006/09/25/measuring-social-impact/

Bryman, A. (2008). *Qualitative Research in Organizations and Management,* 3, pp. 159-168.

Chiesa, V & Masella, C. (1996). *Searching for an effective measure of R&D performance,* Management Decision, Vol. 34 Iss 7 pp. 49 - 57

Dahler-Larsen, P. (2005). *Vaikuttavuuden arviointi.* STAKES - hyvät käytännöt Menetelmä käsikirja.

DeNisi, A.S. (2000). *Performance Appraisal and Performance Management: A Multilevel Analysis*. Multilevel Theory, Research and Methods in Organizations, pp. 121–156.

Donovan, C. (2008). *The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research*. New Directions for Evaluation, 47–60.

Figueras, J., Robinson, R. & Jakubowski, E. (2005). *Purchasing to improve health systems performance.* European Observatory on Health Systems and Policies Series. Open University Press, Berkshire. 298.

Gertler, P., Martinez, S., Premand, P., Rawlings, L. & Vermeesch,C. (2011). *Impact evaluation in practice.* The World Bank.

Godin, B. & Doré, C. (2005). *Measuring the impacts of science: Beyond the economic dimension.* Available on: http://www.csiic.ca/PDF/Godin_Dore_Impacts.pdf

Guthrie, J. & English, L. (1997). *Performance information and programme evaluation in the Australian public sector.* International Journal of Public Sector Management, Vol. 10 Iss 3 pp. 154 - 164.

Harding, A. (2014). What is the difference between an impact and an outcome? Impact is the longer term effect of an outcome. Available in: http://blogs.lse.ac.uk/impactofsocialsciences/2014/10/27/impact-vs-outcome-harding/

Hjelt, M., Luoma, P., Pesola, A., Saario, M., Kämäräinen, V., Maksimainen, A. & Jarkko, V. *Selvitys Tekesin toiminnan vaikutuksista yhteiskunnan ja ympäristön hyvinvointiin.* Available in: http://www.tekes.fi/globalassets/julkaisut/tekesin_toiminnan_vaikutukset.pdf

Huuhtanen, P., Järvenpää, P., Tuomivaara, S., Kähkönen, L. & Hannonen, H. (2009). *Kuntaorganisaatiot muutoksessa: Tilaaja-tuottaja -toimintatapojen käyttöönotto ja vaikutukset.* Työterveyslaitos. Työsuojelurahaston hanke nro 107093.

Humphries, A. & Wilding, R. (2004). *UK defence supply chain relationships: a study of sustained monopoly.* Management Decision 42(2): 259–276.

Hyytinen, K. & Konttinen, J. (2006). *Vaikuttavaa tutkimusta. Arviointikäytäntöjä julkisissa tutkimusorganisaatioissa: Puolustusvoimien Teknillinen Tutkimuslaitos PVT.* VTT Research Notes 2006.

Hyytinen, A. & Toivanen, O. (2010). *Tutkimuksen hyödyntäminen poliittisessa päätöksenteossa.* Kansantaloudellinen aikakauskirja 106.

Kahn, C. & McGaurty, S. (2009). *Performance Management at R&D organizations: Practices and Metrics from case examples.* Mitre Technical Report.

Kaplan, R. S. & Norton, D. P. (2003). *Strategy Maps: Converting Intangible Assets Into Tangible Outcomes*. Boston: Hardvard Business School Press.

Kellogg Foundation. (2004). *Logic Model Development Guide.*

Kettner, P. & Martin, L. L. (1998). *Accountability in Purchase of Service Contracting*. The privatization of human services: Policy and practice issues, 183-204.

Kuopila, A. *et al*. (2006). *Tilaaja-tuottaja-toimintatapa - Ideasta käytäntöön. Näin me sen teimme*. Suomen Kuntaliitto, Helsinki.

Liikennevirasto. (2013). *Liikenneviraston tutkimus ja kehittäminen: T&K suunnitelma 2013-2015.*

Loikkanen, T. & Kutinlahti, P. (2005). *Towards Systemic Future-oriented Innovation Policy Studies & Perspectives of Finnish Knowledge Based Economy*.

Lyall, C., Bruce, A., Firn, J., Firn, M. & Tait, J. (2003). *Assessing end-use relevance of public sector research organisations.* Research Policy 33 (2004) 73–87.

Lähteenmäki-Smith, K., Hyytinen, K., Kutinlahti, P. & Konttinen, J. (2006). *Research with an impact: Evaluation practises in public research organisations.* VTT tiedotteita 2336. Available in: http://www.vtt.fi/inf/pdf/tiedotteet/2006/T2336.pdf

Martin, B.R. (2007). *Assessing the impact of basic research on society and the economy*. Rethinking the impact of basic research on society and the economy (WF-EST International Conference)

Ministry of Transport and Communications. (2015). *Transport Policy.* Available in: *https://www.lvm.fi/web/en/transport_policy*

Mustajoki, A. (2005). *Tutkimuksen vaikuttavuus: Mitä se on ja voidaanko sitä mitata?* Article on Tieteessä tapahtuu.

Paasio, P. (2003). *Vaikuttavuuden arvioinnin rakenne ja mahdollisuus sosiaalialalla.* Stakes, Finsoc Työpapereita 3/2003. ISBN 951-33-1389-1.

Penn, The University of Pennsylvania. (2015). *Improving your social impact measurement practices: Three distinctions on the path to impact.* Jewish Federation General Assembly 2014.

Rantala, J. (2008). *Tilaaja–tuottaja-malli ja metsätalouden julkisrahoitteinen palvelutuotanto*. Metsätieteen aikakauskirja 2/2008: 97–115.

Rooney, P. (2011). *Evaluation Process.* The Leading Business Publication for Nonprofit Management. Available in: http://www.thenonprofittimes.com/news-articles/evaluation-process/

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation a systematic approach* (7th ed.). Thousand Oaks: Sage Publications

Rouse, M. (2014). *What critical success factors drive your business' success?* Project Management Strategies, CIO Guide.

Ruegg, R. & Jordan, G. (2007). *Overview of Evaluation Methods for R&D Programs.* US Department of Energy. Available in: https://www1.eere.energy.gov/analysis/pdfs/evaluation_methods_r_and_d.pdf

Savas, S. 2000. *Privatization and public-private partnerships*. Seven Bridges Press, New York. 368.

Steffens, P. & Matthews, J. H. (2004) *Network Governance of R&D: Purchaser-Provider as Partnerships in a Public Sector Agency*. British Academy of Management Conference.

Steinbach, R. (2013). *Policy decisions in the transport sector.* National Institute for Health Research.

Tassey, G. (2003). *Methods for Assessing the Economic Impacts of Government R&D*. Planning Report 03-1. National institute of Standards & Technology, NIST.

Taysir, E. & Taysir, N. (2012). *Measuring effectiveness in nonprofit organizations: An integration effect.* Journal of Transnational Management, 220-235.

Terry, F. (2000). *Transport: beyond predict and provide. What works? Evidence-based policy and practice in public services*. Policy Press.

Valkama, P. (2004). *Kilpailuneutraliteetin toteutuminen kuntapalveluiden näennäismarkkinoilla*. Acta Universitatis Tamperensis 1039. Tampereen yliopistopaino Oy, Tampere.

Walter, A. I., Helgenberger, S., Wiek, A. & Scholz, R. W. (2007). *Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method.* Evaluation and Program Planning 30 (2007) 325–338.

Weiss, C. (1995). *Evaluation*. Second Edition. Upper Saddle River, Prentice Hall.

Wright, K. E. (2003). *Problems? What problems? We have none at all. Qualitative data collection for impact assessment.* Journal of Microfinance 5, 115–38.

Yin, R. (2012). *Applications of case study research.* 3rd edition. SAGE Publications.

APPENDIX 1. Interview questions

A. R&D purchasers / R&D experts of FTA

Theme 1: R&D operations

1. How FTA has succeeded in the following activities? → How the success can be seen and how these activities could be developed?
    - Bringing out new research needs
    - Bringing out central R&D objectives
    - Bringing out central R&D impacts
    - Bringing out R&D results
    - Implementation of R&D results
2. Does FTA's R&D produce such information that the actors (companies, research institutes, decision-makers, consultants etc.) can utilize? → If no, how this issue could be improved?

Theme 2: Societal effectiveness of R&D

1. What does 'societal effectiveness of R&D' stand for according to your opinion?
2. On scale 1 to 10 how significant societal effectiveness evaluation is in relation to other FTA's evaluations?
3. What is the key starting point for purchasing R&D according to you?
4. How the aspect of FTA's R&D's end-user is taken into account in the R&D process?
5. Who or which party defines which impacts are to be aimed for in R&D?
6. Has R&D produced those impacts, which were originally seeked for and to what extent?
7. What kind of societal impacts R&D has already produced? Have these impacts been measured or evaluated, how?
8. Which issues are central in measuring societal effectiveness?
9. What kind of issues would you point out as the central challenges in measuring societal effectiveness?
10. How the FTA's R&D's societal effectiveness could be increased according to your opinion?
11. Do you have anything else to add related to FTA's R&D or societal effectiveness?

B. R&D end-users

Theme 1: R&D operations

1. Could you tell how you are involved in FTA's R&D activities?
2. How would you describe the FTA's communication related to R&D activities?
3. How would you describe FTA's R&D data's and results';
   a. Quality
   b. Utizibility
   c. Visibility
   d. Availability?
4. What are the key issues in utilizing and implementation of R&D results? What about challenges?
5. To what extent FTA's R&D has noticed yours or your organization's aspect and needs in R&D activities and projects?
6. On scale 1 to 10 how relevance FTA's R&D is as social influence?

Theme 2: Societal effectiveness of R&D

7. What does 'societal effectiveness of R&D' stand for according to your opinion?
8. What kind of societal impacts FTA's R&D have produced according to your opinion?
9. Has FTA's R&D produced those impacts, which were seeked for in the beginning and to what extent?
10. In relation to FTA's R&D activities, which issues require improvement according to your opinion?
11. Do you have anything else to add related to FTA's R&D or societal effectiveness?