



LUND UNIVERSITY

Evaluating for a transition - Advancing evaluation of research and policy for a more sustainable energy system and society

Sandin, Sofie

2021

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Sandin, S. (2021). *Evaluating for a transition - Advancing evaluation of research and policy for a more sustainable energy system and society*. [Doctoral Thesis (compilation), The International Institute for Industrial Environmental Economics]. IIIEE, Lund University.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00



Evaluating for a transition

Advancing evaluation of research and policy for a more sustainable energy system and society

SOFIE SANDIN | IIIIEE | LUND UNIVERSITY





How can evaluation of policy instruments and research initiatives support large-scale transitions towards a more sustainable energy system and society? This question is addressed in this thesis, which focuses on how to advance evaluation practices to provide the essential learning needed to address complex sustainability challenges.

Evaluation of policy and research is key for understanding how different initiatives work and their potential contribution. However, evaluation is often focusing on individual programmes and does not provide insights about transformative contributions. Thus, there is a need to consider how evaluation practices need to be complemented in order to capture how different policy incentives - alone and together – are supporting a transition. This thesis proposes frameworks for how to design such transformative evaluations, by drawing on interdisciplinary foundations and combining key aspects to consider.

The research builds on the case of current evaluation practices in Sweden, focusing particularly on initiatives for energy efficiency in buildings. By exploring how evaluations are conducted and used, the thesis illuminates areas for further advancement, as well as probes the potential for a transformative evaluation approach to be adopted among Swedish state agencies.

EVALUATING FOR A TRANSITION
Advancing evaluation of research and policy for a
more sustainable energy system and society

Evaluating for a transition

Advancing evaluation of research and policy for a
more sustainable energy system and society

Sofie Sandin



LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Engineering, Lund University, Sweden.
To be defended at the International Institute for Industrial Environmental
Economics, Aula and online, 21 May 2021, 10:00.

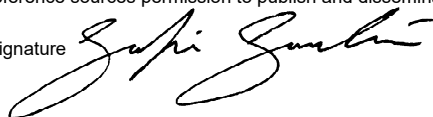
Faculty opponent

Erik Arnold, Adjunkt professor KTH,
Honorary professor University of Manchester

Organization LUND UNIVERSITY International Institute for Industrial Environmental Economics (IIIEE)		Document name DOCTORAL DISSERTATION
Author Sofie Sandin		Date of issue 2021-04-30
		Sponsoring organization
Title and subtitle Evaluating for a transition - Advancing evaluation of research and policy for a more sustainable energy system and society		
Abstract In response to global sustainability challenges, there are multiple public research and policy incentives being implemented to support a more sustainable energy system. These are aimed at advancing technology and innovation as well as guiding citizens to change behaviour. But are those incentives accurately leading to a transition towards a more sustainable energy system and society? Evaluation of these incentives is key to addressing such a question, and such evaluations need to be able to provide knowledge that is oriented around transformative aspects and effects on system levels, and that is grounded in rigorous assessment approaches. The objective of this thesis is thus to advance knowledge on how the evaluation of research and policy incentives can be further developed to support a transition. The research presented will focus on Swedish evaluation practices of research and policy incentives aimed at energy efficiency in buildings. In order to provide foundations for how <i>transformative evaluation</i> can be designed and conducted, this thesis combines insights from transition research, evaluation theory, sociology of science, and policy analysis. In essence, transformative evaluation, as proposed in this thesis, will consist of two main tracks: it is transformative in that it departs from transition research; and it is rigorous and knowledge-based in that it supports assessments that build on and combine relevant evaluation-oriented insights. As such, transformative evaluation seeks to provide a broader approach that addresses the complexities of a systems-level perspective, and that contrasts and complements more traditional evaluation approaches that have a predominant orientation around the programme level. The review of the Swedish evaluation practices shows that evaluations are largely not currently used for, or focusing on, assessing transformative contributions. Key suggestions for advancing evaluation practices to support transformative insights include articulating the role of the evaluator in relation to societal goals and <i>visions</i> , heeding <i>experimentation</i> , and capitalizing on the <i>learning</i> that can be drawn from it. Suggestions for advancing evaluations towards rigorous and knowledge-based assessments include an increased application of triangulation, counterfactual analyses, and a wider variety of criteria than what is currently seen, including e.g. relevance, legitimacy, acceptability, adequacy, equity and responsiveness. A conclusion is that moving towards transformative evaluation can support both an <i>alignment</i> of research and policy instrument evaluations, as well as an understanding of <i>cascading</i> effects from individual research and policy incentives. In terms of adopting a transformative evaluation approach among Swedish state agencies, both benefits and challenges are noted by state representatives. Benefits include increased collaboration and enhanced evaluation use, while challenges relate to institutional and structural barriers, and issues surrounding agency and mandates to govern change processes.		
Key words research evaluation, policy evaluation, transition, energy efficiency, buildings		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		Language English
ISSN 1402-3016 IIIEE dissertations		ISBN 978-91-87357-66-4 (print) ISBN 978-91-87357-67-1 (pdf)
Recipient's notes	Number of pages 214	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature



Date 2021-04-08

Evaluating for a transition

Advancing evaluation of research and policy for a
more sustainable energy system and society

Sofie Sandin



LUND
UNIVERSITY

Cover picture by Kevin Amrulloh.

To me, the Rubik's cube represents the versatility of approaching evaluation. Each side has its own colour, its own perspective and story to tell, but solving one side does not give a complete result. And, as all sides stand connected to each other, trying to solve one side will inevitably affect all other sides. For evaluation, this illustrates the complexity of evaluands, and points to that different approaches will show various results. In the quest to see the whole picture, all sides need to be considered.

The difference between policy evaluation and the Rubik's cube is that while the cube has an answer, an algorithm for solving it, evaluation and knowledge-creation should continue to turn the sides to find new angles, connections, and approaches.

Copyright pp 1-110 CC BY-SA 4.0 Sofie Sandin
Paper 1 © Sofie Sandin & Mats Benner (manuscript)
Paper 2 © Creative Commons CC BY
Paper 3 © 2020 Elsevier Ltd. All rights reserved.
Paper 4 © Attribution 4.0 International (CC BY 4.0)

International Institute for Industrial Environmental Economics (IIIEE)
at Lund University

ISBN 978-91-87357-66-4 (print)
ISBN 978-91-87357-67-1 (pdf)
ISSN 1402-3016

Printed in Sweden by Media-Tryck, Lund University, Lund 2021



Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN 

*To Bengt & Sonja,
to Michel & Ruben,
who inspire and encourage me*

Table of Contents

Acknowledgements	11
Abstract.....	13
Popular Science Summary.....	15
List of Papers	17
Preface	19
1. Introduction	21
1.1 An outlook on current evaluation practice	22
1.2 Towards transformative evaluation	24
1.3 Research objective	26
1.4 Research overview	28
1.5 Research scope.....	29
1.5.1 Sectoral and geographical scope.....	29
1.5.2 Theoretical scope	30
1.6 Thesis structure	30
2. Theoretical background and foundations of analytical frameworks	33
2.1 Transition research.....	33
2.1.1 Key concepts for transformative evaluation.....	34
2.2 Evaluation theory	36
2.2.1 Methods.....	38
2.2.2 Valuing	40
2.2.3 Use	42
2.3 Sociology of science.....	43
2.3.1 Evaluation of research	44
2.4 Policy analysis	47
2.5 Analytical frameworks	50
2.5.1 Framework for Paper I	50
2.5.2 Framework for Paper II.....	51
2.5.3 Framework for Paper IV.....	51
3. Research methodology.....	53
3.1 Research positioning and design	53
3.1.1 An interdisciplinary approach.....	53
3.1.2 Research positioning	54

3.1.3	Research design	55
3.2	Methods for data collection and analysis.....	57
3.2.1	Document studies and systematic reviews.....	57
3.2.2	Focus group sessions	58
3.2.3	Structured interview questionnaire	60
3.2.4	Ethical considerations	61
4.	Key findings organized by papers.....	63
4.1	Paper I	63
4.1.1	Objective and framework	63
4.1.2	Results	64
4.1.3	Conclusions	65
4.2	Paper II	65
4.2.1	Objective and framework	65
4.2.2	Results	66
4.2.3	Conclusions	67
4.3	Paper III.....	67
4.3.1	Objective and framework	67
4.3.2	Results	68
4.3.3	Conclusions	69
4.4	Paper IV.....	69
4.4.1	Objective and framework	69
4.4.2	Results	70
4.4.3	Conclusions	70
5.	Concluding discussion.....	71
5.1	Empirical contributions.....	72
5.1.1	Research question 1.....	72
5.1.2	Research question 2.....	75
5.1.3	Research question 3.....	78
5.2	Theoretical contributions	79
5.3	Implications for evaluation practitioners.....	81
5.4	Reflection on research approach and methods	83
5.5	Suggestions for future research	85
6.	Conclusions.....	89
	References.....	91
	Appendix A – D.....	103

Acknowledgements

Being part of this interdisciplinary research project and institute has truly been a journey. Since I embarked on it in February 2016, the road has been leading me to new places and to new insights. Now, as I reach the end of this journey, I look back at a rewarding, challenging, and enriching experience. What stands out the most is the people I have met along the way. I have had the pleasure to meet so many inspiring and committed people who have shared their knowledge and experiences: during the project meetings, the coffee breaks in the IIIIEE kitchens, and at the international conferences and PhD courses throughout Europe.

I want to start by thanking my project colleagues in the Transition Governance project, Lena Neij, Mats Benner, Maria Johansson, and Per Mickwitz, for giving me the opportunity and privilege to join the team, for sharing your knowledge, and for always providing an open and creative atmosphere in the research project. There are some special thanks to my three supervisors. Lena, thank you for all your support and feedback, for leading me through this journey. I have always trusted that you knew the way even during the times when I felt lost. Mats, thank you for guiding me into the fields of sociology of science, for the discussions about research and society, and for the encouragement. Maria, thank you for supporting me with your knowledge on research methodologies, and for sharing the perspectives of an environmental psychologist.

I also want to thank the representatives at the Swedish state agencies who shared their time and viewpoints in the workshops held in 2019, and to those who shared their knowledge and insights in a questionnaire. And of course, a big thank you to the Swedish Energy Agency in particular, for funding the research project Transition Governance and making this research possible. On his note, I also wish to acknowledge the financial contribution from the Foundation for the International Institute for Industrial Environmental Economics (IIIIEE), and the Foundation in memory of Lars Inge Grundberg for supporting the finalization and the printing of this PhD thesis.

Next, I want to express my gratitude for having the opportunity to be part of the interdisciplinary and diverse IIIIEE. To all of my IIIIEE colleagues: you have made this journey truly enjoyable! Thanks for the chats, the support, the fikas, the walks, the seminars. Special thanks to the amazing PhD group and the postdocs, both you who came before me and showed me the path, and those of you who came after me, who brought new ideas and enthusiasm. You have become my work family and you have

enriched my journey. To Jessika, thanks for the discussions about policy and evaluation and life; Roland, for keeping me connected to the municipal world; Björn, for the support and encouragement, to Lucie for many good games on the badminton court, and Heather for the always-enjoyable chats about music. And, to Steven, thanks for being my companion in all the things we have taken on together, for inspiring me, for challenging me to do better, for the support, and for all the fun.

Last but not least, I want to thank my family and friends who have been there along the way. You all know who you are, and I am privileged to have you in my life. Thank you for the game nights, the dinners, and the laughs. And to Michel, for sharing the life journey with me.

Abstract

In response to global sustainability challenges, there are multiple public research and policy incentives being implemented to support a more sustainable energy system. These are aimed at advancing technology and innovation as well as guiding citizens to change behaviour. But are those incentives accurately leading to a transition towards a more sustainable energy system and society?

Evaluation of these incentives is key to addressing such a question, and such evaluations need to be able to provide knowledge that is oriented around transformative aspects and effects on system levels, and that is grounded in rigorous assessment approaches. The objective of this thesis is thus to advance knowledge on how the evaluation of research and policy incentives can be further developed to support a transition. The research presented will focus on Swedish evaluation practices of research and policy incentives aimed at energy efficiency in buildings.

In order to provide foundations for how *transformative evaluation* can be designed and conducted, this thesis combines insights from transition research, evaluation theory, sociology of science, and policy analysis. In essence, transformative evaluation, as proposed in this thesis, will consist of two main tracks: it is *transformative* in that it departs from transition research; and it is rigorous and *knowledge-based* in that it supports assessments that build on and combine relevant evaluation-oriented insights. As such, transformative evaluation seeks to provide a broader approach that addresses the complexities of a systems-level perspective, and that contrasts and complements more traditional evaluation approaches that have a predominant orientation around the programme level.

The review of the Swedish evaluation practices shows that evaluations are largely not currently used for, or focusing on, assessing transformative contributions. Key suggestions for advancing evaluation practices to support transformative insights include articulating the role of the evaluand in relation to societal goals and *visions*, heeding *experimentation*, and capitalizing on the *learning* that can be drawn from it. Suggestions for advancing evaluations towards rigorous and knowledge-based assessments include an increased application of triangulation, counterfactual analyses, and a wider variety of criteria than what is currently seen, including e.g. relevance, legitimacy, acceptability, adequacy, equity and responsiveness. A conclusion is that moving towards transformative evaluation can support both an *alignment* of research

and policy instrument evaluations, as well as an understanding of *cascading* effects from individual research and policy incentives.

In terms of adopting a transformative evaluation approach among Swedish state agencies, both benefits and challenges are noted by state representatives. Benefits include increased collaboration and enhanced evaluation use, while challenges relate to institutional and structural barriers, and issues surrounding agency and mandates to govern change processes.

Popular Science Summary

This thesis sets out to advance knowledge on how a transition towards a more sustainable energy system and society can be supported by insightful and deliberate evaluation of research and policy incentives.

It does so by promoting *transformative evaluation*, which is an interdisciplinary foundation for evaluation that synthesizes key insights from transition research, evaluation theory, sociology of science, and policy analysis. Transformative evaluation, as proposed in this thesis, seeks to support rigorous knowledge creation that can provide insights into how transformative contributions of an *evaluand* - the object that is being evaluated - relate to and influence large-scale changes in society. Thus, it proposes a more holistic approach to evaluation than what is seen in traditional evaluation approaches, which are mainly oriented around programme processes and goal attainment.

The research is oriented around the case of evaluation of research and policy incentives for energy efficiency in buildings in Sweden, which is used for identifying how current evaluations are conducted in order to outline how these can be further advanced and complemented to move towards a transformative evaluation approach. Data collection and analyses build on reviews of existing evaluation reports of both research and policy incentives, as well as information gained from representatives working with evaluation at Swedish state agencies.

A key finding is that there are many evaluations performed of incentives aimed at energy efficiency in buildings, and these are commissioned and conducted by a limited number of actors. This provides a good foundation to move towards a transformative evaluation approach. Suggestions on how to complement current practices include a need to concretely articulate the energy vision that research and policy incentives seek to support, and to assess how the evaluand supports this vision. The research also suggests that capturing experimentation and learning generated by the evaluand can be further complemented, since testing and piloting of new solutions and practices, and learning from these activities, are required to bring about changes. Evaluations can also expand the involvement of key actors, and adopt a systems perspective to a larger degree. Furthermore, in order to make rigorous assessments of the transformative contributions, one key suggestion is to apply multiple methods and criteria. This in order to facilitate validation of findings through triangulation, and to provide broader

assessments that go beyond the prevalent use of programme-oriented criteria such as effectiveness and impact. The application of counterfactual assessments can also be extended to ensure that effects and impacts are indeed attributable to the evaluand and do not stem from other sources.

Furthermore, evaluations should seek to take a more holistic perspective, where the transformative contribution or potential of an incentive is put in relation to insights generated in other evaluations. This can also support an enhanced *alignment* of research and policy evaluations, and a more comprehensive understanding of *cascading* effects, which illuminate how outcomes and effects stemming from an individual research or policy incentive in turn support or spur transformative effects from other incentives.

Another key for successful evaluations is that they are *used*. This research shows that the use of evaluations, as perceived by informants at Swedish state agencies, is mainly proactive, as use for improvement and for knowledge creation is frequently recognized, rather than use for promoting agendas or simply meeting evaluation requirements. The processes that underpin evaluation use among Swedish state agencies are reported to build on involvement of various actors and channels, which can further support an extended use and sharing of evaluations. In all, this provides good foundations for adopting a transformative evaluation approach, but there are also challenges noted that relate to overcoming structural and institutional boundaries associated with evaluation practices among and between Swedish state agencies.

Moving towards transformative evaluation will require a revision of current evaluation practices, mainly concerning how state agencies with mandates to govern transformative change work with acquiring and sharing knowledge. In order to promote deliberate and well-designed research and policy incentives that can address the need for supporting a transition towards a more sustainable energy system and society, evaluations must be supported to purposefully combine transformative insights with rigorous evaluation approaches.

List of Papers

Paper I

Sandin, S. & Benner, M. (2021). Research evaluations for an energy transition? Insights from a review of Swedish research evaluation reports. (*submitted, minor revision*)

SS and MB designed the study. SS collected evaluations and performed a review of 20 existing evaluations. MB designed the outline of the analytical framework, which SS contributed to. Both authors wrote the manuscript.

Paper II

Sandin, S., Neij, L. & Mickwitz, P. (2019). Transition governance for energy efficiency – insights from a systematic review of Swedish policy evaluation practices. *Energy, Sustainability and Society* 9/1

SS and LN designed the study. SS collected evaluations and performed a review and assessment of 33 existing evaluations, following an analytical framework designed by PM, LN and SS. SS was a major contributor in writing the manuscript; LN and PM supported the text writing.

Paper III

Neij, L., Sandin, S., Benner, M., Johansson, M., & Mickwitz, P. (2021). Bolstering a transition for a more sustainable energy system: A transformative approach to evaluations of energy efficiency in buildings. *Energy Research & Social Science* – 72, 101864

LN designed the study in cooperation with SS, MB, MJ and PM. Evaluations were collected and analysed by SS, MJ and LN. LN and SS were the major contributors in writing the manuscript supported by MB, MJ and PM.

Paper IV

Sandin, S. (2020). Making Use of Evaluations to Support a Transition towards a More Sustainable Energy System and Society—An Assessment of Current and Potential Use among Swedish State Agencies. *Sustainability* 2020, 12, 8241.

Other publications

Mickwitz, P., Neij, L., Johansson, M., Benner, M. & Sandin, S. (2021). A theory-based approach to evaluations intended to inform transitions toward sustainability. *Evaluation*. DOI: 10.1177/1356389021997855

Neij, L., Sandin, S., Benner, M., Johansson, M. & Mickwitz P. (2019). Vägledning för utvärdering av transformativ omställning – med fokus på energisystemet [A guide for evaluation of transformative change – with a focus on the energy system]. Media-Tryck Lunds Universitet, Lund. ISBN: 978-91-87357-44-2

Sandin, S., Neij, L. & Benner, M. (2019). Evaluating for a transition – a systematic approach to aligning research and policy evaluation. *Proceedings of the European Forum for Studies of Policies for Research and Innovation (EU-SPRI) Conference*, Rome, Italy, 2019 (Short paper, peer reviewed)

Sandin, S., Neij, L. & Mickwitz, P. (2017). Review of evaluations of policy instruments for energy efficiency in buildings in Sweden. *Proceedings of ECEEE summer study*, Hyères, France, 2017 (Full paper, peer reviewed)

Preface

This research has been carried out at the International Institute for Industrial Environmental Economics at Lund University, an institute oriented around interdisciplinary research for advancing sustainable solutions moving towards carbon-neutral and resource-efficient societies. The core activities of the Institute are characterized by collaboration with other actors in society and in academia to maintain societal relevance and provision of timely research output.

The PhD project was performed as part of the Transition Governance project (2016-2019), an interdisciplinary research project funded by the Swedish Energy Agency (grant number 39938-1). The aim of this project was to advance evaluation knowledge for supporting a transition towards a more sustainable society, with a particular focus on research and policy instruments for energy efficiency in buildings. The project members consisted of professors Lena Neij, Mats Benner, Maria Johansson and Per Mickwitz, and PhD candidate Sofie Sandin. Together, the members represented different disciplinary scientific backgrounds: energy system studies, policy studies, sociology, environmental psychology, and environmental science, which truly supported the interdisciplinary approach of the research.

The project also had a reference group consisting of representatives from Swedish state agencies: Peter Stern (Energy Agency), Lisa Eriksson (Environmental Protection Agency), Tobias Persson (Growth Analysis), and Professor Hanne Foss Hansen (University of Copenhagen).

1. Introduction

The world is currently facing sustainability challenges that call for action and for large-scale changes (European Environment Agency, 2019a, 2019b; IPCC, 2018) affecting entire socio-technical systems (Kemp, 1994; Markard et al., 2012). Such large-scale changes are known as *transitions* – deep and foundational alterations of the current system configurations that move societies from one equilibrium to another, shaped by new norms, practices and technology (Geels, 2002; Kemp, 1994; Markard et al., 2012). The call for transitions thus clearly encompasses the entire socio-technical system, and there are many areas to focus on for supporting just and equitable changes. A recent and prominent example of this is the 2030 Agenda for Sustainable Development and its 17 Sustainable Development Goals (United Nations General Assembly, 2015).

One key area for addressing these sustainability challenges is the energy sector, which has received notable attention due to its high potential for improving both environmental and social conditions, e.g. through reducing use of resources, mitigating negative climate impacts, increasing energy security, and alleviating social inequities and improving people’s health and well-being (Bruckner et al., 2014; Buonocore et al., 2016; GEA, 2012). Pathways for supporting a more sustainable energy system and society include advancements in various sectors such as industry, transport and building, by e.g. converting to renewable energy, and notably by moving towards more energy-efficient practices and technologies. This thesis takes a particular focus on energy efficiency in buildings (see Box 1).

To accelerate the required energy transition, there is a need for strong and deliberate research and policy incentives that support innovation and governance of societies towards more sustainable practices (IPCC, 2018; IRENA, 2019). Thus, globally, new policies and innovations are called for to guide the development and pace of the transition at large, and existing policies need to be revised and become more ambitious if energy targets are to be met (Global Alliance for Buildings and Construction et al., 2019; International Energy Agency, 2019, 2017). Seen from a systems perspective, this variety of research and policy incentives may support a transition by promoting *transformations* in certain parts of the system. This thesis sees transformations as the local changes or effects that are yielded by a research or policy incentive, which then in combination with, or support from, other incentives may spur system-wide changes: a transition.

Thereby, in order to make deliberate research and policy decisions that lead towards a more sustainable energy system, this thesis builds on the argument that there is a need to consider the roles and contributions of such different incentives from a transformative perspective, and to explore their potential to support a transition. This is where evaluation of such different incentives becomes crucial, and where *transformative evaluation* comes into the picture.

Box 1. A focus on energy efficiency in buildings

Energy efficiency is referred to as the ‘*first fuel*’ of all energy transitions’ (International Energy Agency, 2019), and energy efficiency in buildings in particular holds great potential to be harnessed in terms of adoption of more efficient technologies, policies, and changes in energy behaviours (International Energy Agency, 2018, 2017; IPCC, 2018). The building sector is singled out as a key area for reaching energy and climate targets, as it accounts for 40% of the energy consumption, and 36% of the greenhouse gas emissions within the European Union alone, and with similar figures also seen globally (European Commission, 2020; Global Alliance for Buildings and Construction et al., 2019).

Moreover, the building sector is characterized by long time spans, as measures and technologies adopted in buildings become fixed for a long time, which further adds to the urgency of promoting energy-efficient solutions. Thus, there is a call for prompt actions to reduce the risk of inefficient solutions becoming locked in both new constructions and in refurbishments of existing buildings (International Energy Agency, 2017; IPCC, 2018).

1.1 An outlook on current evaluation practice

In order to make informed policy decisions, there is a need for knowledge and learning about how policies have worked and what they have contributed, as well as on how they will perform under current and changing conditions. In essence, to acquire an understanding about what works and what does not under the auspices of a transition. This does, however, call for evaluations that can account for system-wide effects, and that can address the complexities of transitions.

How then, do present evaluations fare in relation to providing transformative insights? On an international level, the evaluation of both research and policy incentives is recognized as a means to generate informed and timely insights about processes and impacts from different implementations (IPCC, 2014). The importance of evaluating environmental policy instruments in particular is emphasized in the EU’s 7th

Environment Action Programme, and more detailed guidelines for how to perform evaluations are outlined in documents issued by e.g. the European Commission (2017) and by the European Environment Agency (2016). However, it needs to be noted that there is no 'one' evaluation practice, but rather, the approaches to evaluation and the institutional structures and discourses surrounding it vary between countries (Furubo et al., 2002; Jacob et al., 2015; Stern, 2009).

Broadly, the evaluation of research has become an integrated part of the academic world and research policy, where it supports decision-making concerning e.g. which programmes are to receive funding, but also assists in determining societal impact, efficiency, quality, and research excellence (Geuna and Martin, 2003; Godin, 2007; Hellström, 2011). Likewise, the evaluation of policy instruments is an established part of policy processes, where it serves to provide ex-ante or ex-post assessments concerning e.g. impact and effectiveness of programmes, and to support decisions on which policy instruments to select, continue, and terminate (Dunn, 2018; Parsons, 1995).

While there are international examples of elaborate and innovative evaluation structures and practices, for example in the Netherlands with approaches oriented around assessing programmes in relation to societal goals (Knaap, 2000; Leeuw, 2009), in the UK with efforts to address the need for cross-sectoral evaluation approaches (CECAN, 2018), as well as a report commissioned by a Swedish state agency to structure evaluation of research and innovation in relation to a transition (Arnold et al., 2018), it should be noted that evaluation frameworks tend to be oriented around traditional evaluation practices. These commonly focus on the programme level and on improvement of programme design, in terms of attaining programme goals and objectives, and with few or no outlooks to systemic effects (Arnold, 2004; Cunningham et al., 2016; Sanderson, 2000). This fragmentation of evaluation is seen in many sectors and locations: in European countries (Huitema et al., 2011; Sullivan, 2011) and beyond (e.g. the USA and Canada) (Chouinard, 2013; Christie and Fleischer, 2010), where the focus on the programme level is mirrored in the mandates and expectations of the evaluations (Chouinard, 2013; Hildén, 2011). Other reasons for the fragmentation of evaluations stem from changing or conflicting political contexts and interests; from boundaries and barriers between different institutional bodies where evaluations are performed; and from tensions between expectations on either structural (aggregate policy-level insights) or operational (programme-level insights) areas of focus in evaluations (Hildén, 2011; Hildén et al., 2014; Pollitt, 1993; Stame, 2008; Stern, 2009; Sullivan, 2011).

Nevertheless, research has provided suggestions for evaluation practices seeking to uncover aspects surrounding transformations. Examples from evaluation theory include theory-based approaches and intervention theories that can, if applied as such, be seen as methods for providing transformative insights with regard to how a programme yields intended (and unintended) outcomes and effects towards fulfilling expected goals

(Chen, 1990; Vedung, 1997). Examples from the energy sector in particular include methods for evaluation of programmes for market transformation in relation to introducing and adopting energy-efficient products and services (Neij, 2001; Vine et al., 2001), and theory-based approaches for understanding how programmes may support the emergence of such transformative effects (Blumstein et al., 2000). Altogether, the efforts made for evaluating research and policy incentives today are in many respects valuable and recognized, but if evaluations are to be made more ready to inform about a transition, there is arguably a need to complement the evaluation approaches in order to equip them to gear focus towards transformative change. Moreover, rendering evaluations able to support a transition also entails expanding the use of evaluations beyond use for programme accountability and improvement, towards use practices that build on knowledge exchange between sectors, actors, and institutions in order to support a more systemic overview of the contributions to energy efficiency from various research and policy incentives.

1.2 Towards transformative evaluation

The next question is then: what insights can be gained from research on how to further advance traditional evaluation practices to support transformative insights?

In transition research, a transition is described as a system-wide reconfiguration, influencing multiple actors and institutions, and building on multiple transformational processes, factors and technologies (Geels, 2002). Moreover, transitions are often understood or exemplified from the perspectives of certain contexts or breakthroughs (e.g. the electric light or the steam ship), yet each such transformation stands connected to *other* societal transformations and sectors as well, where it affects, or is affected by, other transformative efforts (Geels, 2002). This paints the picture of a system in which different research and policy incentives are parts of movements of a transition, but their contribution on a larger, societal scale may be seen only from a systems perspective. Clearly, this makes evaluation of transitions inherently complex, and especially so on a granular level where the transformative contributions and roles of individual implementations are sought to be uncovered. Here, *transition research* provides valuable conceptualizations and understandings of how transformative and transitional processes come about, and of how different aspects may factor in and affect the transition pathway (Geels and Schot, 2007; Kemp et al., 1998; Schot et al., 1994). However, while meeting the need to understand the processes of transitions, transition research may not be adequate for providing tools for evaluating how different research and policy incentives work in relation to a transition.

Thus, in order to render evaluations able to inform about the transformative potential, role, or contribution of an *evaluand* - the object that is being evaluated - in a complex system, this thesis builds on the argument that there is a need to consider how to couple

transition research with relevant evaluation literature for additional guidance on how to perform rigorous assessments of both research and policy incentives: to support *transformative evaluation*.

The evaluation practices of today hold well-established expectations surrounding evaluations, as evidenced by the frequent conducting of evaluations and their established roles as knowledge providers in decision-making contexts. Arguably, these practices can advantageously be extended to provide transformative insights. A challenge that arises when coupling the understanding of a transition with evaluation practices is, however, what the perspectives taken in the evaluation should enfold in order to provide relevant and valuable knowledge concerning different aspects of a system transition. Thus, if evaluations are to become more encompassing and able to capture transformative efforts in a complex system, there is a need to reconsider the epistemological foundations of these evaluations, and allow them to take a broader focus that is rooted in different and relevant theoretical underpinnings. Based on this reasoning, this thesis argues that an essential pathway to move towards transformative evaluation practices is to combine transition research with several relevant evaluation-oriented theoretical foundations, in order to accommodate a broad and encompassing systems approach.

Transformative evaluation, as suggested in this research, seeks to further advance and complement evaluation practices with insights from different theoretical fields, in interdisciplinary approaches. It does so by building on two main tracks: firstly, it is *transformative*, meaning that it builds on insights from relevant transition research theories, and secondly, it is *knowledge-based*, meaning that it is grounded in relevant evaluation-oriented theories and methods that can provide rigorous assessments of transformative contributions. Thus, the suggested approach departs from insights from transition research about the processes and drivers of a transition, which are coupled with insights from other disciplines that can support a broad approach to generating knowledge about how research and policy incentives lead to transformative change. There are many disciplines to consider for supporting rigorous and knowledge-based transformative evaluation, but in this thesis, evaluation theory, sociology of science, and policy analysis are applied to provide the theoretical foundations for evaluating transformative efforts (see Figure 1).

The rationale for this selection is that these are fields that together can provide a more holistic approach to evaluation of both research and policy instruments, in a complex field like energy efficiency in buildings. Starting with evaluation theory, this is a field that hosts rigorous knowledge about approaches and methods for conducting evaluations of research and policy (see, for example, Crabbé and Leroy, 2008; Shadish et al., 1991). Yet, since evaluation theory has traditionally not been aimed at providing tools for assessing how transformative outcomes of different incentives may lead to transitions, it is argued to be less ready to address the complexities and changing

contexts pertaining to a sustainability transition (Julnes, 2019a). While transition research and evaluation theory together are arguably well equipped to complement each other to identify and support core ideas for transformative evaluation, this thesis also argues that deeper and specific insights are needed into the features of energy-related research and policy incentives in order to further assess their roles in transformative changes. Thus, in order to provide deeper insights of the processes and factors that shape research initiatives in particular, sociology of science is applied; and similarly, in order to understand processes underpinning policy instruments, policy analysis is applied in order to further outline essential traits of transformative and knowledge-based evaluations.

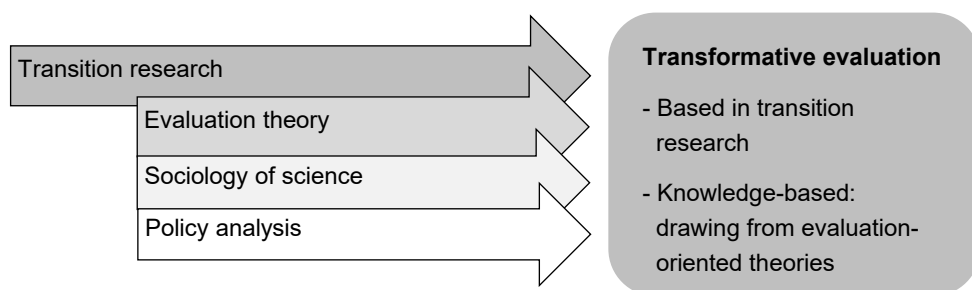


Figure 1. An overview of the different theoretical foundations proposed for transformative evaluation. This research departs from transition research, and turns to evaluation theory, sociology of science, and policy analysis for complementary insights on how to generate knowledge-based learning about how research and policy incentives perform from a transformative perspective.

1.3 Research objective

The objective of this thesis is to advance knowledge on how evaluation of research and policy incentives can support a transition towards a more sustainable energy system and society. This is done by looking into how theory can support transformative evaluation, and how evaluation practices can be advanced and complemented to assess transformative efforts.

The case used for operationalizing this objective is Swedish evaluation of research and policy incentives for energy efficiency in buildings, which was chosen based on the relevance of energy efficiency for addressing the sustainability challenges of today. More specifically, the building sector is an interesting focus area due to its inherent complexities involving multiple building types, usages, and actors (Cantin et al., 2012; World Business Council for Sustainable Development, 2009). There are multiple barriers located in various parts of the system which hinder the potential of energy efficiency in buildings (Tuominen et al., 2012; Weber, 1997), and thus, there is

arguably great potential for transformative evaluation approaches to provide valuable insights about how research and policy incentives together may support transformative efforts for energy efficiency in a complex system such as buildings.

This objective seeks to support two main contributions: a theoretical contribution and an empirical contribution.

The *theoretical contribution* intends to support a broader approach to evaluation by putting forward theoretical frameworks that outline how different disciplines can support knowledge-based evaluation of research and policy incentives, and render these evaluations more informing about transformative efforts. Thus, this contribution is grounded in the interdisciplinary approach that the suggested transformative evaluation approach builds on. Seeking to advance transformative evaluation of both research and policy incentives, insights are drawn from four disciplines that are identified as essential for supporting transformative evaluation: transition research, evaluation theory, sociology of science, and policy analysis. By identifying and combining key aspects from these disciplines, the theoretical contribution builds on the two suggested main tracks of transformative evaluation: (I) presenting a foundation in transition research, and (II) supporting knowledge-based evaluation through key aspects from relevant disciplines.

The *empirical contribution* seeks to provide suggestions on how to further advance and complement current Swedish evaluation practices. This is done in part by providing insights about how evaluations today are designed and conducted, and then, building on the suggested transformative evaluation approach, by outlining key areas where concrete suggestions can be made for how to make evaluation of research and policy incentives support a transition towards a more sustainable energy system and society. The empirical contribution also intends to shed light on the current use practices of evaluations among Swedish state agencies, and to identify and assess the perceived benefits and challenges related to adopting a transformative evaluation approach.

As such, the theoretical contribution supports the empirical contribution in three aspects. Firstly, by providing theoretical outlines of how to design transformative evaluation; secondly, by providing frameworks to generate insights about current Swedish evaluation practices and how they can be further advanced and complemented; and lastly, by providing insights about how to support the use of transformative evaluation.

The research objective crystallizes into three main research questions, which are geared towards the empirical contribution:

RQ1: Are current evaluation practices for research and policy instruments, within the field of energy efficiency in buildings, allowing a capturing of transformative contributions of the evaluand?

RQ2: How do the evaluation practices for research and policy instruments need to be complemented to be able to support a transition towards a sustainable energy system and society?

RQ3: What are the current processes for use of evaluations among Swedish authorities; and what are the major issues of concern among Swedish state agencies for adopting a transformative evaluation approach?

1.4 Research overview

In addressing the research objective and questions, this thesis is composed of four papers. Figure 2 illustrates how these papers are connected and how they relate to the posed research questions. Papers I and II look separately at the current Swedish evaluation practices for research programmes/institutions, and policy instruments for energy efficiency in buildings, respectively, and elaborate on how the current practices may be further complemented to support transformative insights. Paper III outlines the features of transformative evaluation and, by reviewing literature from several selected knowledge areas, provides an approach for how evaluations targeting research and policy can be developed in terms of supporting a transition, and in terms of providing knowledge-based learning. Lastly, paper IV looks at the processes pertaining to use of evaluations at Swedish state agencies, with a focus on perceived use and the readiness among the agencies to move towards transformative evaluation.

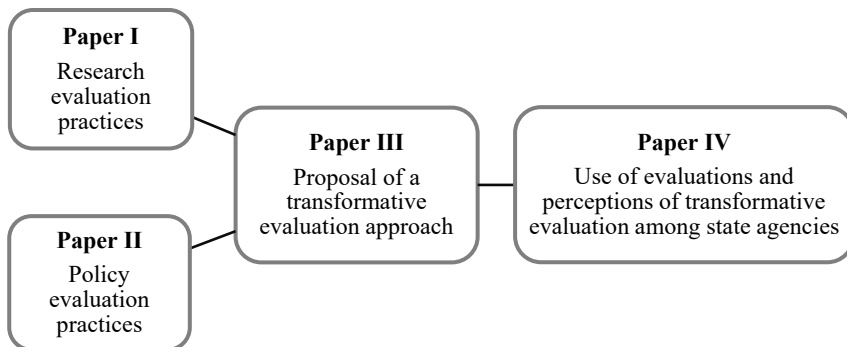


Figure 2. Overview of research papers included in this thesis and their relation to the posed research questions.

1.5 Research scope

Based on the case used in this research – Swedish evaluation practices for research and policy incentives for energy efficiency in buildings – this thesis takes a sectoral, geographical, and a theoretical scope.

1.5.1 Sectoral and geographical scope

The case study undertaken in this thesis focuses on evaluation of research and policy incentives for energy efficiency in buildings in Sweden. This concerns the practices of evaluation, as well as relevant state agencies' perceived use of evaluations and their readiness to move towards transformative evaluation.

The energy sector is essential in our societies for supporting health, food production, services and security (GEA, 2012). The focus in this thesis, energy efficiency in buildings, was selected as a key area within the research project Transition Governance due to its large potential for addressing sustainability challenges, and due to the range of enforced public incentives and attention that the sector is receiving both in Sweden and internationally.

The energy politics in Sweden are guided by the European Union's energy regulations, which are operationalized in the national energy targets. On this note, it should be emphasized that while a key driver for undertaking energy efficiency measures on a European or global scale often is to mitigate greenhouse gas emissions, this has not been the main driver for Sweden as Sweden's energy production is dominated by non-carbon-intensive sources, mainly hydropower and nuclear power (Swedish Energy Agency, 2019). Instead, guided by the political energy agenda adopted already in 1997, which is aimed at facilitating a transition of the energy system to support a sustainable society, the rationales for increasing energy efficiency include resource efficiency and addressing dependence on nuclear power and fossil fuels (Government bill 1996/97:84; Baylan et al., 2017, p. 147). Next, Sweden is seen a suitable case for studying evaluation practices due to Sweden's prevalent and ingrained evaluation culture in the public sector. Evaluation as an articulated public practice can be traced back to the 1960s, but evidently, similar inquiries were made even earlier, and were intended to promote knowledge production and assist decision-making (Furubo, 1994; Vedung, 2010).

In all, the national and global potential of energy efficiency in buildings and the range of different research and policy incentives enforced, in combination with the established status of evaluation as an important activity for providing information in public processes in Sweden, makes this a relevant case for studying the possibilities of transformative evaluation to support a transition towards a more sustainable energy system and society. The potential disadvantages of the selected case study pertain to the limitations of how the scoping to energy efficiency in buildings and Sweden will affect

the opportunities to generalize and transfer learning about how to complement evaluation practices to other contexts. Sweden represents a context with strong institutional structures as well as a largely stable political situation, which may shape the opportunities for evaluation and challenging of public incentives. Energy efficiency in buildings is a delimited part of the energy system, and the building stock and environmental factors (e.g. need for heating/cooling) differ between countries and regions, which in part makes assessments of transformative efforts contextual, but also in part affects the available alternatives for supporting a transition. It is difficult to safeguard for external validity in case studies, but the scope taken in this research may yet provide insights into how to chart and assess current evaluation practices, and how to leverage their strengths to support systemic insights about transformative efforts in a given context. The research methodology is further discussed in Section 3.

1.5.2 Theoretical scope

This thesis takes an interdisciplinary approach in order to provide a broad foundation to address the complexities involved in assessing transformative efforts. It does so by connecting insights from transition research with insights from evaluation-oriented disciplines in order to provide knowledge about how different research and policy incentives may support a transition. Thus, transition research, evaluation theory, sociology of science, and policy analysis have been selected to satisfy the specific knowledge needs to support a transformative and a knowledge-based approach to evaluation. On this note, it needs to be acknowledged that there are other additional disciplines that would also have been valuable to apply for furthering transformative evaluation, which are not accounted for in this scoping (see Section 3). The theoretical scope has been guided by the research task, which is interdisciplinary and largely performed in an applied research context (Bruce et al., 2004). This means that this thesis seeks to provide a representable account of what the different selected disciplines can offer to transformative evaluation, without elaborating deeply on any conflicts or debates within them, and that identification and presentation of insights from these various disciplines is normative.

1.6 Thesis structure

Section 2 presents the theoretical underpinnings drawn from four scientific disciplines that were deemed relevant for studying current evaluation practices and for moving forward towards proposing the outlines of transformative evaluation. This section concludes with a presentation of the analytical frameworks that are built on insights from the different disciplines and used for structuring data collection and analysis in Papers I, II and IV. Section 3 outlines the methodological basis of the thesis, starting

with research positioning and design, and then proceeding with a description of the methods used for data collection and data analysis, as well as ethical considerations. The main findings from the four papers are presented in Section 4, followed by a concluding discussion concerning main findings and contributions in Section 5. The thesis ends with a final conclusion in Section 6.

2. Theoretical background and foundations of analytical frameworks

A transition towards a more sustainable energy system and society involves complex interactions between multiple system components and levels, energy efficiency in buildings being one of them. This means that an understanding of how transformative efforts may create ripples or impacts in these systems requires a wide theoretical angle and approach. Four main research areas have been identified within the research project Transition Governance as important to form the theoretical basis and central features of transformative evaluation: transition research, evaluation theory, sociology of science, and policy analysis. These were selected because they represent various approaches to considering how transformative processes emerge and evolve, which roles different research and policy incentives have in such processes, and how these can be gauged.

The following sub-sections present key insights from these disciplines to support understanding of how transformative changes are conceptualized, and how drivers, components, and other factors can be readily approached and assessed when moving towards transformative evaluation. The section concludes with a brief description of the analytical frameworks to which these insights have contributed.

2.1 Transition research

With a focus on the drivers and components of societal transitions, transition research is an essential point of departure for moving towards transformative evaluation. By and large, transition research seeks to understand, explain and conceptualize the emergence and processes of large-scale changes in society. Building on different core literature strands, including science and technology studies, innovation theory, and technological regimes, a focus on transitions and technological development for sustainability in particular was attracting growing attention in the 1990s, and has since developed a range of orientations for describing the mechanisms and concepts pertaining to societal change processes for sustainability (Kemp et al., 1998; Markard et al., 2012; Rip and Kemp, 1998).

In order to identify key aspects of a transition to be regarded in transformative evaluation, there is a need to understand how the mechanisms and components of a transition interact. Scholars in the transition research field have proposed various approaches for conceptualizing these processes, of which the multi-level perspective has become well established and recognized among scholars (Geels, 2002). In this model, societies can be described as socio-technical systems, which are composed of micro, meso, and macro levels that are interlinked and interacting, and which together dictate the current and stable status quo (Rotmans et al., 2001). This status quo is often referred to as the socio-technical regime, the established and institutionalized practices which are stable and reluctant to undergo rapid change. Socio-technical regimes are minted by an entrenchment in incumbent trajectories and technologies, in lock-in patterns where current structures are protected and upheld by those benefitting from them (Kemp et al., 1998; Meadowcroft, 2009; Smith et al., 2005). Thus, changes that do occur in socio-technical regimes are rather incremental and unlikely to deeply alter the current configuration.

However, changes needed for more sustainable practices and societies will require deep, radical, changes in current systems (Markard et al., 2012), and, important for this thesis, a shift away from analysing individual effects of products and innovation and towards regarding their implications on a society-wide scale (Smith et al., 2010, 2005). Uncovering the role of the different components acting for or against this change process is thus essential for transformative evaluation. A transition happens when radical changes do occur in a socio-technical system, which is reconfigured into a new status quo, entailing an establishment of new technologies, practices, norms and institutions coming into domination (Kemp, 1994; Markard et al., 2012). The socio-technical regime is found on the meso level of the system, and changes and pressures may come from the micro or macro levels, or be the result of pressures arising from both. The micro level is recognized as niches (Geels, 2002), spaces where other norms and visions than what are otherwise prevalent in the system are nurtured (Smith et al., 2010), whereas the macro level is recognized as the landscape, which hosts overarching demographic and environmental conditions, as well as political, economic, scientific and social paradigms and structures that support or challenge current regimes (Geels, 2002; Smith et al., 2010).

For transformative evaluation, this means that there are many aspects to consider to gain an understanding of the contributions or the role of an evaluand in a transition context, where pathways for change are complex, multiple, and intricate.

2.1.1 Key concepts for transformative evaluation

Drawing from this overarching conceptualization of the emergence and processes of transformative change, certain aspects can be identified as essential to consider when

seeking to capture contributions to a transition in an evaluation. In this thesis, these transformative aspects are identified as main components of the transition, and main processes that drive the transformative change.

Firstly, the components of the transition are identified as *systems*, *scale*, and *multi-actor perspectives*. These components seek to illustrate the system in which the transition is (intended) to take place. A *systems perspective* means to acknowledge the various elements that shape the system, such as the actors, structures and institutions. Such a systems perspective also seeks to capture the potential transition pathways in order to elucidate how changes may emerge and gain momentum (Geels and Schot, 2007). Relatedly, the *scale* of the evaluand is of importance for its effects on a transition; and assessments may concern whether it is still a fairly isolated experiment, if it is about to be upscaled to a regime level, or what the potential for upscaling is with regard to barriers and opportunities. The last component suggested as important concerns the actors involved in, or affected by, the implementation, and their roles in the system. Again tightly connected to the systems perspective and the scale perspective, socio-technical systems are minted by the actions of people with different agendas, interests, and power (Avelino and Wittmayer, 2016; Rotmans et al., 2001). Evaluation seeking to inform about transformative change efforts and potential will need to consider the roles of these people in relation to the evaluand, in terms of their supporting or counteracting it (see, for example, Rotmans et al., 2001). This requires a deliberate *multi-actor* approach and transparency about which actor groups are involved, and consequently which are not.

As for the second category of aspects important for capturing transformative change in an evaluation, i.e. the processes that drive a transition, transition research outlines some key drivers which can be described as *visioning*, *experimentation* and *learning*. In order to destabilize and reconfigure socio-technical regimes, there need to be forces coming either from the landscape level, in terms of e.g. new sustainability criteria or environmental pressure, or experimentation and innovation coming from niches that manage to break into and become adopted in the regime. While it is argued that transitions cannot be fully managed or controlled, there is still acknowledgement that they may be guided, e.g. by governance or by *visions* (Loorbach, 2010; Meadowcroft, 2009). Visioning is thus, in this sense, the visioning among stakeholders of where a transition needs to go and what trajectory the regime shifts should follow (Meadowcroft, 2009; Rotmans et al., 2001). Linked to the multiple actors in the system, moreover, it needs to be acknowledged that there may be different visions about what a desirable transition is according to different actor groups. Thus, in order to assess an evaluand in relation to supporting a transition, the underlying vision of what the evaluand supports is of the essence.

The other two processes that underpin transformative change are *experimentation* and *learning*, which are closely related and support each other mutually (Geels, 2002; Schot

et al., 1994). Experimentation is the action of developing and testing new ideas, innovation and technology; and learning is the essential process of taking stock of gained experiences and taking subsequent steps for scaling up (Schot et al., 1994). Moreover, experimentation may arguably not only concern technological innovation in niches, but may also take place at landscape level in experimentation with governance structures, science agendas and policy instruments (Loorbach et al., 2017). In order to promote and support a sustainability transition, there is a need to deliberately evaluate experiments and efforts, and to be explicit about lessons learned so that these can be reiterated in future initiatives. Evaluation is in itself a measure for supporting learning, and thus holds great potential to provide insights to be used for designing research and policy instruments for a transition (Luederitz et al., 2017).

The aforementioned key components and drivers are argued in this thesis to be particularly important for advancing and designing transformative evaluation, due to their prevalence within the transition literature, and with regard to their pointing to core aspects of society, actors, and ambitions for a transition. While transition research offers many levels of complexity in describing and understanding transitions, these concepts are proposed because they together encapsulate a wide range of aspects that may be essential for capturing transformative contributions of an evaluand, while at the same time being flexible enough to be adapted in level of granularity to the specific contexts of each evaluand and evaluation.

To sum up, as Patton (2019, p. 103) aptly points out, “*Evaluating transformation means transforming evaluation*”. As such, there is a need to understand the foundations of evaluations in order to identify essential aspects that can complement transition research with knowledge-based learning about how research and policy incentives may support a sustainability transition. The following sub-sections outline some key contributions from three disciplines in this regard.

2.2 Evaluation theory

Evaluation as an activity is ancient and has been conducted for centuries, for reasons as varied as the evaluation of the quality of crafts and goods, to the evaluation of the performance of public efforts (Scriven, 1991; Weiss, 1998). The evaluation of policy programmes as we recognize it today has, however, only since the second half of the 20th century developed its own set of theories and become acknowledged as its own discipline (Rossi et al., 2004; Scriven, 1991; Shadish et al., 1991; Weiss, 1998). This section seeks to establish some core positions taken in this thesis, with regard to the identification and synthesizing of relevant insights from evaluation theory for advancing rigorous and knowledge-based learning in transformative evaluation.

As evaluation has been applied in various sectors and contexts (evolving from a type of meta-theory belonging to specific disciplines into its own theory (Scriven, 1991)), there are varied views on what the term entails. Thus, it is necessary to establish what is considered ‘evaluation’ in this thesis. While there have been some seminal suggestions made by scholars for defining evaluation (see, for example, Scriven, 1991; Weiss, 1998), this research adheres to the definition proposed by Vedung (1997, p. 3):

“Evaluation = df. Careful retrospective assessment of the merit, worth and value of administration, output, and outcome of government interventions, which is intended to play a role in future, practical action situations.”

From the perspective of advancing transformative evaluation, this definition captures the essence of supporting rigorous and knowledge-based assessments in evaluations, and rendering evaluations parts of decision-making and action – in this case to support a transition towards a more sustainable energy system and society. However, while Vedung (1997) excludes prospective ex-ante assessments from the definition of evaluation, this research does include such assessments. Relating this to transition research, transitions may take long time, which consequently demands some prospective assessments for providing knowledge about any potential future outcomes relating to transformative processes. It is also based on the recognized importance of prospective assessments for environmental policies in particular (Mickwitz, 2003), and on the fact that ex-ante assessments are included in guidelines for evaluation by the European Commission (2017), as well as called for in the European Union’s 6th and 7th Environmental Action Programmes (Schoenefeld and Jordan, 2019).

Evaluation belongs to and is an essential component of many sectors and disciplines, but since this thesis takes a focus on sustainability and energy, a few words on the role of evaluation in the sector of environmental sustainability are warranted. First, a key feature of sustainability policy is that as knowledge about the complexity of sustainability problems has increased, the span of policy instruments that are to address these challenges has also grown (European Environment Agency, 2016). This increased understanding of the complexity of sustainability has also meant that the complexity related to evaluation of these implementations has grown, as it is now tasked with addressing the intricacy of the systemic connections from multiple measures, as well as the complexity of the sustainability challenges (Birnbaum and Mickwitz, 2009; European Environment Agency, 2016; Julnes, 2019b). Recently, it has also been recognized that if an evaluation practice that is sensitive to the needs of policy-makers and stakeholders facing complex sustainability decisions is to be secured, evaluations need to transform in order to address such challenges and provide useful and relevant insights and knowledge (Patton, 2019; Rowe, 2019). Arguably, such evaluation needs to build on plurality, both in disciplines and in evaluation approaches.

What are then essential insights from evaluation theory to bring forward for supporting transformative evaluation with rigorous and knowledge-based assessments? The field of

evaluation theory hosts many different approaches to evaluation processes, analysis and assessment, but this thesis uses Christie and Alkin's (2013) evaluation tree to further articulate some key aspects to be carried forward from evaluation theory. The tree has three branches that represent main considerations for evaluation, namely *methods*, *valuing* and *use*, which will guide the following sub-sections.

2.2.1 Methods

The evaluation field has undergone a broad evolution in terms of development of theory and methodology, moving from predominantly quantitative approaches towards a complementation of qualitative approaches (Cook, 1997; Vedung, 2010). Methods for evaluation are underpinned by ontological and epistemological standpoints, where the credibility of the knowledge production is minted by an evaluator's or a stakeholder's perception of what reality is and how it can be measured or construed; and how validity is assessed and which claims can be made to causality (Shadish et al., 1991).

Broadly speaking, two main epistemological traditions can be distinguished for approaching evaluation: the rationalistic and the argumentative approach (Bovens et al., 2008), which build on positivist and constructivist ontologies (see also Vedung (2010) for additional main evaluation approaches). Modern evaluation was originally mainly oriented in rationalistic approaches, with an emphasis on experimental and quasi-experimental methods (see, for example, Campbell, 1969; Suchman, 1967). Such approaches depart from a positivistic stance where methods for analysis and assessments are rooted in natural sciences, seeking to provide empirical and objective data which can be controlled for bias, and which can provide generalizable results (Bovens et al., 2008; Vedung, 2010).

The argumentative approach, on the other hand, disputes the separation of facts and values, and departs from the stance that evaluation cannot be objective since analysis and results will be subjected to interpretation from the evaluator (Bovens et al., 2008). Based on a constructivist ontology, this approach holds that there is not one reality that can be readily measured, but that instead, what is perceived as real is subjective and is depending on each individual (Christie and Fleischer, 2009). Thus, rather than striving for value-free empirical data, evaluators acknowledge that there will be bias in results, and instead place stakeholders and their perceptions of an evaluand at the heart of inquiry (Christie and Fleischer, 2009; Guba and Lincoln, 1989). Evaluation thus revolves around stakeholders' views and a pluralistic approach to how different viewpoints affect the constructions of what is true.

These two main approaches to evaluation bring different implications for providing knowledge-based learning in transformative evaluation, in terms of *which methods will be applied*, and *which stakeholders are involved*. Tying them to the core concepts from

transition research, there is an emphasis on actor involvement as emphasized in argumentative approaches, in order to uncover e.g. power structures and agendas. However, transition research also suggests that drivers and components interact in certain ways towards a transition, which calls for some causal assessments as provided by more rationalistic approaches.

Another methodological approach suggested in evaluation theory that is interesting to consider for transformative evaluation is the theory-based approach. This method seeks to assert the *intended* workings and processes of interventions, thus addressing what is known as the 'black box' (see, for example, Bickman, 1987; Chen, 1990; Chen and Rossi, 1989; Rogers et al., 2000; Weiss, 1997). Theory-based approaches build on *intervention theories*, also known as e.g. programme theories, or logic models, which are constructed to elucidate the causal relationships between programme activities and target groups, through intermediate and final outputs and outcomes (Bickman, 1987; Rogers et al., 2000; Rossi et al., 2004; Vedung, 1997; Weiss, 1998). By outlining the intended functions of programmes, theory-based approaches are used to focus and drive the evaluation (Meyer and Stockmann, 2013; Rogers et al., 2000), and to separate between programme failures (implementation issues) and theory failures (flawed programme design) (Bickman, 1987; Weiss, 1997). For transformative evaluation, theory-based approaches hold the potential to provide overviews of mechanisms that may promote transformative outcomes, as well as provide knowledge-based learning on where to aim efforts to support such transformative outcomes; for example, in terms of knowing if an absence of effects is due to a poor implementation scheme, or a faulty design in how mechanisms would interact.

This overview of roots and methodological approaches to evaluation shows that there is a plethora of approaches to consider when planning an evaluation, and clearly, this will also apply to transformative evaluation. Different methods have different strengths and are suitable for addressing different aspects, and there is thus no evaluation approach that meets all needs and expectations (Shadish et al., 1991; Weiss, 1998). Instead, methods should be selected that cater to specific needs and uses. The middle ground is found in the use of *mixed methods*, both quantitative and qualitative (Christie and Fleischer, 2009; Greene, 2008; Mertens and Hesse-Biber, 2013; White, 2013), in order to provide opportunities for methods *triangulation* for control and validation of results (Denzin, 1971; Howe, 2012; Patton, 2002). Moreover, another important methodological aspect to knowledge-based approaches, as emphasized in evaluation research, is to provide *counterfactual assessments* for gauging impacts of an evaluand (Ferraro, 2009). This in order to elucidate any effects that can in fact be attributed to the evaluand, and to further understand what role the evaluand has in relation to other implementations and societal developments.

A key consideration to be made, however, in moving towards transformative evaluation concerns how the methods used in an evaluation relate to key drivers and components

of transformative efforts, as outlined from transition research. In order to couple these drivers and components with evaluation theory insights, methods applied should, when appropriate, seek to be inclusive and support a multi-actor perspective, and to adequately articulate the role of the evaluand in relation to the vision, while at the same time providing robust assessments of causal effects and their societal implications.

2.2.2 Valuing

Valuing is central to evaluation, and evaluation is thus not just about determining facts but certainly also about providing some sort of value judgement to the facts (Alkin et al., 2012; Shadish et al., 1991). Consider the multiple value systems that environmental policies are subjected to, which include ecological aspects as well as social dimensions such as democracy and equity. This calls for transparency concerning value constructions in evaluations of these policies, if assessments that are just and reliable to different stakeholders are to be provided (Crabbé and Leroy, 2008; Mickwitz, 2003).

For transformative evaluation, the need to consider value judgements is arguably more important than ever, given that a transition will come with deep changes that affect stakeholders system-wide. For this, evaluation theory offers the broad distinction between prescriptive and descriptive approaches to valuing (Shadish et al., 1991). In a prescriptive perspective, values are derived from certain selected theories that are placed at the heart of the inquiry by the evaluator. These values are thus seen as superior to other value constructions as far as the evaluation is concerned, and guide the evaluation to assess an implementation in relation to these values (Shadish et al., 1991; Vedung, 1997). In such evaluations, valuing may be guided by the evaluator's knowledge about the programme, by evaluation excellence, or by morals or other values held by the evaluator (Alkin et al., 2012). In descriptive valuing, values are instead derived from stakeholders' perceptions or opinions about an implementation, or from its goal formulations. These values can be determined before the evaluation is conducted, and are not seen as superior or more appropriate than others. Instead, they serve the purpose of understanding stakeholders' views on the performance of a programme, and can as such provide guidance on measures for improvement, etc. (Shadish et al., 1991; Vedung, 1997). A third category for valuing builds on value constructions drawn from both the evaluator and the stakeholders. In such approaches, the evaluator may support or guide stakeholders in developing frameworks or other approaches for making their own value judgements (Alkin et al., 2012).

Value judgements can be operationalized in an evaluation using *criteria* that seek to determine the quality and performance of an implementation (Stake, 2004). These criteria are then commonly based on some established standard that dictates how well the evaluand meets the criteria, which provides a comparison for the performance of the evaluand (Brewer and DeLeon, 1983). There are multiple criteria to be considered

for an evaluation, the selection depending on the purpose of the evaluation and the rationales for value judgements (Dunn, 2018, p. 197; Weiss, 1998). One key for transformative evaluation is arguably that criteria appropriately represent different stakeholders, while also providing rigorous and knowledge-based assessments that are thoroughly anchored in relevant and justified approaches.

More concretely, the European Environment Agency (2016) emphasizes some criteria which are considered particularly relevant for environmental and climate policy, namely: *effectiveness*, the extent to which effects correspond to programme objectives; *efficiency*, if costs can be justified in relation to benefits; *relevance*, if programme objectives address societal needs; and *coherence*, if a programme is coherent with other programmes, or internally coherent with regard to objectives and goals. Another criterion that is important for the particular case of climate and energy policy is *sustainability*. Initially, this criterion was intended to measure whether effects from an implementation would sustain themselves after the termination of an implementation (OECD/DAC, 2019; Patton, 2019). However, in light of current sustainability challenges, this criterion has been extended to not focus only on statically lasting effects, but also on solutions that work at system levels, under dynamic, complex, and context-bound conditions (Patton, 2019). Other criteria that are frequently proposed in evaluation theory broadly include *equity*, *adequacy*, *responsiveness*, *persistence*, *flexibility*, and *predictability* (Brewer and DeLeon, 1983; Dunn, 2018; Mickwitz, 2003).

Clearly, the criteria that are used to make value judgements, and their operationalization in an evaluation, will affect the outcome of the evaluation, and conflicting interests among different stakeholders call for a just and transparent treatment of the criteria selection. *Applying multiple criteria* in the same evaluation is thus a means to balance different perspectives of stakeholders, and the many potential outcomes of an evaluand (Brewer and DeLeon, 1983; Crabbé and Leroy, 2008, p. 26), and is suggested in this thesis as a key for promoting knowledge-based learning which sheds light on the many workings and effects of a research or policy incentive from a transformative perspective.

To conclude the value section, there are some key implications for transformative evaluation to be considered. These relate to the complexity of the evaluand as embedded in a system with multiple actors with various expectations and demands. Navigating these in an evaluation means carefully considering prescriptive and descriptive approaches to valuing, and acknowledging the actors who are represented (or not) in the criteria adopted for making assessments of an evaluand. It also encompasses a consideration of the level of *reflexivity* that the chosen criteria support, with regard to elucidating the assumptions that underpin the evaluand, and challenging the goals set. Related to the key components from transition research, a multi-actor approach and placing the evaluand in relation to a vision are essential aspects to be tied to the value judgements used in transformative evaluation.

2.2.3 Use

The third key area of evaluation, as outlined in the evaluation tree, is use. Clearly, if an evaluation is not put to use, it becomes a missed opportunity to act upon the learning and knowledge generated. The necessity of utilizing evaluations to improve and make informed decisions going forward cannot be stressed enough, which certainly also is key for transformative evaluation in terms of supporting a transition towards a more sustainable energy system and society.

What are then the underlying processes for use, and for promoting enhanced use? Firstly, evaluations are conducted for a certain purpose and with a certain idea of how they are to be used (Vedung, 1997; Weiss, 1998), which means that in order to advance transformative evaluation, the reason for and intended use of an evaluation need to be turned towards how it is proposed to support a transition.

In terms of then supporting actual use of an evaluation, evaluation theory describes the immediate use as largely distinguished between *process-related* use and *findings-related use*, where process-related use takes place during the ongoing evaluation, and findings-related use means acting upon the results generated in the evaluation (Alkin and Taut, 2002). These two main types of use can be further expanded in a wide range of potential *models of use*, which in turn correspond to various purposes for evaluating. These models of use often represent purposes pertaining to accountability, knowledge production for basic knowledge, and improvement, but also illustrate alternative purposes for evaluating, such as for supporting tactical or persuasive agendas (Vedung, 1997; Weiss, 1998) (see Table 1 in Section 2.5.3 for a more elaborate account of different models of use).

The move towards transformative evaluation arguably has room for many different models of use, both process- and findings-related. In order to promote a just evaluation use, the purpose of supporting a transition should be openly communicated, along with an outline of how the results from the evaluation are intended to serve such a purpose.

There are some key measures for enhancing use that are important to bring forward in order to support a transformative use of evaluations. First, involvement of stakeholders and a clear purpose are essential (Patton, 2008; Rossi et al., 2004; Vedung, 1997), as well as a credibly executed evaluation which is grounded in robust methodologies, and which is sensitive to the value constructions and contextual factors pertaining to political aspects and intended use of the evaluation (see, for example, Chelimsky, 2015; King and Alkin, 2019; Leviton, 2003). It may also be important to understand why evaluation use fails. Two main categories for non-use are organizational resistance, where evaluation results raise critique towards programme quality, administration and the like; and political resistance, which is grounded in reluctance towards evaluation findings due to their negative influence on e.g. elections, funding, etc. (Chelimsky, 2015; Weiss, 1998).

As transformative evaluation seeks to support insights about transformative effects on a systems level, and thus seeks to broaden the horizons about how effects may have influence beyond the evaluand itself, it is about bridging the immediate use of evaluation processes and findings with what is called influential use, characterized by utility of an evaluation outside of the evaluated programme or institution (Patton, 2008). Thus, rather than focusing on the more traditional use of evaluations, revolving around instrumental improvement and accountability (Patton, 2008), evaluations are intended to be used for knowledge production concerning systemic effects. However, transformative evaluation – like traditional evaluation – may bring challenging results or calls for action. Thus, a key for its success is an openness among users and stakeholders to the learning and knowledge that an evaluation brings, and a transparent consideration of the implications of evaluation findings for different stakeholders.

2.3 Sociology of science

Moving towards transformative evaluation requires that evaluations take a comprehensive approach concerning how to understand and assess the processes of science and research, and to consider how these processes are shaped in relation to the social context they are part of. This includes e.g. how research can be understood from the perspectives of societal and political visions, and as an undertaking shaped by contextual factors of governance and funding. In order to provide some foundations for how transformative evaluation may address such issues, this thesis turns to sociology of science, a broad discipline which addresses the role of research in society, its relevance to the needs of people, and its effects on social, ecological, and economic factors (Bucchi, 2004).

Sociology of science presents a range of different perspectives on research as a part of societal processes. Against this background, this thesis departs from the articulated need to realize a transition towards a more sustainable energy system and society, and sees research and science as important factors driving such change. It also departs from the viewpoint that such targeted efforts are not necessarily spurred on their own, or used effectively, and thus, there is a need to understand how to support, design, govern and direct resources to incentives with appropriate characteristics for the envisioned transition. This means that research is seen as an activity that can be steered towards set objectives, and which consequently can be assessed and evaluated against these.

Understanding the role of science and research in relation to society has been described as a progression from *Mode 1* to *Mode 2*. *Mode 1* sees science as oriented around disciplines, where knowledge production is fundamental rather than applied, and upheld and performed by a clearly identified scientific community (Gibbons et al., 2010). From this very disciplinary approach to knowledge production, there has been a progression towards *Mode 2*, where research problems began to be viewed with a

curiosity towards how this theoretical, foundational knowledge could be applied in practical areas of use. As such, *Mode 2* is problem-oriented, and emphasizes transdisciplinarity to allow knowledge from various disciplines to merge in the pursuit of addressing a specific issue. This kind of knowledge production brings together a range of researchers, while also facilitating communication between science and society, as the investigated problem creates a contact surface for actors representing various sectors, including academia, industry, and business (Gibbons et al., 2010). The two modes are at work in parallel, and are mutually exchanging in that basic science can be applied, but applied science and its discoveries can also prompt further basic research (Gibbons et al., 2011, 2010). Thus, both types are called for to support a functioning and effective model of knowledge production.

The actors involved in the current state of science policy thus go beyond academia and the *Mode 1* disciplinary perspective. Instead, there are many actors representing science, politics, industry, market, and civil society that create self-organizing networks and communities of *actors that are affecting and shaping the governance of research* (Colebatch, 2009; Haas, 1992).

The implication of this for transformative evaluation is that research initiatives are found in *multiple locations in a societal system*, bound by contextual factors, and promoted and undertaken by different stakeholders with different goals and agendas. This means that *boundaries between science and society* have become flexible as politics and science have become merged (Borrás, 2012; Guston, 2001), and as scientific knowledge is deliberately used to affect political decision-making (Ezrahi, 1980). Thus, the various shapes of utility of research clearly challenge its evaluation, particularly with regard to defining the *aims and objectives of research*, as these may span clear-cut goals which can be more readily evaluated, and goals wide enough to accommodate the diverse kinds of objectives and uses the research may have (Ezrahi, 1980). For transformative evaluation in particular, it also calls for a need to connect the research incentives with the political visions for a transition, heeding the involved actors' agendas and interests and rationales for undertaking the research.

2.3.1 Evaluation of research

Evaluation of research plays an important role in shaping research policy, e.g. in the case of determining, justifying, funding, and undertaking the research needed for realizing a sustainability transition. Consequently, evaluations can also become powerful tools for steering agendas, and the way evaluations are composed will shape the outcome of the evaluation. Research evaluation thus needs to navigate an intricate field of disciplines, governance and political agendas. This section outlines some key insights on research evaluation practices drawn from sociology of science that can be brought forward into transformative evaluation.

The evaluation of research is a well-established part of academic processes and in many ways has shaped the evolution of the field, e.g. through assessments and evaluations related to publishing procedures, to applications for funding, in the shape of accountability checks, and not least in rankings of universities (Hazelkorn, 2015). Due to the central role of evaluation in research governance, the field of sociology of science provides a variety of approaches to performing evaluations, which will be briefly outlined in order to illuminate key aspects for transformative evaluation.

Among the main rationales for carrying out an evaluation – accountability, basic knowledge, and improvement (Vedung, 1997) – accountability is actualized in discussions surrounding the use of public funding for different research initiatives: what money is spent on, and how the outcomes benefit society (Elzinga, 2012; Geuna and Martin, 2003). Over the decades, as science as a societal activity has grown and attracted increasing funding, there has been an increasing call for science to be more efficient (Elzinga, 2012). This can be seen in accountability audits, which have moved focus towards assessing the output of research in relation to its input (Elzinga, 2012). This assumed level of causality and correlation between *input and output* has seen quantitative metrics such as *bibliometrics* being increasingly applied to assess efficiency and accountability, and to generate insights into how to improve and optimize the return of investment in terms of desired output (Godin, 2007). However, the trust in bibliometrics and quantitative approaches to research evaluation has been questioned, as not all research output can be – or should be – measured quantitatively, as they also provide other types of benefits to society which quantitative approaches neglect to account for (see, for example, Toulmin, 1966). The methods for evaluating research are thus broader than just quantitatively guided bibliometrics and efficiency assessments, but also span *qualitative approaches* such as peer review (Geuna and Martin, 2003; Langfeldt, 2001; Vedung, 1997) and *network analyses* (Leite and Pinho, 2017), as well as other approaches to evaluation, as is illustrated by the different types of research evaluation systems that are employed in different countries (Gläser, 2007).

Carefully established value judgements are essential for performing just assessments, and particularly so when assessing transformative efforts that may affect stakeholders differently. However, for qualitative research in particular, evaluation practices are scarcely formalized (Cho, 2017; Mårtensson et al., 2016), and inter- and transdisciplinary research largely does not have any agreed criteria for assessing conducted research (Belcher et al., 2016). This may have implications for evaluations that target research areas that are interdisciplinary and that include qualitative research, such as energy-related research, as key decisions related to methods, value judgement, and use of evaluations may lack guidance in the form of recommended best practices. This tension may, however, also provide fertile grounds for developing, testing, and adopting alternative approaches to evaluating research.

Criteria that are often applied in research evaluations include *quality* and *excellence*. These can be geared towards the output of research, or towards the processes of research that can in turn lead to excellent output (Hellström, 2011). However, while ‘excellence’ was originally oriented around efficiency criteria, as proposed by the input/output models, more recent metrics instead revolve around capacity building through e.g. networking, mobility and exchange (Hazelkorn, 2015). To evaluate quality and excellence of research, some kind of operationalization of what excellence is, and how it can be measured and attained, is required. Such operationalization will to a large extent be individual to the academic discipline, which may have different goals and standards of scientific output. This means that excellence quickly becomes contextual, and thus not readily comparable between different research programmes or institutions (Hellström, 2011).

In order to address such issues in transformative evaluation, evaluations need to be grounded in an understanding of how research is seen to interact with society from many perspectives, as different actor groups may see this differently. Thus, a multi-actor approach is emphasized, which can assist in uncovering how a research initiative relates to a transformative goal or vision. It is also critical to consider how the upholding of certain research structures may favour already-existing research fields, and how a co-creation of knowledge between a limited range of actors has the power to further solidify (unsustainable) structures, research agendas and scopes (Rip, 2016), which may counteract a sustainability transition. Relatedly, the potentially adverse effects of an evaluation on actor groups also need to be uncovered, such as inequality and loss of jobs resulting from technological advancements (Sarewitz et al., 2004), if a just assessment of how a research incentive indeed supports a sustainability transition is to be made.

In sum, there are some essential aspects to be drawn from sociology of science for producing knowledge-based assessments in transformative evaluation. First, a look to the *operationalization of research* is suggested. This concerns the *actors* who are shaping and conducting the research, how the research is envisioned to contribute to *visions*, and which *boundaries* are set up for delimiting the research area, in part in relation to how knowledge production is understood and conducted, and in part in relation to other *governance areas* that shape how it is conducted. Secondly, the variety of approaches to evaluating research calls for transparent and deliberate approaches to *analysis*, concerning the methods used, the actors involved, the focus and *directionality* of inquiry, and how these support robust and just assessments. Lastly, a consideration is suggested of how *assessments* in evaluations are made, with regard to the value judgements and criteria that are applied, and whether the evaluation is communicated to and used by its audience as intended to support various end goals.

2.4 Policy analysis

Policy analysis provides insights into the roles and processes underpinning public policies in society and, as such, is essential for transformative evaluation as a means to guide evaluation designs and assessments.

Broadly speaking, policy analysis takes a comprehensive approach to understanding public policy processes, and is intended to provide insights that support decision-making and policy-making (Dunn, 2018; Hogwood and Gunn, 1984). Policy processes have been conceptualized as *policy cycles*, describing the chronological and cyclical steps following policies. These are, in general terms: agenda setting, policy formulation, selection and implementation of policies, evaluation of the implementation and subsequent adaptation, and ultimately, policy termination (Dunn, 2018; Spicker, 2006). Thus, decision-making is an integrated part in all steps of the policy process, taking place at different times and engaging different actors. Policy analysis seeks to support this decision-making, for example by assisting in needs and problem formulation, in advocacy for either a certain policy or a certain underlying policy process, in the identification and selection of possible alternatives and solutions in terms of policies, and in evaluation of the policy (Dunn, 2018; Parsons, 1995; Spicker, 2006). As such, policy analysis is oriented around assessments and knowledge production that is intended to be *informed by actors* and to be fed into the decision-making process by actors. A key in policy analysis is thus that it builds on *insights held by different actor groups*, who are guided by their own value systems, and whose perceptions and opinions are shaped by *social systems and contexts* (Spicker, 2006; Wagenaar and Cook, 2003).

From a transformative evaluation perspective, there are certain key issues to highlight from policy analysis. First, value judgements in policy analysis can, like general approaches in evaluation, be seen as either prescriptive or descriptive, as revolving around values based either on the perceived optimal or ideal workings of a policy, or on how a policy is performing and seeking to explain and understand workings based on causal relationships (Dunn, 2018; Hogwood and Gunn, 1984). Next, policy analysis is inherently political, in that there are underlying political viewpoints and processes that underpin decisions for policy-making. As such, policies are shaped by relationships between *multiple stakeholder groups* in various places in complex networks (Hogwood and Gunn, 1984). Involvement of stakeholders in policy analysis is thus crucial for navigating different perceptions and agendas. More concretely, actors to be considered are those with key roles in the implementation and those who are affected by the policy, as well as the general population (Spicker, 2006). Furthermore, the intricacy in actor relations and political agendas also calls for *ethical* considerations, e.g. in terms of making issues surrounding accountability and democracy transparent, since public policy handles public assets and affects people, and since what is to be deemed 'societal improvement' varies and is dependent on the viewpoints held by different

stakeholders (Dunn, 2018; Spicker, 2006). These are certainly key issues that are also highly relevant to consider when designing transformative evaluation approaches.

Relatedly, another key aspect to highlight is the *various effects and outcomes* that a policy incentive may bring. While a policy carries with it some embedded goals or targets that articulate which kinds of outcomes are desired or expected (Spicker, 2006), this may not be how it actually plays out. Thus, although there is a clear idea of what outcomes *ought* to be, policies are likely to have *both expected and un-expected outcomes* (Hogwood and Gunn, 1984), which need to be acknowledged in evaluations (Mickwitz, 2003; Neij and Åstrand, 2006; Schalock, 2001). In order to address and uncover such varied outcomes, the policy cycle outlines assessment activities both prior to implementation and after (or during) the implementation has taken place. These assessments are often referred to as either prospective or retrospective, or *ex-ante and ex-post* evaluation (Dunn, 2018). Ex-ante assessments are aimed at forecasting and understanding the future effects of any decisions in order to support the selection of which alternatives to go for (Dunn, 2018; Hogwood and Gunn, 1984). Ex-post evaluation takes place after the implementation has started, and seeks to provide insights on how the implementation has performed against selected criteria. The intention is to support decisions on whether to change, continue or terminate the implementation.

The *timing of evaluation* is crucial when considering which kind of knowledge is needed at which point in time of the policy cycle. The rough distinction of formative and summative types of knowledge production has been contested as being too simplifying (Patton, 1996; Wholey, 1996), yet it serves the purpose of articulating the use potentials of different types of knowledge. As such, formative knowledge is produced at the initial stages of the policy cycle and during a policy implementation, and is intended to support decisions on improvement and processes, while summative knowledge takes over at the end of (a phase of) the implementation and provides insights about how the implementation worked and delivered on intended outcomes (Scriven, 1996; Weiss, 1998). Both types are arguably important in transformative evaluation in order to provide a continuous overview of how policy incentives are planned to work in relation to a vision of a transition, and how their actual effects and outcomes are supporting this goal.

Another key issue related to time, which is especially relevant for environmental policies and transformative evaluation, is what time horizons should be applied in assessments (Crabbé and Leroy, 2008; Mickwitz, 2003). The point in time when an evaluation is conducted in relation to the policy cycle influences the time dimension considered in the evaluation (Hildén, 2009), and with that the potential considerations of future effects. Environmental policy in particular targets large, complex systems in which effects are not necessarily linear (Hildén, 2009), and policy is in itself not prone to rapid change as it is formed as a result of an incumbent system with *lock-in structures* (Dunn, 2018; Hildén, 2009). Thus, the time horizons considered for transformative evaluation in particular

need to be sensitive to the fact that societal systems are slow to change, thus potentially hindering immediate effects, and environmental effects may be incremental until a certain point in time, after which they may be substantial (e.g. considering environmental tipping points or thresholds (Lenton, 2011)). In order to address this, it has been proposed that the *ex-ante/ex-post* distinction in policy cycles be loosened, and that prospective evaluations be allowed to include a retrospective perspective, and vice versa. As such, evaluations become a continuum in which knowledge is used and exchanged to accommodate longer time-horizons (Hildén, 2009).

The long timescales associated with environmental issues is one part of the complexity of challenges concerning environmental policy evaluation, which is further complicated when considering the *interconnections between different policies* in relation to addressing social, economic and technical aspects (Crabbé and Leroy, 2008; Mickwitz, 2003). There have been calls for cross-sectoral policy integration for environmental sustainability in order to address these challenges on a broad scale (Jordan and Lenschow, 2010). Furthermore, it is also emphasized that policy instruments should not be seen as isolated in a delimited policy arena, but instead as parts of numerous policies working in parallel to achieve similar – or different – overarching objectives. Consequently, the concept of policy mixes has been emphasized for supporting sustainability transitions, as it addresses the challenges related to the real-world policy landscape which comprises multiple actors and governance levels, intricate interlinkages and interdependencies of policies, and the long-term visions of the policy mix (Flanagan et al., 2011; Rogge and Reichardt, 2016).

In summary, what policy analysis offers to transformative evaluation is importantly the outlines of the underlying processes of a policy, with its multitude of actors involved in decision-making, the politicized context and the call for ethics, accountability and democracy, and the identification of the complexity and intricateness of the policy landscape. Transformative evaluations of policy instruments need to be sensitive to the interrelatedness of the system in which they are performed, and appropriate inclusion of key actors needs to be secured in order to support a just and reliable evaluation. Moreover, the importance of the time horizons applied in evaluations for assessing potentially transforming effects, as well as an integrated learning and application of knowledge between prospective and retrospective evaluations, are other crucial elements to capitalize on in order to promote evaluations that support a transition.

2.5 Analytical frameworks

The four disciplines that have been reviewed from the perspective of supporting transformative evaluation inform three analytical frameworks, which present key aspects that are proposed for consideration in evaluations. These aspects seek to capture core ideas, while at the same time keeping a manageable level of detail. Thus, they are oriented around breadth rather than depth.

These frameworks have been used in separate studies of this thesis. The first two were used to guide structured and systematic reviews of existing evaluation reports, as presented in Papers I and II. The third framework was applied in Paper IV to structure an understanding of how evaluations may be put to use. An overview of the compositions of each of these frameworks is presented in the sub-sections below, and presented in their entirety in Appendices A-C.

2.5.1 Framework for Paper I

Paper I focused on the Swedish practices for evaluating research initiatives and institutions within the area of energy efficiency in buildings or building-related research. The overall objective of the study was to shed light on how the evaluations were framed in relation to supporting an energy transition, and the analytical framework that was constructed to guide a comprehensive and structured review of existing evaluation reports was grounded in transition research, sociology of science and evaluation theory.

The framework consisted of three main categories: (I) operationalization of research fields, (II) analysis as performed in evaluations, and (III) assessment and use (see Figure 3). These categories in turn consisted of 12 sub-categories, which outlined concrete aspects to help structure data extraction from the reviewed evaluation reports (see Appendix A).

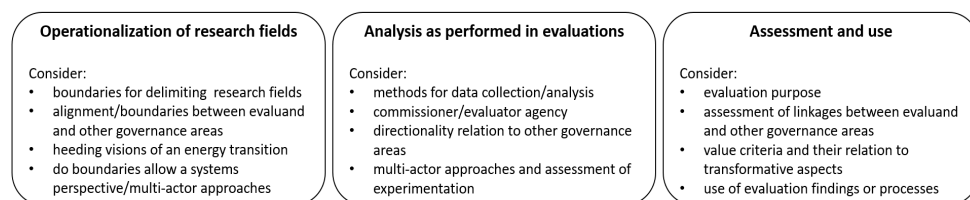


Figure 3. Overview of the analytical framework used for reviewing Swedish research evaluation reports.

2.5.2 Framework for Paper II

Paper II focused on the current Swedish evaluation practices for policy instruments for energy efficiency in buildings. Similar to Paper I, the objective was to illuminate how evaluations performed with regard to supporting an energy transition. The analytical framework that was designed to structure the review and analysis was informed by insights from transition research, evaluation theory, and policy analysis.

Following the conceptualization of the evaluation field as a tree with three main branches (Christie and Alkin, 2013), the framework was arranged around the three categories of: (I) methods applied in evaluations, (II) value judgements in evaluations, and (II) use of evaluations (see Figure 4). These categories were equipped with a total of 16 sub-categories (see Appendix B).

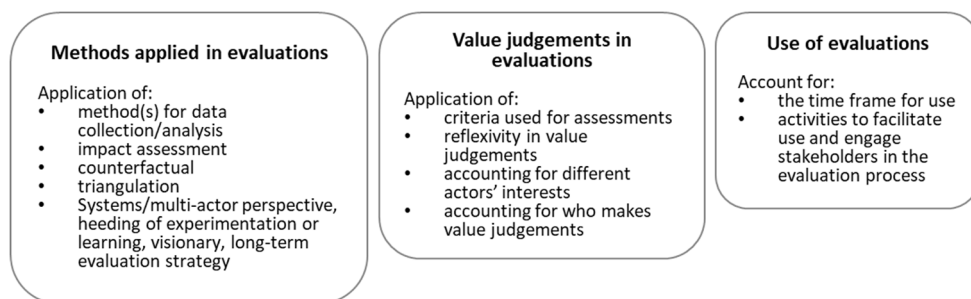


Figure 4. Overview of the analytical framework used for reviewing Swedish policy instrument evaluation reports.

2.5.3 Framework for Paper IV

Paper IV focused on the use of evaluations among Swedish state agencies, with regard to both process-related use and findings-related use. A conceptual framework based on evaluation scholars' work on use of evaluations was compiled to create an overview of different models of use of evaluations (see Table 1). This framework was then used to construct statements to be used in the structured interview questionnaire sent to representatives at the state agencies in order to probe into the current use as reported by the informants (see Appendix C).

Table 1: A list of eleven models of use of evaluations, compiled based on the work of evaluation scholars, as indicated in the table. The models were used to probe into current use processes at Swedish state agencies in a structured interview questionnaire sent to representatives working with evaluation.

Model of use	Characteristics
1. Instrumental Rossi et al. (2004); Vedung (2015, 1997); Weiss (1998)	<i>Evaluation findings are applied immediately for specific actions concerning the evaluand.</i>
2. Enlightenment / conceptual Rossi et al. (2004); Vedung (2015, 1997); Weiss (1998)	<i>Evaluation findings are not used for the evaluand per se, but rather to provide insights about issues or implementations in general.</i>
3. Legitimizing / Reinforcing use Vedung (2015, 1997); Weiss (1998)	<i>Evaluation findings are used to legitimize and justify decisions already made, e.g. to confirm current knowledge and beliefs, or to support confidence in a standpoint.</i>
4. Interactive Vedung (1997); Weiss (1979)	<i>Evaluation findings form part of a larger decision-making process, where other sources of information and actors influence the decisions.</i>
5. Ritual / Symbolic use / Mechanical use Patton (2008); Vedung (2015)	<i>Evaluations are conducted symbolically, because this is expected by current customs and practice; or evaluations are performed mechanically, and evaluands seek only to fulfil requirements and get a good 'score'.</i>
6. Mobilizing use / Persuasive use Rossi et al. (2004); Weiss (1998)	<i>Evaluation findings are used to create support for a particular standpoint, for mobilizing actors.</i>
7. Overuse Patton (2008)	<i>Evaluation findings are put to use as definitive facts, without considering the contextual factors, or without exploring why certain outcomes have emerged.</i>
8. Process use Patton (2008); Vedung (2015)	<i>The evaluation process in itself contributes with insights and learning and spurs action.</i>
9. Constitutive / Anticipatory use Vedung (2015)	<i>An evaluation process spurs action and has impacts already in its initial phases by making evaluands and stakeholders aware of its coming into force.</i>
10. Tactical use Vedung (2015, 1997)	<i>Evaluations are conducted in order to postpone decisions, by showcasing that there is an ongoing evaluation. It is the process rather than the findings that is at the focal point of use.</i>
11. Unintended use Patton (2008)	<i>Evaluation findings are used outside of an evaluated program; e.g., by spurring other investigations concerning issues separate from the program.</i>

3. Research methodology

3.1 Research positioning and design

3.1.1 An interdisciplinary approach

As this thesis is predominantly oriented towards practical application rather than disciplinary advancements (*Mode 2*), and seeks to provide a broad approach for addressing complexities related to supporting an energy transition through evaluation, an interdisciplinary approach was deemed appropriate. This approach was operationalized by combining insights from four disciplines in an effort to promote transformative evaluation.

Interdisciplinary research is characterized by approaching a research problem that is too complex or too expansive to be addressed by only one discipline or one method (Klein, 1990; Pye, 2018), by applying multiple disciplinary perspectives, skills, methods and theoretical frameworks (Aboelela et al., 2007; Klein, 2017). This is notably the case in areas pertaining to sustainability, where problems and issues are clearly related to many different disciplines, including natural sciences, social sciences, and humanities (Frodeman, 2014; Klein, 2004). In interdisciplinary research, insights from various knowledge areas are *combined* and *integrated* to support comprehensive outcomes, and thus a key feature is to transcend disciplines to achieve synergies and a more holistic approach (Bruce et al., 2004; Siedlok and Hibbert, 2014).

As outlined before, supporting a transition towards a more sustainable energy system and society faces complex challenges that span the entire system, including its institutions, norms, and stakeholders. Thus, seeking to understand how transformative evaluation can be viably achieved is complex and arguably requires insights from multiple disciplines. However, bearing in mind that *Mode 1* and *Mode 2* research is not necessarily mutually excluding (Gibbons et al., 2011, 2010), the interdisciplinary approach taken in this thesis also leverages the opportunity to explore how key insights from the selected disciplines can support the specific context of transformative evaluation, and thus also expand into areas of theoretical application.

In terms of selecting relevant disciplines to consider, *Mode 2* interdisciplinary research sees that the research task largely guides this delimitation (Bruce et al., 2004). Thus,

four key disciplines are applied in this work, as outlined in Section 2. While these four disciplines form the theoretical boundaries of this thesis, there are clearly several other scientific disciplines that may be considered as well for the undertaking of transformative evaluation, depending on the selected context and focal areas. Examples of such include behavioural studies, e.g. environmental psychology or other knowledge areas that revolve around behaviour and psychology for assessing the role of and implications for the individual in an energy transition, or organizational studies and institutionalization for capturing the processes behind learning and adoption of practices, to mention a few.

There are always trade-offs to consider when deciding upon research approaches, and one main challenge associated with an interdisciplinary approach is the balance between breadth and depth. In this thesis, the incorporation of four main theoretical areas came at a cost in terms of the level of detail that could be accounted for, while keeping a manageable and useful approach. Consequently, a single disciplinary approach, or reducing the number of disciplines considered, would have provided opportunities to go into more depth in each. The importance of this alternative was not disregarded or dismissed, but as the nature of the research project Transition Governance was oriented around taking a systemic and wider grasp of how to design and promote transformative evaluation, the foundation of drawing from multiple knowledge areas was deemed to strengthen the research design. The roles of each discipline in this endeavour were to provide insights that could be used to support a proposed transformative evaluation approach, and which could be synthesized into comprehensive frameworks for concretely guiding transformative evaluation in a systemic perspective. Thus, the balance between breadth and depth was guided by the intention to acquire and present an overview of key insights from each discipline that could be relevant for advancing transformative evaluation in a manner that was broad enough to encompass and account for the richness of aspects to be considered, and yet detailed to the level that each aspect could be concretely operationalized and applied in transformative evaluation.

3.1.2 Research positioning

Inevitably, the interdisciplinary approach taken in this thesis means navigating multiple ontologies and epistemologies. As all selected disciplines that delimit this thesis in themselves represent multiple approaches to viewing research and policy incentives in relation to society and transformational change, some discussion is warranted on how this research has been conducted with regard to ontological and epistemological standpoints.

This work is guided by a critical realist stance, in brief meaning that the research adopts the ontology of there being a reality that is independent of our interpretation of it

(Cruickshank, 2003), and that our understanding of it is shaped by how we approach and interpret the relationship between structures of this reality and the events that it forms (Bhaskar, 2010). Thus, critical realism is related to the post-positivist schools of thought, in highlighting that underlying theories or knowledge may affect the perception of the evidence production (Chalmers, 2013). Critical realism is, moreover, suitable for taking an interdisciplinary approach, as it acknowledges the need for multiple viewpoints when seeking to gain an understanding of complex issues, such as sustainability or other system-wide aspects such as transitions (Bhaskar et al., 2010).

Epistemologically, adopting a critical realist stance means that this research acknowledges that the choices of disciplines to be considered, along with the researcher's previous knowledge, can influence the research output. Measures that are taken in this research to provide insights that build on validity include transparency about the thesis's presuppositions and an account of the theoretical underpinnings of the frameworks used, as well as inclusive and comprehensive data collection and triangulation of methods.

3.1.3 Research design

The overall research design in this thesis is grounded in a case study focusing on how transformative evaluation can be advanced. The case study builds on generating insights from current Swedish evaluation practices for research and policy incentives for energy efficiency in buildings, which guide suggestions on how to move towards evaluation approaches that are transformative, rigorous and knowledge-based. By and large, case studies seek to take a more holistic approach, rather than a reductionist one (Verschuren et al., 2010, p. 178), which in the context of this research enables learning about how evaluations are conducted and used broadly, while also laying the foundations for understanding the implications for state agencies in adopting a transformative evaluation approach.

The case study applied in this research is oriented towards the exploratory kind (Yin, 2009), intended to allow broad insights into an issue for which relevant boundaries and characteristics are still largely open and undefined (Swanborn, 2010). This approach was chosen over other kinds of case study designs, such as descriptive or explanatory, since the research sought to provide an understanding of an issue that was guided by a direction and a rationale (advancing evaluation for a transition towards a more sustainable energy system), rather than actual events. As the research seeks to uncover the opportunities for evaluations to support a transition, the outlined case is not oriented around establishing causal relationships (as would be an explanatory case), or describing well-defined events over time (as in a descriptive case) (Yin, 2009).

A benefit of using a case study is that participants can be involved in the research, and as such, the relatability of the research output to stakeholders is supported, as results

are grounded in perceptions and understandings of these actors (Simons, 2009). This is an important aspect for the overall objective of the research, since establishing a relationship and dialogue with key practitioners working with evaluation is essential to proposing an evaluation approach that is perceived as feasible and realistic. In this research, stakeholders from Swedish state agencies were involved in focus groups and completed a structured interview questionnaire to provide insights about how transformative evaluation and current evaluation practices were perceived.

The limitations of case study research designs can concern bias of the researcher, and the validity of findings (Simons, 2009). The subjectivity of the researcher is addressed in the section about research positionality, and is addressed by seeking to keep methodologies and interpretation of data transparent. In terms of internal and external validity of the case study, issues surrounding internal validity can be disregarded as the nature of this case study design is exploratory and does not seek to establish any causal relationships (Yin, 2009). As for external validity and the ability to make generalizations of findings applicable to instances outside of the case study, the specific insights gained are bound to the specific context of Swedish evaluation practices for research and policy incentives related to energy efficiency in buildings. This means, among other things, that specific areas suggested for further complementing current evaluation practices may not be applicable to other cases. Nevertheless, case studies can provide overall additions to the pool of knowledge in an area, where learning drawn from specific cases can be valuable in and of itself (Flyvbjerg, 2006). Case study results can also assist in refining how a phenomenon is understood, through ‘modified generalizations’, where new pieces of knowledge alter the previous understanding (Stake, 1995). Such contributions from this research may concern insights regarding how evaluation methods are operationalized in a certain context, and what they are – and may – contribute with in terms of providing transformative insights. Thus, it can be argued that the empirical findings from the case study at hand, in combination with the provided theoretical approaches that assist analysis, can be useful to other cases provided that the specific contextual differences and the characteristics of the studied case are heeded (see Yin, 2009).

More concretely, for this research it is thus suggested that insights gained from the Swedish evaluation practices in the field of energy efficiency in buildings, can be used as sounding boards in other evaluation contexts that are geographically or sectorally different, provided the generic structures or other characteristics relating to evaluation are similar. This argument is, moreover, arguably true for evaluation practices at large, since each evaluation of a research or policy programme will always be unique in terms of the evaluand and the context in which it is operating. Thus, there is no generic formula for how to conduct an evaluation, but instead a range of approaches and theories exist that can be applied onto multiple types of evaluands.

3.2 Methods for data collection and analysis

Methods for data collection and analysis were selected according to the research objectives of each paper, which were guided by the overarching research questions (see Figure 5). The following section describes the methods applied and provides reflections on their justifications.

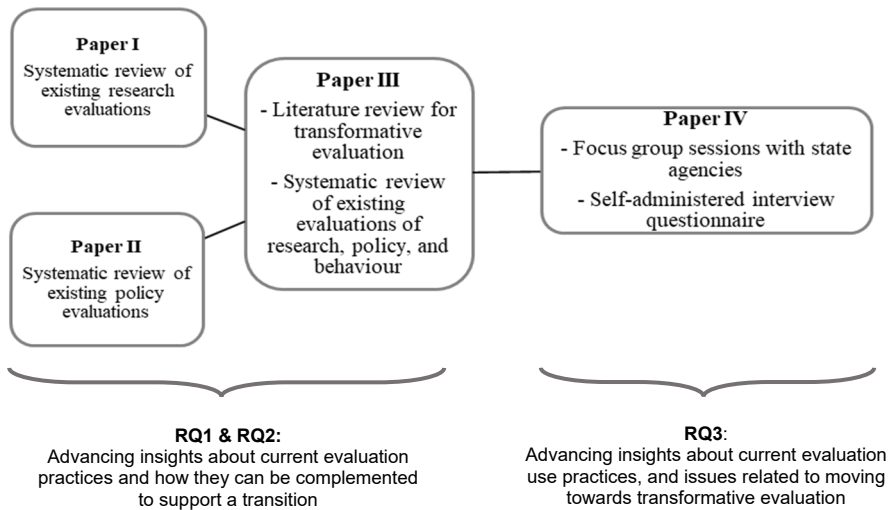


Figure 5. An overview of the different methods applied in each paper and their relation to overarching research questions.

3.2.1 Document studies and systematic reviews

Document studies of existing evaluation reports formed the empirical foundations of Papers I-III. In addressing research questions 1 and 2, regarding the evaluations' capturing of transformative efforts and how to further complement the evaluation practices, document studies were seen as an appropriate means to gain an understanding of how evaluations were performed.

The reviews thus covered existing Swedish evaluations of research programmes/institutions and policy instruments for energy efficiency in buildings, which were commissioned by Swedish state agencies or funding agencies. The identification and collection of potential evaluations was carried out through a close reading of Sweden's National Communications on Climate Change to UNFCCC (Ministry of the Environment, Sweden, 2014, 2009), through interpersonal contacts with representatives at relevant state authorities and actors conducting evaluations (including The Swedish Energy Agency, the National Board of Housing, Building and

Planning, Formas, IQ Samhällsbyggnad, Mistra, and the Swedish Environmental Protection Agency), through web searches using search engines and the authorities' webpages, and through snowballing approaches from acquired material.

Scoping and delimitation of which evaluations were to be included in the review was guided by the following criteria:

- The evaluand was relevant for supporting energy efficiency in buildings.
- The evaluation was commissioned by governmental authorities or funding/steering organizations.
- Commissioning date of the evaluation (2005-2017 for Paper I; 2005-2015 for Paper II).

The evaluations that met the criteria amounted to 20 research evaluations, and 33 policy instrument evaluations. These were all closely read, and data was extracted qualitatively and structured according to the theory-based specific frameworks described in Section 2.5.

A content analysis approach was applied for analysing the data generated from the reviewed evaluations. The frameworks were thus constructed to guide a systematic review of each reviewed evaluation (Littell et al., 2008), which allowed a qualitative categorization of the content and which also provided a quantitative overview of the sample in accordance to the sub-categories of the frameworks (Bryman, 2012). One potential drawback of this method concerns the representativeness of documents used in the review, as the robustness of the method is dependent on the documents reviewed. However, as the review was exhaustive in that it entailed all evaluations identified nationwide that met the outlined criteria, the representability of the sample was deemed reliable as far as possible.

3.2.2 Focus group sessions

Addressing research question 3, Paper IV aimed to uncover benefits and challenges connected to moving towards transformative evaluation in practice. In order to acquire an overarching and initial understanding of how Swedish state agencies working with evaluations perceive transformative evaluation, focus group sessions were deemed suitable. Focus group interviews are typically arranged to discuss a defined topic in order to acquire insights from participants in groups, rather than as individuals. They typically consist of people with specific experiences or training, and in the group discussion certain core issues and consensus can be identified (Bryman, 2012, p. 501).

The selection of focus groups was delimited to state agencies that were working with evaluation of energy-related research and policy instruments, by either planning,

commissioning or conducting evaluations. Based on this, requests were made to five¹ agencies to participate, of which four agreed: the Swedish Energy Agency, the Swedish Environmental Protection Agency, Vinnova, and Growth Analysis.

The focus group sessions were arranged by the project researchers Sofie Sandin and Lena Neij, and were joined by project members Maria Johansson at the Swedish Environmental Protection Agency, and Per Mickwitz at Vinnova. They were framed as workshops, where researchers reached out to a representative at each state agency with the inquiry for the agency to participate and share their experiences about evaluation, as well as their thoughts about moving towards transformative evaluation. Invitations to join were extended to co-workers by the agencies themselves. The sessions primarily targeted those working with planning/commissioning/conducting evaluation of energy-related research or policy instruments, but participant selection was left to the discretion of each department in order to maintain a session that was attractive to the agencies. The sessions were held separately at each agency in April and May 2019. Each session lasted for approximately two hours, and the number of participants ranged from three to approximately fifteen. Attendees were able to leave the session at any time, and there were instances in which some attendees were unable to join for the full duration of the workshop due to other duties; hence the approximate number of attendees.

Data was collected by Sofie Sandin and Lena Neij, who took notes during the sessions, as the sessions were not recorded in order to support an informal environment where discussion would be as open as possible. The notes were compiled after each session and were then analysed and categorized qualitatively in an inductive and exploratory manner to gain an understanding of the perceptions among state departments with regard to possibilities for adopting a broader transformative evaluation approach.

The limitations of the method relate in part to the agencies that agreed to participate in the focus group sessions, as the representation of agencies determined the scope in terms of who were able to share insights. Limitations also relate to taking notes while in the session and then interpreting them accurately afterwards. These limitations were addressed by transparency concerning which state agencies participated in the focus group sessions, and by having two project researchers taking notes during the meetings. When interpreting the notes, they were grouped and categorized to ensure that core ideas remained accounted for and intact.

Another limitation of focus groups is the potential risk of participants withholding or altering their views on a topic due to group dynamics and social relationships and expectations, or due to the lack of anonymity (Bryman, 2012). Since these sessions were located at the participants' workplaces and held with their colleagues, there may have been social structures and hierarchical aspects that influenced how participants shared their thoughts and insights. The focus group method was nevertheless prioritized over

¹ The National Board of Housing, Building and Planning (Boverket) did not participate.

individual interviews since the main goal at this stage was to acquire insights that could provide an initial scoping about perceptions of transformative evaluation, rather than detailed accounts. Another rationale was that the format of a joint session at the workplace would be more attractive to participate in for the state agencies, as it also provided an opportunity for an inclusive and creative discussion within the department about evaluation, as well as learning for the researchers about the agencies' work with evaluation. In order to balance potential limitations in the focus group sessions, insights from these focus groups were later coupled with a structured interview questionnaire that was to be filled in anonymously and individually by state agency representatives.

3.2.3 Structured interview questionnaire

In accordance with research question 3, Paper IV also aimed to shed light on the current processes of using evaluations at state agencies. A structured interview questionnaire was used for this purpose, the reasoning for the choice of method being twofold: in part to complement the focus group sessions and acquire deeper information about how individual employees at Swedish state agencies who are working with evaluation perceive processes and issues related to evaluation use at their departments; and in part to triangulate findings from the focus group sessions and to illuminate any misleading results drawn from the focus group sessions, either from the researchers' notes or stemming from group dynamics in the focus groups (Bryman, 2012).

The questionnaire was guided by insights from evaluation theory pertaining to the use of evaluations, in terms of how to support further use and in terms of different models of use, as described in Sections 2.2.3 and 2.5.3. The questionnaire was designed as a self-administered online questionnaire and contained 25 questions (one question was conditional, however, and was only posed if the informant had answered *yes* to the previous question). The response formats were of different kinds, consisting of open-ended, multiple choice, and Likert scales (see full questionnaire in Appendix to Paper IV).

The selection of state agencies to receive the questionnaire was based on their working with commissioning and/or conducting evaluation of research or policy related to the areas of climate and/or energy. The scoping of relevant state agencies was thus extended from the focus group sessions to enable a wider sample of informants. Seven state departments were identified as relevant: the Swedish Energy Agency, the Swedish Environmental Protection Agency, Vinnova, Growth Analysis, The Swedish Agency for Economic and Regional Growth, Formas, and the National Board of Housing, Building and Planning. The population was thus rather limited from the beginning, which consequently led to a limited number of responses. This was, however, deemed a necessary trade-off in order to secure relevant responses on use processes for evaluation

within the specified fields of climate and energy-related research and policy instruments.

The request to participate in the questionnaire was made via e-mail to one representative at each state agency, who was asked to disseminate the request to relevant colleagues who were involved in the agency's work with evaluation of research and policy incentives. The request included a description about the questionnaire with regard to its background, objective and scope, as well as a link to where the questionnaire could be accessed. This approach to dissemination was deemed appropriate in order to secure anonymity as far as possible, as the researcher thus had minimal contact with potential informants.

Consent to participate was requested digitally upon accessing the questionnaire from each informant who were asked to read a more detailed account of the study and then actively select to proceed to the questionnaire. The questionnaire was open from 18 October to 15 November 2019. Data was compiled in an Excel sheet upon its closing and analysed using a mainly qualitative approach.

3.2.4 Ethical considerations

The study presented in Paper IV included data collected from representatives working at state agencies. The topic concerned perceived evaluation practices and initial discussions concerning the application of transformative evaluation, and was thus not of a sensitive nature and therefore no approval was needed from the Ethical Review Board. Participation in both the focus group sessions and the self-administered questionnaire was voluntary and participants could cancel at any time. No compensation was provided for participation.

Requests about the arrangement of focus group sessions were made by the researchers to a representative at each state agency, who then organized the event and invited relevant co-workers to attend. The researchers were mindful that these sessions were requesting time from employees, and the format was therefore largely flexible to suit the time availability and expectation of each state agency. The objective of the focus group sessions was to acquire an initial understanding of perceptions about transformative evaluation; thus, the format was also highly informal and no details about the attendees were required. The sessions were not recorded, and any potentially identifying traits of participants collected in relation to individual statements were removed at the analysis phase. The data from all four focus groups were aggregated to further remove connections of statements to particular state agencies.

The self-administered questionnaire was filled in individually online, and was introduced with a page about the study that accounted for background, scope, selection of respondents and treatment of data. The informants were then asked to actively agree to participate in the study, in order for the study to begin. The questionnaire was

anonymous to the extent that it did not require informants to disclose their name, and the request to participate was disseminated by a representative at each authority in order to prevent the researcher accessing information about potential informants. However, since the study aimed to reach a fairly small population of informants, scoped to representatives at state agencies working with evaluation of climate or energy-related incentives, the study was reported to the *Personal Data Lund University (PULU)* register. The analysis of the material was performed on an aggregate level, where connections between individual comments and responses and particular traits of the informants were not presented, in order to further secure anonymity for informants.

4. Key findings organized by papers

Four papers constitute this thesis. They take empirical areas of focus, seeking to shed light on current evaluation practices and use processes, with regard to how these can be advantageously complemented to support assessments of transformative contributions, and with regard to benefits and challenges related to adopting a transformative evaluation approach among Swedish state agencies. The following sub-sections present the key findings from each paper, structured around the main categories of *objective and framework*, *results*, and *conclusions*.

4.1 Paper I: Research evaluations for an energy transition? Insights from a review of Swedish research evaluation reports

4.1.1 Objective and framework

Understanding how to move towards transformative evaluation approaches in practice requires an understanding of how evaluations are currently conducted. Paper I looked at existing evaluations of research programmes and institutions for energy efficiency and building-related research in Sweden, and focused on ascertaining how the knowledge created in the evaluations relates to an energy transition.

Evaluations of energy-related research in Sweden are conducted in relation to the articulated national goal of seeing Swedish research and development (R&D) support a transition to a sustainable and secure energy system (Government bills 2012/13:21 and 2016/17:66). Against this backdrop, the ability of research evaluations to actually deliver relevant knowledge about the evaluand's ability to spur or support a transition is crucial.

An interdisciplinary framework was developed to articulate key areas of research evaluations in relation to an energy transition, which was used to guide a review of

existing evaluation reports (see Section 2.5.1 and Appendix A). The number of reviewed evaluation reports amounted to a total of 20, all of which were commissioned and/or conducted by Swedish state agencies or funding/steering organizations between 2005 and 2017. The evaluations covered both individual research programmes and institutions that engage in energy-related research.

The contributions of the paper were twofold: an empirical review of current evaluation practices concerning R&D for energy efficiency in Sweden, and a theoretical contribution through the development and application of an interdisciplinary framework aimed at guiding transformative evaluations and articulating how the knowledge produced in research evaluations relates to an energy transition.

4.1.2 Results

Core strengths identified in the reviewed material that were in line with capturing transformative efforts included a frequent operationalization of the evaluand as forming part of a societal vision, and articulation of the system in which the evaluand was found with regard to governance areas and other factors such as funders and stakeholders that may influence the research and its outcomes. However, the analysis of the evaluand's functions and outcomes were predominantly oriented on a programme level, focusing on e.g. the processes, goal attainment, and quality of the research programme, while not explicitly assessing the implications of effects on a societal scale. Notions of experimentation and learning were seen mainly as expressed through evaluations outlining learning and knowledge dissemination as an intended outcome of a research programme, and as a criterion for gauging programme achievements.

In terms of making rigorous and knowledge-based assessments, the review indicated a frequent practice of applying multiple criteria such as relevance, impact, and knowledge dissemination and learning, which paves the way for broad assessments of the evaluand. It is noteworthy that the most frequently applied criterion was relevance, which sought to establish the position of the evaluand in relation to societal needs – an essential feature for supporting transformative assessments and an energy transition. The methods used for analysis were mostly qualitative (mainly document analysis and interviews), but while multiple methods were employed in more than half of the reviewed evaluations, triangulation was scarce, as was the establishment of counterfactuals, which leads to concerns related to the level of reflexivity in challenging the role and contribution of the evaluand in transformative processes.

Lastly, the review showed that external actors representing academia and consultancies conducted the majority of the research evaluations. Actors who were involved in the evaluations included researchers and representatives from industry, companies, organizations, and authorities. The level of involvement was predominantly as informants for data collection in interviews.

4.1.3 Conclusions

The review showed that evaluation of energy-efficiency research is frequently undertaken in Sweden, and that the practices outline strengths to capitalize on, as well as areas to complement further to advance transformative evaluation.

In order to generate knowledge about how the evaluand supports an energy transition, there is a need to frame and analyse the evaluand as a part of realizing transformative goals. Evaluations can thus be complemented to provide transformative assessments by extending the operationalization of research programmes and institutions as parts of societal processes towards reaching overarching energy visions, and also towards *analysing* and *articulating* what the evaluand's transformative contributions are in relation to this vision. This includes taking a grander evaluation focus and scope that can account for systemic effects that the research spurs or yields, by making explicit how the evaluand is interconnected with other governance levels and actors, and assessing how the evaluand aligns with other implementations on a societal scale. Moreover, articulating how the knowledge produced in a research evaluation can be used to support a transition could further support the establishment of links between individual evaluations, and thus develop insights into how different incentives leverage each other in relation to transition processes.

Knowledge-based assessments can be further advanced by expanding the methodological toolbox towards an increased application of triangulation and the construction of counterfactuals that can support conclusions about transformative impacts and contributions of research initiatives.

4.2 Paper II: Transition governance for energy efficiency – insights from a systematic review of Swedish policy evaluation practices

4.2.1 Objective and framework

Similar to Paper I, Paper II aimed to shed light on current evaluation practices in order to outline key areas for moving towards transformative evaluation. The objective of Paper I was to acquire insights about the current evaluation practices for policy instruments for energy efficiency in buildings in Sweden, by looking at existing evaluation reports commissioned and/or conducted by state agencies and, based on these, identifying core areas to complement for advancing transformative evaluation.

An interdisciplinary framework was designed to guide the review of existing evaluations (see Section 2.5.2 and Appendix B). The number of reviewed evaluations amounted to a total of 33 evaluation reports of individual policy instruments conducted between 2005 and 2015. These covered a wide range of policy instruments: legislative, financial, informative, and other types of instruments such as technology procurement programmes and cooperative networking programmes, which provided a broad foundation for generating insights on evaluation practices within the sector.

While the paper's objective was geared towards generating empirical insights into current evaluation practice, it also articulated a theoretical contribution in terms of developing and applying an interdisciplinary framework for guiding evaluations towards a more transformative approach.

4.2.2 Results

A key message to bring forward is the fact that multiple evaluations have been conducted of individual policy instruments for energy efficiency in buildings in Sweden over the investigated period. These were predominantly conducted by external consultancies, but two state agencies also conducted evaluations. The overall scope of the evaluations showed a focus on the programme level, mainly oriented around programme goals, processes and impacts. Thus, there was generally little attention paid to side-effects and rebound effects.

In terms of methods and value judgements, the review showed that evaluations often did employ knowledge-based approaches such as multiple criteria and methods, yet without capitalizing on the opportunity to crosscheck results using triangulation. Methods used were mainly qualitative, building on interviews and document analysis, and the actors that were involved – mainly for data collection – included authorities, beneficiaries, and organizations. The most frequently applied criteria were effectiveness, impact, and instrumental feasibility and processes. One evaluation assessed the relevance of the programme. Application of counterfactuals was not a predominant practice for ascertaining impacts attributable to the evaluand.

Key drivers and processes outlined for capturing transformative efforts in evaluations, such as visioning, experimentation and learning, and systems and multi-actor perspectives, were heeded to varying extents. Visioning was seen applied in terms of relating the evaluand to overarching energy goals or by combining retrospective and prospective assessments, and experimentation and learning were expressed in some evaluations, emphasizing both experimental contributions as well as failures to support new practices. However, as these aspects were not central to the evaluations at large, the connections between transformative contributions and effects from experimentation on a societal or visionary level were not thoroughly assessed.

Lastly, in terms of use, it was clear that the intended use was predominantly programme-oriented, generally either for revising a current programme, or for determining final impacts and outcomes.

4.2.3 Conclusions

The review shows that Sweden has a sound and established evaluation practice in place for policy instruments related to energy efficiency in buildings, which can be further complemented to advance transformative and knowledge-based evaluation approaches.

Firstly, while the reviewed evaluations did not showcase clearly stated objectives to support transformative assessments, key components outlined for transformative evaluation were still applied to varying degrees. In this regard, evaluations can be complemented to take a more inclusive systems perspective, e.g. by moving away from the emphasis on criteria concerning effectiveness and impacts, which largely gears the evaluation focus towards a programme level, and towards criteria that can support insights on how the evaluand may support transformative change on a societal level, through e.g. relevance, acceptability and predictability. A more deliberate and thorough multi-actor approach is also encouraged, one that involves actors in the evaluation processes and places focus on assessing how an evaluand supports visions and spurs transformative effects.

Secondly, a knowledge-based approach can be further supported by triangulating findings in order to verify them and to gain additional perspectives, and by applying multi-criteria analyses that allow a focus beyond the programme boundaries. Lastly, counterfactual analyses and constructions can add further rigour to impact assessments, as attributable effects are analysed in further depth.

4.3 Paper III: Bolstering a transition for a more sustainable energy system: A transformative approach to evaluations of energy efficiency in buildings

4.3.1 Objective and framework

Building on insights from Papers I and II, Paper III set out to provide the foundations for supporting governance towards a transition to more energy-efficient buildings, by proposing an evaluation approach that deliberately seeks to design and combine evaluations with a transformative focus. As such, this paper was designed as an opinion piece, yet equipped with an empirical part as described below.

Firstly, the paper described and suggested how to develop and move towards a transformative and interdisciplinary evaluation approach. The suggested evaluation approach is built on two key components: (I) it is anchored in transition research and incorporates key aspects for capturing and assessing transformative efforts of an evaluand, and (II) it is rigorous and knowledge-based, i.e. it promotes that evaluations should be grounded in relevant theories and that key aspects from these should be applied to secure robust and reliable assessments of the contribution of the evaluand.

Thus, the paper presented a literature review aimed at providing the theoretical foundations of a transformative evaluation approach, covering knowledge areas that were suggested to be of importance in the case of promoting energy efficiency in buildings. These were transition research, evaluation theory, sociology of science, policy analysis, and behavioural sciences.

Secondly, based on the identified theoretical key aspects for moving towards transformative evaluations, the paper also provided some empirical insights on the performance of a total of 79 existing evaluations of research, policy instruments, and behavioural aspects, conducted in Sweden from 1997 to 2018, with regard to their ability to make transformative and knowledge-based assessments.

4.3.2 Results

The results from the literature review identified key aspects to consider for moving towards a transformative evaluation approach. From transition theory, three drivers that are essential for supporting a transition were highlighted: *visioning* as regards what the long-term goals are and how to achieve them; *experimentation* for supporting innovative efforts and risk-taking that may affect current practices; and *learning* for a deliberate and well-informed improvement of the designs and implementations of research and policy incentives. Additional factors that were identified as important to consider were systems and multi-actor perspectives, and acknowledgements of lock-in effects and path dependences, which hinder reconfigurations of socio-technical regimes.

The review of essential knowledge areas for supporting knowledge-based evaluation practices outlined key aspects for promoting robust evaluations, such as the application of multiple methods for securing triangulation, the use of counterfactuals for supporting robust impact assessments, and the application of multiple criteria for supporting versatile and more inclusive assessment and judgement in evaluations.

The review of existing evaluations showed that current evaluation practices in Sweden were to a large degree building on outlined key aspects for supporting a transformative evaluation approach, yet without a deliberate focus to make assessments of the transformative contributions of the evaluand. In terms of securing knowledge-based evaluations, which build on robust methods and approaches, triangulation and a more

thorough application and construction of counterfactual cases could be more frequently applied to further strengthen current evaluation practices.

4.3.3 Conclusions

A key conclusion for moving towards transformative evaluations lies in the potential to further establish an evaluation practice that builds on interdisciplinary approaches, and that deliberately plans how evaluations can support each other in providing a more holistic perspective on transformative change spurred by various research and policy instruments. This can be done by articulating a joint vision for energy efficiency in buildings to be used as a benchmark for evaluations, by outlining a plan for the key aspects to be considered in an evaluation of implementations, and by increasing the coordination between evaluations.

4.4 Paper IV: Making use of evaluations to support a transition towards a more sustainable energy system and society — An assessment of current and potential use among Swedish state agencies

4.4.1 Objective and framework

Building on the outlined characteristics of a transformative evaluation approach outlined in Paper III, Paper IV sought to further generate knowledge on the possibilities for moving towards evaluation practices that support a transition towards a more sustainable energy system and society. If transformative evaluation approaches are to indeed be made a viable means to support governance for a sustainability transition, there is a need to understand how state agencies, with mandates to support transformative research and policy, perceive a potential adoption of such an evaluation approach. There is also a need to understand how evaluations are currently used, and whether the processes for supporting an increased use can be further strengthened.

The objective of Paper IV was thus to shed light on how representatives working with evaluation at relevant Swedish state agencies perceived the adoption of a transformative evaluation approach with regard to benefits and challenges, and how they perceived current evaluation use practices.

Focus group sessions were arranged with four state agencies (the Swedish Energy Agency, the Swedish Environmental Protection Agency, Vinnova and the Swedish Agency for Growth Policy Analysis), and the transformative evaluation approach was

introduced and discussed in groups at these sessions. This was coupled with a structured interview questionnaire sent to five state agencies (the four abovementioned, and the Swedish Agency for Economic and Regional Growth) in order to further probe the processes underpinning use of evaluations, and to triangulate the perceptions of adopting a transformative evaluation approach.

4.4.2 Results

The results outline some key benefits and challenges related to adopting a transformative evaluation approach, as perceived by representatives at the state agencies. Benefits included that a transformative evaluation approach can bring increased collaboration and a broader application of evaluation results, whereas challenges pertained to current structures for commissioning and conducting evaluations, including e.g. agency, ownership and mandate to make executive decisions, and acting on evaluations.

As for current practices for using evaluations, the results showed a predominant perception of use revolving around instrumental use and enlightenment, and that less proactive models of use, such as ritual and tactical, were less agreed with. There was also a seemingly solid foundation for facilitating use, which built on the involvement of different actors in the evaluation process, as well as a wide range of channels used for communication.

4.4.3 Conclusions

Although the results showed some promising foundations for evaluation use, there were still some concerns expressed among informants that evaluations are not used to their full potential. Thus, identifying good practices, such as collaboration and well-designed communication through various channels, and capitalizing on these, can extend the use of evaluations.

Informants at the Swedish state agencies outlined that a transformative evaluation approach can be a means to support a more holistic perspective of transformative efforts, but that there are structural and institutional challenges to overcome in order to adopt such an approach. Overall, models of use that were perceived as commonly recognized at the state agencies were largely suitable for moving towards a transformative evaluation approach.

5. Concluding discussion

The overall objective of this thesis has been to advance knowledge about how to support a transition towards a more sustainable energy system and society, by supporting evaluations of research and policy incentives to make assessments about transformative efforts. As has been outlined throughout this thesis, a transition comes with many dimensions that are interconnected. Addressing this complexity in evaluations therefore requires varied perspectives that target and account for different levels: from the individual policy incentive to the more holistic system level. An important note to make is that there is valuable knowledge generated through the collected evaluation activities undertaken in Sweden, but to further capitalize on the potential of these insights, there is a need to make connections between the knowledge created on the programme level, and its relation to a systems perspective.

A key argument of this thesis is that transformative evaluation advantageously departs from transition research, which provides conceptualizations of transformative processes but does not provide broader guidance for how to evaluate or gauge transformative efforts from research and policy incentives. Thus, insights from other disciplines have been synthesized in order to provide tools and approaches for moving towards transformative evaluation approaches that support knowledge-based assessments of contributions from research and policy incentives. Together, the different disciplines presented in this thesis highlight some key aspects that can advantageously be applied to complement current evaluation practices.

This section provides some concluding reflections and discussions of the contributions of this thesis, with regard to empirical findings about current Swedish evaluation practices and how to further complement these along the lines of transformative evaluation; and with regard to theoretical contributions related to evaluation with a transition in the focal point. As these results are grounded in the case study of Swedish evaluation practices, it needs to be remembered that the suggestions are context-specific in many respects. Nevertheless, the insights generated from the Swedish case may still outline guidelines for how to also complement evaluation practices in other contexts and countries.

5.1 Empirical contributions

The empirical contributions made in this research are oriented around the three research questions posed in the beginning of the thesis. The reviews of current research and policy evaluation practices in Sweden are thus used as a point of departure in this research in order to outline areas where efforts can be directed to move towards transformative evaluation. These aspects will be elaborated upon in addressing the research questions one by one.

5.1.1 Research question 1

Are current evaluation practices for research and policy instruments, within the field of energy efficiency in buildings, allowing a capturing of transformative contributions of the evaluand?

In response to the first research question, Papers I and II found that evaluation practices applied in 53 existing evaluations commissioned by Swedish state agencies and funding/steering organizations are not currently oriented towards generating insights about transformative contributions of the evaluand. However, there is a solid foundation in current evaluation practices that can advantageously be further developed in terms of gearing focus towards transformative aspects, and by promoting transformative and rigorous, knowledge-based approaches. These are explored further below, oriented around three main areas of discussion: the focus of evaluations, the foundations for capturing transformative efforts, and the foundations for knowledge-based approaches.

The focus of evaluations

An important first note to make is that the reviews of existing evaluations in Papers I and II mainly covered individual research programmes or policy instruments, and, as may have been expected, these evaluations predominantly took a programme focus, which meant that they were not prioritizing knowledge production concerning aspects related to effects beyond the evaluand. Looking particularly at energy efficiency in buildings, the reviewed evaluations showed a variety of incentives for promoting energy efficiency through e.g. development of new technology, provision of avenues for knowledge exchange between actors, and financial subsidies. The evaluations of these incentives often pointed to successful implementations in terms of e.g. saved kilowatt-hours and awareness among stakeholders, but the transformative effects of these different efforts were largely not addressed.

There are certain key conclusions to draw from this. One is that on one hand, programme evaluations are delivering important knowledge that may be used on the programme level, e.g. for instrumental improvement or accountability checks. Another

conclusion, however, is that these assessments also can provide insights about how outcomes or impacts of the programme are (potentially) influencing other system levels, other implementations and programmes, and stakeholders beyond the programme. While one research or policy incentive may not bring about a transition on its own, it may be a crucial component that either yields or supports a transformative outcome that in turn supports another programme to provide transformative effects. This type of cascading effect may prove to either support or counteract a transition, and thus needs to be considered from the perspectives of the envisioned transition and what role the evaluand and its effects have in relation to it.

Foundations for capturing transformative efforts

Drawing from transition research, the identified drivers and components of a transition were visioning; experimentation and learning; and systems, scale and multi-actor perspectives. Arguably, if an evaluation heeds these aspects, it is better suited to provide a solid foundation for assessments of transformative effects of an evaluand. Although the majority of the reviewed evaluations were focused on the individual programme level with few or no outlooks to the wider system level, the results still – rather unexpectedly – revealed that these key concepts were indeed applied in the evaluations to varying degrees, which provides fertile grounds for further expansion.

Visioning was often apparent in the reviewed evaluations as a means to place the evaluand in relation to societal goals, as was seen in Papers I-III, but the concrete links between the evaluand and the societal vision were disrupted by the focus on programme performance and effects. Assessments of experimentation and learning, in terms of developing, testing or demonstrating e.g. new technology and practices, were treated differently in different research and policy instrument evaluations. As outlined in Papers I and III, notions of experimentation and learning were in large part seen as inherent to research programmes and institutions, but, as seen in Papers II and III, were less pronounced in policy evaluations. This relates in part to the nature of the evaluand with regard to how prone it is to support experimentation, acknowledging that different research and policy instruments have different objectives, but it also relates to the focus of the evaluations and how they were set up to assess the evaluand. Arguably, due to the predominant programme focus in the reviewed evaluations, outlooks to experimental efforts with potential transformative implications may have been parsimonious as the evaluations assessed other programme-related aspects that were not geared towards experimentation and learning. For example, if an evaluation focused on goal attainment and effectiveness, and the goals were not oriented around experimentation, such efforts are likely to have been either disregarded, or mentioned only superficially. Nevertheless, there were outcomes supporting energy efficiency of various kinds outlined in the reviewed evaluations that may support transformations. Some examples include new development and adoption of technologies, such as

lighting products or heat pumps, and knowledge transfer and best practices facilitated through e.g. cooperative network programmes.

Evaluations of both research and policy instruments often engaged different stakeholders to acquire insights about the evaluand from certain perspectives. While the level of involvement was often limited to data collection with a programme focus, it shows a foundation for deeper assessments of the roles of these actors in terms of promoting change and transformation. As has already been outlined, a thorough systems perspective was not commonly employed in current evaluations – given the programme focus – but both research evaluations and policy evaluations did to varying extents rely on interactions with, and data collection from, different system actors. These actors included government representatives, programme leaders, funders, researchers, representatives from the industry and private sector organizations, as well as private citizens, to mention a few. Thus, if these evaluations would be seen as complementing each other, a more holistic picture would emerge concerning where in the system which outcomes and effects are taking place, which also could promote a more deliberate alignment between research and policy evaluations.

Foundations for knowledge-based approaches

In terms of providing rigorous knowledge and assessments, the reviewed evaluations showed a rather robust approach, but one that can advantageously be further advanced.

A key finding concerning both research and policy evaluations was the prominent use of multiple methods for analysis and data collection, which provides versatile and broad insights about the evaluand. While this in theory provides an opportunity to triangulate results to validate their accuracy, such triangulation was, however, uncommon among the reviewed evaluations.

Another general strength seen in both research and policy evaluations was the frequent use of multiple criteria, which allowed broader assessments of the different workings of the programmes: from the administrative processes, to the effects, to programme effectiveness. However, there was an emphasis on programme-related criteria such as impact and effectiveness, and other criteria that arguably would have been better able to mirror the system effects to a larger degree, e.g. relevance, equity, and legitimacy, were less frequent. One difference between research and policy evaluation in this regard was the rather frequent application of *relevance* as a key criterion in research evaluations, which was significantly less frequent in policy evaluations.

An important discussion point before closing the discussion on research question 1 is the balance that individual evaluations must strike between providing essential insights for programme development, accountability and improvement, and the potential they hold for providing insights about transformative contributions beyond the evaluand. Transformative evaluation does not suggest that each evaluation needs to address all aspects pertaining to both programme learning and transformative contributions.

Instead, evaluations should be planned to complement each other, and results can advantageously be *synthesized* in order to generate more holistic overviews, as outlined in Paper III. In terms of advancing and complementing evaluation practices towards making transformative assessments, suggestions as to where to direct efforts to address these issues are further discussed in relation to the second research question.

5.1.2 Research question 2

How do the evaluation practices for research and policy instruments need to be complemented to be able to support a transition towards a sustainable energy system and society?

Seeking to outline some key suggestions for how to complement Swedish evaluation practices to provide rigorous knowledge that is relevant for supporting a transition, insights from all four papers are synthesized. The suggestions made below depart from the empirical findings from research question 1 and are oriented around how to complement practices for a more transformative approach, and how to complement practices for a knowledge-based approach.

Complementing for a transformative approach

To start by picking up the discussion from research question 1, a key area for moving towards transformative evaluation is to make evaluations take a broader focus that allows outlooks on how outcomes and effects stemming from the evaluand affect a system level, particularly as seen in relation to other enforced or planned research and policy incentives. However, the prevalent focus on accountability and programme development seen in the reviewed evaluations leads to two largely separated tracks for research and policy instrument evaluation, with few connections made between them in the evaluations. This has the consequence that any combined effects or issues surrounding alignment and cascading of supporting effects between different incentives tend to be overlooked. While learning on a programme level is also clearly important and not to be disregarded in order to continuously sharpen programme designs, there is room to complement evaluations to consider how the evaluand depends on or supports other research or policy incentives in generating transformative contributions. Thus, a first measure to develop current evaluation practices towards supporting a transition is to allow evaluations to take a broader focus that reaches beyond the programme boundaries, and thus to support alignment between research and policy evaluations.

As for supporting alignment, there is a promising foundation to be noted. When summarizing the 53 reviewed evaluations of both research programmes and institutions, and policy instruments for energy efficiency in Sweden that inform the research in Papers I-III (see Appendix D), it becomes apparent that the actors who are

involved in the commissioning and conducting of evaluations of these incentives constitute a limited number of organizations. This shows that knowledge production for both research and policy incentives is in many cases procured by the same state agency, which clearly provides opportunities for an enhanced alignment of these evaluations. Such alignment can inform about evaluation practices, e.g. how to assess transformative contributions, it can support deliberate planning of different evaluation efforts to account for various parts of the system, and it can support a more holistic picture of the combined efforts from the various evaluated programmes and institutions if these are synthesized, as proposed in Paper III.

An increased alignment of evaluations could potentially also support a more frequent combination of prospective and retrospective evaluations, and thus support a reinforcing continuum of knowledge creation where insights and learning from ex-post assessments could support forecasting and ex-ante assessments, which in turn could guide decision-making for securing transformative outcomes and effects.

Another key aspect for supporting assessments of contributions towards a transition is to further make evaluations connect the evaluand to a societal goal or vision. The results from the reviewed evaluations in Papers I-III showed that the evaluand's relation to a societal vision was often mentioned, but assessments of contributions towards it were limited, as the focus was on the programme rather than the system level. Thus, the individual evaluation practices can be further complemented in seeking to assess how the effects or outcomes of the evaluand may lead to or support a vision by e.g. providing new technology, practices, regulations, funding, experimentation, etc. Moreover, an aligned evaluation approach, which also considers the timing of different evaluations, can advantageously support synthesized assessments of how evaluands together fare in relation to long-term visions for the energy system and society.

Furthermore, current Swedish evaluation practices can be complemented regarding the articulation of how the evaluand relates to experimentation and learning. As experimentation was not addressed as a core concept in the reviewed evaluations, there is much to be gained in terms of further complementing evaluations in this regard. In gauging the experimental efforts of an evaluand, it can be valuable to assess what kind of experimentation is provided, e.g. experimental outcomes in terms of new technology or practices, or experimental policy designs, and then to capitalize on the learning that can be drawn from the evaluand in terms of how to further advance energy efficiency. This can, moreover, advantageously be related to the overarching vision as well.

On this note, if evaluations are to be complemented in accounting for visioning and experimental efforts, it may also be valuable to critically seek to address any adverse transformative effects that are supported by the evaluand. These may be rooted in the upholding of unsustainable societal structures or practices, or in supporting a pathway that leads to new, unsustainable practices.

Complementing for a knowledge-based approach

Turning to guidance for how current evaluation practices can be complemented to generate knowledge-based learning that can support rigorous assessments of contributions from an evaluand, Papers I-III identify some key areas to focus on.

Firstly, current Swedish evaluation practices for energy efficiency in buildings do show a variety of methods used for data collection and analysis, which is good for generating versatile insights. However, the extent to which methods were used to ensure robust results through methods triangulation, rather than being used in parallel with each other, was infrequent and can thus be complemented – arguably without requiring a lot of additional efforts or resources. Triangulation can support reliability of the results, and may also be able to illuminate an issue from various angles and perspectives, which further may support a broader evaluation approach that assesses aspects based on various sources and methods for analysis.

Secondly, the reviewed evaluations also showed a variety of criteria used for assessing the evaluands: research evaluations showing an even distribution but with a slight emphasis on relevance and impact, and policy instrument evaluations with an emphasis on impact and effectiveness. Criteria concerning impact and effectiveness were expected from a programme accountability perspective, and are valuable for learning on a programme level, but in order to inform about the evaluand's performance on a systems level, these criteria need to be either put in relation to societal goals and needs, or be complemented with other criteria that can inform about transformative potentials. If evaluations are aligned to a larger extent, then insights regarding impacts and effects can be combined with other evaluation results, which can support an overarching view of transformative contributions.

Relevance is important for establishing a link between the programme objective and societal needs, and its application can be further extended, particularly in the policy evaluation practices, to secure assessments about whether the instrument is indeed appropriate for addressing societal needs. Other criteria to consider for a transformative assessment approach include – but are not limited to – *adequacy* to determine whether the instrument is adequate for reaching goals pertaining to supporting a transition; *legitimacy* and *acceptability* for assessing the extent to which transformative efforts and solutions are agreed upon and supported by stakeholders; and *equity* and *responsiveness* for ensuring that benefits and costs are shared equitably and that the outcomes that may support a transition are acceptable to different actor groups. Such criteria were largely absent throughout the reviewed evaluations, and thus illuminate an important area for further complementing current evaluation practices to enable them to become more transformative.

5.1.3 Research question 3

What are the current processes for use of evaluations among Swedish state agencies, and what are the major issues of concern among Swedish state agencies for adopting a transformative evaluation approach?

If the potential of evaluations to support a sustainability transition is to be harnessed, the learning and the knowledge generated in the evaluations need to be acted upon. This means that the evaluations have to be used, either during the evaluation process, or by using the final findings. In addressing the third research question, Paper IV found that the underpinning processes at Swedish state agencies for increasing the use of evaluation results are largely solid and correspond well to the suggestions made by evaluation scholars, e.g. in terms of using targeted and varied channels of communication, and in terms of inviting external actors to take part in conducting evaluations.

As evaluations can be used in various ways, the results generated regarding how informants at Swedish state agencies perceive current use further illuminate a promising foundation for moving towards transformative evaluation. Among the wide range of models of use that the informants were asked to rate, it was predominantly models of use that are deemed more appropriate for an enhanced use of evaluations (e.g. enlightenment and process use) that received high scores, whereas less-constructive models that tend to push hidden agendas or promote inaction (e.g. tactical use and overuse) received lower scores. This shows that there is a perception among the informants at Swedish state agencies that evaluations are conducted to be useful, and that there is a will to act and capitalize on the learning generated. The results also showed that the informants were not adverse to the possibilities to use evaluations more interactively and to have them form parts of larger pools of information for decision-making, or even to have them apply to areas outside of the evaluated programme itself (unintended use).

Moreover, the potential of evaluations to support transformative efforts was largely recognized among representatives from state agencies, which can be seen as a prerequisite for even considering moving towards transformative evaluation. However, while state agencies did see benefits associated with adopting the aforementioned transformative evaluation approach, such as increased use and cross-agency collaboration, they also expressed concern about potential challenges related to e.g. agency and ownership.

These concerns are clearly relevant. Even though this research has shown that current Swedish evaluation practices have good foundations to elaborate and build further on, it is also acknowledged that moving towards transformative evaluation does require a revision of the current practices, particularly concerning ownership, agency, and

supervision and overview of different evaluation efforts. It is also acknowledged that while increased collaboration between state agencies, academia, and private actors may be desirable in order to provide well-performed and targeted assessments of implementations, it also comes at a price in terms of time, money, and resources needed, e.g. for developing new routines and skills, and for maintaining efficient communication between actors. On this note, bearing in mind that the actual use of evaluations can be hindered or be sub-standard due to timing, resources, structural issues, etc. (see, for example, Chelimsky, 2015; Weiss, 1998), as was also outlined by informants in paper IV, there are certainly reasons to consider how to secure an enhanced and facilitated use of evaluations, and consider the benefits of this use in relation to the resources needed for realizing it.

As has been outlined before, a transition will entail deep socio-technical configurations, and is as such not an isolated event but something that will come about as the product of multiple interventions in multiple governance areas. Thus, if to make evaluations support a transition, there is a need to promote and secure knowledge transfer and use of evaluations, not only between evaluators and key stakeholders of the evaluand, but also between authorities and governance areas that are part of the change processes. While this research shows that state agencies outline some challenges in relation to such an undertaking, there are also acknowledgments of the benefits that could come from revising current evaluation practices and extending the collaboration with external parties, both from academia, private sector, and other public authorities.

5.2 Theoretical contributions

The theoretical contributions made in this thesis are rooted in the interdisciplinary research approach, which draws from four knowledge areas that were identified as important to support transformative evaluation of both research and policy incentives. The contributions to theory of this thesis are twofold: in part building on identification and suggestion of relevant aspects from these disciplines that can support transformative evaluation (as presented in section 2), and in part providing concrete suggestions on how to combine insights from these different knowledge areas in interdisciplinary frameworks to support evaluations to provide broader knowledge about how an evaluand performs in relation to a transition.

Starting with the identification of key aspects to consider when approaching evaluation from the perspectives of the four disciplines reviewed, a range of key concepts and aspects have been identified and highlighted. These cover overarching perceptions of systems, processes, and drivers pertaining to both transitions and research and policy incentives, as well as concrete aspects for how to capture these by considering various methods and value constructions. While it is acknowledged that each discipline certainly holds additional valuable input as well, the overall ideas presented from each

are intended to support transformative evaluation. In keeping a rather coarse level of aggregation, this approach seeks to provide outlines that can be applied in individual contexts, with room to consider how the presented overarching idea can be adapted to be relevant for the case at hand. Thus, in reviewing four disciplinary fields and outlining key ideas that may affect an evaluation focusing on a transition, this thesis provides a theoretical point of departure which can be adopted and adapted as is deemed necessary.

Turning to theoretical contributions stemming from the interdisciplinary approach taken, this research rests on a strong use of frameworks which are founded in different constellations of the four disciplines (see section 2.5), and these frameworks thus showcase how insights from various disciplines may support or strengthen each other to provide more holistic designs of evaluations. The functionality and usefulness of these frameworks have also been tested as they have been applied to guide analyses of evaluation practices in existing evaluation reports. This application has proven their worth in structuring evaluations around core ideas for transformative and knowledge-based insights, drawing from different disciplines.

Perhaps a more dynamic theoretical contribution is made in the interdisciplinary interstices between different disciplinary fields, where key ideas from the different disciplines can truly support one another in the pursuit of supporting a transition. One example concerns the key drivers for a transition - visioning, experimentation and learning - that can be used to guide an evaluation towards a focus on transformative contributions. These are crucial aspects for moving a transition ahead, but if to capture them in an evaluation and assess how a research or policy incentive is performing on these regards, they need to be operationalized in a concrete and manageable way. Since transition research in itself largely does not address issues surrounding how to couple evaluation and transition, this is where e.g. evaluation theory, sociology of science, and policy analysis make valuable contributions.

Some concrete examples will now be used to further elaborate this argument. Starting with visioning, a research or policy incentive is implemented to support some more or less articulated (long-term) change. However, in order to assess this contribution in an evaluation, the intended workings of the programme towards this vision need to be uncovered. Evaluation theory provides tools for this, e.g. in theory-based evaluation approaches using intervention theories that traditionally outline how a programme is intended to reach its own goals, but this can be expanded to also capture how the evaluand needs to perform to contribute to fulfilling the desired vision. Such approaches require the involvement of key actors and stakeholders in order to acquire their perceptions and insights concerning how an instrument is intended to work towards a goal or vision, and such heeding of actors is also greatly emphasized in sociology of science and policy analysis. These disciplines provide deeper understandings of the processes that underpin research and policy incentives as parts of

societal and political contexts, and thus emphasize the importance of understanding how different actors influence and affect an implementation. This adds an emphasis on assessing an evaluand as part of societal systems, where both research and policy incentives are shaped and promoted by different actors' value systems and perceptions. Such interdisciplinary insights can further promote and guide a deeper involvement of key stakeholders in evaluations, which may yield a more versatile approach to assessing an evaluand, and a more inclusive picture of how the evaluand is seen from various perspectives.

Looking at experimentation and learning, these are drivers that can easily be argued to build on a need for merging theoretical insights, as outcomes and knowledge do not necessarily have only one theoretical area of application. Again noting that transitions are system-wide, it is key to consider how experimentation may create effects that go beyond the disciplinary scope of one set of theories. Here, sociology of science and policy analysis outline the contexts and systems in which the evaluand is implemented, and provide conceptualizations of how the evaluand is tied to different governance areas, markets, and actors, which all may have interests and incentives to act on the learning that can be acquired from the evaluand and its experimental efforts. Evaluation theory again provides approaches for capturing these different viewpoints, and for gauging and articulating the experimental efforts and outcomes, rendering them tangible and visible, and thus articulating the learning that can be drawn.

In all, articulating the theoretical contributions of this thesis quickly becomes a matter of navigating the complexities that come both with seeking to support a transition through transformative evaluation, and with working interdisciplinarily. Nevertheless, this thesis seeks to make a theoretical contribution based on the argumentation that this interdisciplinary research also supports disciplinary insights by demonstrating how a discipline can be effectively applied in complex matters (such as evaluating for a transition), as well as by outlining the synergies that different disciplines indeed host when brought together in joint frameworks.

5.3 Implications for evaluation practitioners

This research has outlined empirical and theoretical contributions, but what are the implications of these for practitioners working with evaluation of research and policy incentives?

As this research has taken a focus on energy efficiency in buildings, the outlines for how transformative evaluation can be designed and conducted are hoped to be of particular value in this sector. One contribution for practitioners is the suggestions on how to complement current Swedish evaluation practices to support a transition; another is the provision of interdisciplinary frameworks that articulate key areas to consider when

moving towards transformative evaluation. Moreover, it is also hoped that these frameworks can provide concrete categories that facilitate the discussions surrounding evaluations, as terminology and a common and shared set of aspects is introduced. These frameworks are, however, perhaps only an outline of *one* way of moving towards transformative evaluation, and there are likely other disciplinary insights and approaches that can be added to provide more varied and encompassing evaluations. Thus, from here, practitioners are encouraged to build further on the outlined frameworks, to try them, to develop them, and to shape them according to their needs and intended uses.

Furthermore, while the transformative evaluation approach suggested in this thesis is designed in an energy system context, its components are arguably translatable to other evaluation sectors as well, where they might be used to provide guidance on key issues to consider when seeking to allow more transformative perspectives in evaluations. Thus, transformative evaluation should be seen as malleable and able to provide transformative and knowledge-based insights to support sustainability transitions in many different sectors.

However, the suggestions made above on how to support a transition by complementing current evaluation practices do come with some implications for practitioners when commissioning and performing evaluations of research and policy incentives. While the suggestions are grounded in current practices and seek to outline key areas for further development, suggestions still call for some major changes in how evaluations are conducted, both in terms of design and in terms of ownership and agency.

One key implication is that transformative evaluation does call for a more holistic approach to evaluation at large: to move beyond regarding evaluations as locked-in-time snapshots of an intervention, and towards seeing evaluations as interconnected knowledge carriers with a purpose that is larger than programme improvement. Such a holistic approach does seem to contrast with current Swedish evaluation approaches to a certain extent, as evaluations in Sweden are often commissioned and/or performed within delimited jurisdictions or governance areas, and evaluation use is often limited to programme learning rather than embracing a cumulative learning of contributions to a transition. To support a transition that will span across governance areas, there is a need to prepare for integrating evaluation practices further in the policy process, by preparing for shared data access and cross-border collaboration between different departments and agencies, and with actors and stakeholders that hold essential skills or knowledge in conducting evaluations.

Such a cross-border collaboration around transformative evaluation will likely require an increased overview and relocation of mandates and resources to plan, design and perform evaluations in order to ensure that the knowledge exchange and synthesis of insights are secured. The challenges related to this concern the need for time and

resources to establish new evaluation procedures, and also the requirement for a continued plan for management in order to stay operational and effective.

5.4 Reflection on research approach and methods

This research built on interdisciplinarity, in bringing key disciplines together to provide encompassing frameworks and suggestions for advancing transformative evaluation. As already stated, an interdisciplinary approach was considered suitable for attempting to address the complexities that surround an energy transition to a satisfying degree, in order to provide a versatile and inclusive frame for transformative evaluation. As discussed in Section 3, this research approach required a balance between breadth and depth, but in terms of the disciplines applied it is concluded that they have all provided valuable insights, and that removing any of them would have given a narrower perspective and less weight to the many aspects that have been identified as important for advancing transformative evaluation.

Reflecting on the relevance of the four disciplines that were used to build the foundations of this thesis, they have proven valuable for providing insights on various aspects of evaluating for a transition, covering all necessary aspects pertaining to transition knowledge, evaluation insights, and key aspects concerning both research and policy instruments. As such, they have all contributed with specific insights that, when brought together, have supported the thesis's overall objective of advancing transformative evaluation. However, the theoretical scoping applied in this thesis does mean that insights from other disciplines may still provide additional perspectives and angles to transformative evaluation, which may support an even more holistic approach to supporting a transition.

Designing the research as a case study focusing on Swedish evaluation practices for energy efficiency in buildings is still, at the completion of this research, considered a relevant area to focus on when seeking to outline and develop transformative evaluation. When this research was initiated in early 2016, the Paris Agreement had just been adopted, emphasizing the urgency of global actions to address climate change. Thus, the research took place in a context of global action, awareness and concern, which is still very much the case at the time the research was completed. As such, orienting the research around a sustainability transition, focusing in particular on the case of energy efficiency in buildings, was a convenient way to keep the research relevant and timely, while still scoped enough to be manageable and relatable in terms of knowledge production.

Reflecting on the methods used in this thesis for data collection and analysis, they assisted in building the thesis in different ways. The document studies of existing evaluation reports, as well as the process of locating and acquiring these evaluation

reports by interpersonal contacts with authorities, and by reviewing other written documents (e.g. Sweden's National Communications on Climate Change (Ministry of the Environment, Sweden, 2014, 2009)), in all provided an overview of the current efforts undertaken with regard to evaluation for energy efficiency in buildings in Sweden. This stocktaking of current practices fulfilled the purpose of gaining an understanding of the present performance of evaluations, which in turn was used to identify the main strengths and areas to be further complemented in order to move towards transformative evaluation.

What the document studies lacked in terms of up-to-date insights about the processes of using evaluations, and the perceptions of and possibilities expressed by Swedish state agencies in moving towards transformative evaluation, was complemented by the focus group sessions and the structured interview questionnaire. The focus group sessions served multiple purposes, as they provided an opportunity for the researchers to present the work conducted within the research project thus far and share some results and implications, and served to introduce the outlines of a transformative evaluation approach and discuss this with key stakeholders. As this thesis set out to advance evaluation practices and support them in becoming more transformative, it was crucial to reach out to the actors with the mandate and the overall responsibility for performing evaluations of research and policy incentives on a national level. The discussions that the researchers and the representatives at the different state agencies engaged in turned out to be more than just a mere introduction of research aimed at advancing transformative evaluation, as it also gave rise to additional insights into how authorities work with evaluations, their wishes and limitations, and so on. Thus, the focus group sessions were valuable both for outreach and for sharpening the transformative evaluation approach according to the needs of the key receivers, as well as for acquiring timely insights about current workings at state agencies concerning evaluation issues.

Since the focus group sessions were based on presentations and discussions, it was important to ensure that the insights gained were indeed consistent with what had been said. The structured interview questionnaire thus fulfilled the purpose of triangulating and validating the statements drawn from focus group sessions regarding the benefits and challenges of adopting a transformative evaluation approach, yet also served to acquire deeper insights into how evaluations were actually perceived to be used. This latter aspect was critical for understanding where to aim further efforts for increasing use and making it relevant for supporting a transition. As such, the individually administered questionnaire served its purpose well in both complementing the more interactive focus group sessions, and in providing results that were more safeguarded from either misunderstandings or misinterpretations, or from omissions of sharing viewpoints among participants due to perceived peer pressure or other group dynamics.

All in all, the research design and methods have provided the intended data and knowledge, which has then been synthesized in the above outlined overarching

contributions to empirics and theory. Moreover, they have also provided additional valuable aspects such as interpersonal contacts with key stakeholders at Swedish state agencies, as well as an established dialogue about transformative evaluation during the time of the research project.

5.5 Suggestions for future research

Transformative evaluation builds on the deliberate planning and conducting of different evaluations, which when seen together can provide insights into how transformative efforts from a range of research and policy incentives can together support a transition. The transformative effects that each incentive has may not spur a transition on their own, but may provide the fertile grounds for another incentive's effects to emerge or ripen, thus supporting a *cascading* motion towards a transition. In order to understand how these different transformations are linked to each other, further research is encouraged to focus on evaluating the micro-steps of a transition towards a more sustainable energy system in a certain context. Such research could go into more depth concerning the (potential) transformative effects of enforced incentives, e.g. developing insights about how to assess the different transition pathways (as outlined by Geels and Schot, 2007) that successful transformations support, and connecting these with the need for subsequent research and policy incentives.

Further research may also be geared towards refining the theoretical foundations of transformative evaluation, either by supporting breadth in terms of reviewing additional theoretical disciplines, or by supporting depth in terms of exploring specific aspects from the selected disciplines in more detail. For additional breadth, many disciplines may be considered for further supporting transformative evaluation, but a key area is arguably to focus on the individual and behavioural levels of supporting a transition. It has been emphasized in the reviewed disciplines that *actors* and their beliefs, networks, and value systems are important aspects to consider when evaluating both research and policy. Thus, disciplines such as the aforementioned environmental psychology and organizational studies are suggested as essential areas to consult for further coupling the role of actors in a transition, and for outlining how to capture these in transformative evaluation.

In terms of supporting further depth, a detailed focus is suggested to look more deeply into how different kinds of experimental efforts can be gauged from a transformative perspective in evaluations. Kivimaa et al. (2017) outline different kinds of experimentation to support a transition, which take place in different parts of the socio-technical system, and thus involve different actors. By furthering an understanding of how the characteristics and key features of such different types of experimentations can be recognized and thus assessed in evaluations, transformative evaluation may be further supported in the pursuit of placing the transformational contributions of an

evaluated in a societal perspective, and in providing overviews of which kinds of experimental efforts are currently prevailing or lacking.

Another area for further research, which goes slightly more beyond the research presented in this thesis, concerns the *perception* of transformative evaluation among key stakeholders and users. Ultimately, the successful adoption of a transformative evaluation approach comes down to a *willingness* among these actors to support and invest in such an approach. This, in turn, requires that there is a perception and trust among stakeholders that evaluations can in fact become the knowledge providers needed to support a sustainable energy transition.

Thus, the role of *framing* in terms of perceiving evaluations as carriers of knowledge for a transition is suggested as an area for future research – research that may provide not only essential knowledge concerning how to support a transformative approach, but importantly also how to make it widely accepted and sustainable in the long term. Framing may be a central area for further investigation for two main reasons: firstly, for providing deeper insights into existing frames relevant for evaluations in regard to how they fare in relation to supporting a transition, and secondly, for further research into the establishment of frames that support transformative evaluation in a just and equitable manner, with regard to how the conditions and the presuppositions of frames affect stakeholders in a transition context.

Framing can be broadly explained as the process in which people construct meaning in a context or situation (Benford and Snow, 2000; Goffman, 1974), and frames assist in defining, conceptualizing and re-formulating one's understanding of an issue by providing yardsticks for sorting, evaluating, and acting upon this issue or piece of information (Chong and Druckman, 2007; Entman, 1993). One key feature of frames in the particular case of supporting a just and equitable transition towards a more sustainable energy system is that a prerequisite for establishing a frame and rendering it viable is consent among people (Fisher, 1997). New or competing frames can become dominant by causing a questioning of existing standpoints, or by assisting in highlighting new angles not previously included in a receiver's decision-making processes (Carnahan et al., 2019; Kangas et al., 2014), e.g. seeing evaluations as carriers of transformative knowledge. However, on this note – and importantly for using evaluations in a transformative context – framing effects require that the receivers are open to unfavourable evidence, and that they are able and willing to challenge and revise current frames and preferences (Chong and Druckman, 2007). This, in itself, is also a key challenge in regard to increased evaluation use, and thus, there are certainly synergies to be found in considering evaluation use and framing from the perspective of supporting a transition.

Another important aspect to consider for future research regarding framing of evaluations as knowledge providers in a transformational context is the underlying foundations of a frame. Frames are shaped by a set of viewpoints which affect how

people respond to them, and it thus lies in their nature to build on certain viewpoints, while disregarding others (Entman, 1993). As such, framing evaluations to support a transition comes with a responsibility in terms of providing reliable frames that are built on equity and transparency about underlying assumptions, drivers, and goals. This in order to prevent unjust use of frames to support inequitable, unsustainable, or ungrounded claims of supporting a transition.

6. Conclusions

Returning to the quote by Patton (2019, p. 103), saying that *'Evaluating transformation means transforming evaluation'*, this thesis provides a suggestion for how such a transformed evaluation may be thought of, designed and achieved.

In relation to the overall objective of this research – to support a transition through transformative evaluations – the findings presented in this thesis provide some empirical and theoretical contributions. The empirical insights provide deeper knowledge about current Swedish evaluation practices with regard to how they are conducted and how they fare in relation to making transformative assessments. This makes it possible to outline areas where concrete suggestions can be made for how to complement these practices to fully harness their potential of providing *transformative* and *knowledge-based* insights that can support a transition. Suggestions to promote transformative assessments include making evaluations assess the contributions of evaluands in relation to societal energy visions to larger extents, assessing experimentation and learning that can support transformative efforts and their upscaling, moving focus away from the predominant programme level towards a systems perspective, and involving actors from various actors groups to generate versatile accounts of the evaluand. Together, these aspects support knowledge production in evaluations that can advantageously be synthesized to provide a more holistic overview of collected efforts towards realizing a transition.

Suggestions to promote *rigorous and knowledge-based* assessments include capitalizing on the frequent application of multiple methods in the evaluations, and using these to cross-check results using triangulation. The application of different criteria should also be carried forward, and can be complemented by applying a wider selection of criteria that can gear focus towards effects on a systems level.

The theoretical contributions made in this research are both disciplinary and interdisciplinary. They are disciplinary in that they identify and outline key features from four disciplines, in order to concretely outline aspects that can support evaluation of research and policy incentives in relation to a transition; and they are interdisciplinary in that the thesis provides tangible and tested frameworks which synthesize these aspects, in order to provide a shared terminology and guide for designing and conducting transformative and knowledge-based evaluations.

The research also shows that there are both benefits and challenges related to moving towards a transformative evaluation approach among Swedish state agencies, but that current practices for using evaluations provide a good foundation to build on.

References

- Aboelela, S. W., Larson, E., Bakken, S., Carrasquillo, O., Formicola, A., Glied, S. A., ... Gebbie, K. M. (2007). Defining Interdisciplinary Research: Conclusions from a Critical Review of the Literature. *Health Services Research, 42*(1p1), 329–346. <https://doi.org/10.1111/j.1475-6773.2006.00621.x>
- Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation, 29*(1), 1–12. [https://doi.org/10.1016/S0191-491X\(03\)90001-0](https://doi.org/10.1016/S0191-491X(03)90001-0)
- Alkin, M. C., Vo, A. T., & Christie, C. A. (2012). The evaluator's role in valuing: Who and with whom. *New Directions for Evaluation, 2012*(133), 29–41. <https://doi.org/10.1002/ev.20004>
- Arnold, E. (2004). Evaluating research and innovation policy: A systems world needs systems evaluations. *Research Evaluation, 13*(1), 3–17. <https://doi.org/10.3152/147154404781776509>
- Arnold, E., Åström, T., Glass, C., & De Scalzi, M. (2018). *How should we evaluate complex programmes for innovation and socio-technical transitions?* Technopolis Group. Retrieved from <https://www.technopolis-group.com/wp-content/uploads/2020/02/How-should-we-evaluate-complex-programmes-for-innovation-and-socio-technical-transitions.pdf>
- Avelino, F., & Wittmayer, J. M. (2016). Shifting Power Relations in Sustainability Transitions: A Multi-actor Perspective. *Journal of Environmental Policy & Planning, 18*(5), 628–649. <https://doi.org/10.1080/1523908X.2015.1112259>
- Baylan, I., Bäckström-Johansson, M., Gunther, P., Hjalmered, L., Lahti, B., Nilsson, I., ... Westlund Å. (2017). Kraftsamling för framtidens energi [Concerted efforts for the energy of the future]. Swedish Government Official Report, SOU 2017:2.
- Belcher, B. M., Rasmussen, K. E., Kemshaw, M. R., & Zornes, D. A. (2016). Defining and assessing research quality in a transdisciplinary context. *Research Evaluation, 25*(1), 1–17. <https://doi.org/10.1093/reseval/rvv025>
- Benford, R. D., & Snow, D. A. (2000). Framing Processes and Social Movements: An Overview and Assessment. *Annual Review of Sociology, 26*(1), 611–639. <https://doi.org/10.1146/annurev.soc.26.1.611>
- Bhaskar, R. (2010). Contexts of interdisciplinarity—Interdisciplinarity and climate change. In R. Bhaskar, C. Frank, K. G. Høyer, P. Naess, & J. Parker (Eds.), *Interdisciplinarity and Climate Change: Transforming Knowledge and Practice for Our Global Future* (pp. 1–24). London: Routledge. DOI: 10.4324/9780203855317
- Bhaskar, R., Frank, C., Høyer, K. G., Naess, P., & Parker, J. (2010). *Interdisciplinarity and Climate Change: Transforming Knowledge and Practice for Our Global Future*. London: Routledge. DOI: 10.4324/9780203855317

- Bickman, L. (1987). The functions of program theory. *New Directions for Program Evaluation*, 1987(33), 5–18. <https://doi.org/10.1002/ev.1443>
- Birnbaum, M., & Mickwitz, P. (2009). Editors' notes. *New Directions for Evaluation, Special Issue: Environmental Program and Policy Evaluation: Addressing Methodological Challenges*, 2009(122), 1–8. <https://doi.org/10.1002/ev.290>
- Blumstein, C., Goldstone, S., & Lutzenhiser, L. (2000). A theory-based approach to market transformation. *Energy Policy*, 28(2), 137–144. [https://doi.org/10.1016/S0301-4215\(99\)00093-2](https://doi.org/10.1016/S0301-4215(99)00093-2)
- Borrás, S. (2012). Three Tensions in the Governance Of Science and Technology. In D. Levi-Faur (Ed.), *Oxford handbook of governance*. Oxford: Oxford University Press. Retrieved from <http://oxfordhandbooks.com/view/10.1093/oxfordhb/9780199560530.001.0001/oxfordhb-b-9780199560530-e-30>
- Bovens, M., 't Hart, P., & Kuipers, S. (2008). The Politics of Policy Evaluation. In M. Moran, M. Rein, & R.E. Goodin (Eds.), *The Oxford handbook of public policy*. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199548453.003.0015>
- Brewer, G. D., & DeLeon, P. (1983). *The foundations of policy analysis*. Homewood, Ill: Dorsey Press.
- Bruce, A., Lyall, C., Tait, J., & Williams, R. (2004). Interdisciplinary integration in Europe: The case of the Fifth Framework programme. *Futures*, 36(4), 457–470. <https://doi.org/10.1016/j.futures.2003.10.003>
- Bruckner, T., Bashmakov, I. A., Mulugetta, Y., Chum, H., de la Vega Navarro, A., Edmonds, J., ... Zhang, X. (2014). Energy Systems. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, ... J.C. Minx (Eds.), *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge ; New York: Cambridge University Press.
- Bryman, A. (2012). *Social research methods* (Fourth edition). Oxford ; New York: Oxford University Press.
- Bucchi, M. (2004). *Science in society: An introduction to social studies of science* (First English language edition). London ; New York: Routledge.
- Buonocore, J. J., Luckow, P., Norris, G., Spengler, J. D., Biewald, B., Fisher, J., & Levy, J. I. (2016). Health and climate benefits of different energy-efficiency and renewable energy choices. *Nature Climate Change*, 6(1), 100–105. <https://doi.org/10.1038/nclimate2771>
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, 24(4), 409–429. <https://doi.org/10.1037/h0027982>
- Cantin, R., Kindinis, A., & Michel, P. (2012). New approaches for overcoming the complexity of future buildings impacted by new energy constraints. *Futures*, 44(8), 735–745. <https://doi.org/10.1016/j.futures.2012.05.001>
- Carnahan, D., Hao, Q., & Yan, X. (2019). Framing Methodology: A Critical Review. *Oxford Research Encyclopedia of Politics*. Oxford: Oxford University Press. Retrieved from <https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-1026>

- CECAN (2018). *Policy evaluation for a complex world*. Retrieved from <https://www.cecan.ac.uk/>
- Chalmers, A. F. (2013). *What is this thing called science?* (Fourth edition). Maidenhead: McGraw-Hill Education.
- Chelimsky, E. (2015). A strategy for improving the use of evaluation findings in policy. In C. A. Christie, A. T. Vo, & M. C. Alkin (Eds.), *Evaluation use and decision-making in society: A tribute to Marvin C. Alkin* (pp. 73–90). Charlotte, NC: Information Age Pub. Inc.
- Chen, H.-T. (1990). *Theory-driven evaluations*. Newbury Park, CA: SAGE Publications.
- Chen, H.-T., & Rossi, P. H. (1989). Issues in the theory-driven perspective. *Evaluation and Program Planning*, 12(4), 299–306. [https://doi.org/10.1016/0149-7189\(89\)90046-3](https://doi.org/10.1016/0149-7189(89)90046-3)
- Cho, J. (2017). *Evaluating qualitative research*. New York, NY: Oxford University Press.
- Chong, D., & Druckman, J. N. (2007). Framing Theory. *Annual Review of Political Science*, 10(1), 103–126. <https://doi.org/10.1146/annurev.polisci.10.072805.103054>
- Chouinard, J. A. (2013). The Case for Participatory Evaluation in an Era of Accountability. *American Journal of Evaluation*, 34(2), 237–253. <https://doi.org/10.1177/1098214013478142>
- Christie, C. A., & Alkin, M. C. (2013). An Evaluation Theory Tree. In M. C. Alkin (Ed.), *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences* (Second edition). Thousand Oaks: SAGE Publications.
- Christie, C. A., & Fleischer, D. N. (2009). Social Inquiry Paradigms as a Frame for the Debate on Credible Evidence. In S. I. Donaldson, C. A. Christie, & M. M. Mark (Eds.), *What Counts as a Credible Evidence in Applied Research and Evaluation Practice?* (pp. 19–30). Los Angeles: SAGE.
- Christie, C. A., & Fleischer, D. N. (2010). Insight Into Evaluation Practice: A Content Analysis of Designs and Methods Used in Evaluation Studies Published in North American Evaluation-Focused Journals. *American Journal of Evaluation*, 31(3), 326–346. <https://doi.org/10.1177/1098214010369170>
- Colebatch, H. K. (2009). Governance as a conceptual development in the analysis of policy. *Critical Policy Studies*, 3(1), 58–67. <https://doi.org/10.1080/19460170903158107>
- Cook, T. D. (1997). Lessons Learned in Evaluation Over the Past 25 Years. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: SAGE Publications.
- Crabbé, A., & Leroy, P. (2008). *The handbook of environmental policy evaluation*. London ; Sterling, VA: Earthscan.
- Cruickshank, J. (Ed.) (2003). *Critical Realism: The Difference it Makes*. London: Routledge.
- Cunningham, P., Edler, J., Flanagan, K., & Larédo, P. (2016). The innovation policy mix. In J. Edler, P. Cunningham, A. Gök, & P. Shapira (Eds.), *Handbook of Innovation Policy Impact* (pp. 505–542). Cheltenham, UK Northampton, MA: Edward Elgar.
- Denzin, N. K. (1971). The Logic of Naturalistic Inquiry. *Social Forces*, 50(2), 166. <https://doi.org/10.2307/2576935>
- Dunn, W. N. (2018). *Public policy analysis: An integrated approach* (Sixth edition). New York: Routledge, Taylor & Francis Group.

- Elzinga, A. (2012). Features of the current science policy regime: Viewed in historical perspective. *Science and Public Policy*, 39(4), 416–428. <https://doi.org/10.1093/scipol/scs046>
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- European Commission (2017). *Better Regulation Guidelines* [European Commission Staff Working Document]. Retrieved from <https://ec.europa.eu/info/sites/info/files/better-regulation-guidelines.pdf>
- European Commission (2020). *In focus: Energy efficiency in buildings*. Brussels: European Commission, Department Energy.
- European Environment Agency (2016). *Environment and climate policy evaluation*. Luxembourg: Publications Office of the European Union.
- European Environment Agency (2019a). *Adaptation challenges and opportunities for the European energy system: Building a climate-resilient low-carbon energy system*. Retrieved from http://publications.europa.eu/publication/manifestation_identifier/PUB_THAL19005ENN
- European Environment Agency (2019b). *Trends and projections in Europe 2019: Tracking progress towards Europe's climate and energy targets. EEA Report No 15/2019*. Retrieved from https://op.europa.eu/publication/manifestation_identifier/PUB_THAL19016ENN
- Ezrahi, Y. (1980). Utopian and pragmatic rationalism: The political context of scientific advice. *Minerva*, 18(1), 111–131. <https://doi.org/10.1007/BF01096662>
- Ferraro, P. J. (2009). Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation, Special Issue: Environmental Program and Policy Evaluation: Addressing Methodological Challenges 2009*(122), 75–84.
- Fisher, K. (1997). Locating Frames in the Discursive Universe. *Sociological Research Online*, 2(3), 1–24. <https://doi.org/10.5153/sro.78>
- Flanagan, K., Uyarra, E., & Laranja, M. (2011). Reconceptualising the ‘policy mix’ for innovation. *Research Policy*, 40(5), 702–713. <https://doi.org/10.1016/j.respol.2011.02.005>
- Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, 12(2), 219–245. <https://doi.org/10.1177/1077800405284363>
- Frodeman, R. (2014). *Sustainable knowledge - a theory of interdisciplinarity*. Basingstoke: Palgrave Macmillan. <https://doi.org/10.1057/9781137303028>
- Furubo, J.-E. (1994). Learning from Evaluations: The Swedish Experience. In F. L. Leeuw, R. C. Rist, & R. C. Sonnichsen (Eds.), *Can Governments Learn? - Comparative Perspectives on Evaluation & Organizational Learning*. New Brunswick, N.J: Transaction Publishers.
- Furubo, J.-E., Rist, R. C., & Sandahl, R. (Eds.). (2002). *International atlas of evaluation*. New Brunswick, U.S.A: Transaction Publishers.
- GEA (2012). *Global Energy Assessment—Toward a Sustainable Future*. Retrieved from www.globalenergyassessment.org
- Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Research Policy*, 31(8–9), 1257–1274. [https://doi.org/10.1016/S0048-7333\(02\)00062-8](https://doi.org/10.1016/S0048-7333(02)00062-8)
- Geels, F. W., & Schot, J. (2007). Typology of sociotechnical transition pathways. *Research Policy*, 36(3), 399–417. <https://doi.org/10.1016/j.respol.2007.01.003>

- Geuna, A., & Martin, B. R. (2003). University Research Evaluation and Funding: An International Comparison. *Minerva*, 41(4), 277–304. <https://doi.org/10.1023/B:MINE.0000005155.70870.bd>
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzmann, S., Scott, P., & Trow, M. (2010). Chapter 1: Evolution of Knowledge Production. In M. Gibbons, C. Limoges, H. Nowotny, S. Schwartzmann, P. Scott, & M. Trow (Eds.), *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies* (pp. 17–46). London: SAGE Publications.
- Gibbons, M., Limoges, C., & Scott, P. (2011). Revisiting Mode 2 at Noors Slott. *Prometheus*, 29(4), 361–372. <https://doi.org/10.1080/08109028.2011.641384>
- Gläser, J. (2007). The Social Orders of Research Evaluation Systems. In R. Whitley & J. Gläser (Eds.), *The changing governance of the sciences: The advent of research evaluation systems* (pp. 245–266). Dordrecht: Springer.
- Global Alliance for Buildings and Construction (2019). *2019 Global Status Report for Buildings and Construction: Towards a zero-emissions, efficient and resilient buildings and construction sector*. Prepared by the International Energy Agency and the United Nations Environment Programme. Retrieved from <https://www.unep.org/resources/publication/2019-global-status-report-buildings-and-construction-sector>
- Godin, B. (2007). Science, accounting and statistics: The input–output framework. *Research Policy*, 36(9), 1388–1403. <https://doi.org/10.1016/j.respol.2007.06.002>
- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. London: Harper & Row.
- Government bill 1996/97:84. Regeringens proposition 1996/97:84, En uthållig energiförsörjning [A sustainable energy provision]. The Swedish Government.
- Government bill 2012/13:21. Regeringens proposition 2012/13:21 Forskning och innovation för ett långsiktigt hållbart energisystem [Research and innovation for a long-term sustainable energy system]. The Swedish Government.
- Government bill 2016/17:66. Regeringens proposition 2016/17:66 Forskning och innovation på energiområdet för ekologisk hållbarhet, konkurrenskraft och försörjningstrygghet [Research and innovation within the area of energy for ecological sustainability, competitiveness, and security of supply]. The Swedish Government.
- Greene, J. C. (2008). Is Mixed Methods Social Inquiry a Distinctive Methodology? *Journal of Mixed Methods Research*, 2(1), 7–22. <https://doi.org/10.1177/1558689807309969>
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: SAGE Publications.
- Guston, D. H. (2001). Boundary Organizations in Environmental Policy and Science: An Introduction. *Science, Technology, & Human Values*, 26(4), 399–408. <https://doi.org/10.1177/016224390102600401>
- Haas, P. M. (1992). Epistemic Communities and International Policy Coordination. *International Organization*, 46(1), 1–35.
- Hazelkorn, E. (2015). *Rankings and the Reshaping of Higher Education. The Battle for World-Class Excellence*. London: Palgrave Macmillan UK.

- Hellström, T. (2011). Homing in on excellence: Dimensions of appraisal in Center of Excellence program evaluations. *Evaluation*, 17(2), 117–131. <https://doi.org/10.1177/1356389011400891>
- Hildén, M. (2009). Time horizons in evaluating environmental policies. *New Directions for Evaluation*, 2009(122), 9–18. <https://doi.org/10.1002/ev.291>
- Hildén, M. (2011). The evolution of climate policies – the role of learning and evaluations. *Journal of Cleaner Production*, 19(16), 1798–1811. <https://doi.org/10.1016/j.jclepro.2011.05.004>
- Hildén, M., Jordan, A., & Rayner, T. (2014). Climate policy innovation: Developing an evaluation perspective. *Environmental Politics*, 23(5), 884–905. <https://doi.org/10.1080/09644016.2014.924205>
- Hogwood, B. W., & Gunn, L. A. (1984). *Policy analysis for the real world*. Oxford ; New York: Oxford University Press.
- Howe, K. R. (2012). Mixed Methods, Triangulation, and Causal Explanation. *Journal of Mixed Methods Research*, 6(2), 89–96. <https://doi.org/10.1177/1558689812437187>
- Huitema, D., Jordan, A., Massey, E., Rayner, T., van Asselt, H., Haug, C., ... Stripple, J. (2011). The evaluation of climate policy: Theory and emerging practice in Europe. *Policy Sciences*, 44(2), 179–198. <https://doi.org/10.1007/s11077-011-9125-7>
- International Energy Agency (2017). *Energy Technology Perspectives 2017: Catalysing Energy Technology transformations*. Paris.
- International Energy Agency (2018). *Perspectives for the Energy Transition: The Role of Energy Efficiency*. Paris.
- International Energy Agency (2019). *Energy Efficiency 2019*. Retrieved from <https://www.iea.org/reports/energy-efficiency-2019>
- IPCC (2014). *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge ; New York: Cambridge University Press.
- IPCC (2018). *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. In Press.
- IRENA (2019). *Global energy transformation: A roadmap to 2050 (2019 edition)*. Abu Dhabi: International Renewable Energy Agency.
- Jacob, S., Speer, S., & Furubo, J.-E. (2015). The institutionalization of evaluation matters: Updating the *International Atlas of Evaluation* 10 years later. *Evaluation*, 21(1), 6–31. <https://doi.org/10.1177/1356389014564248>
- Jordan, A., & Lenschow, A. (2010). Environmental policy integration: A state of the art review. *Environmental Policy and Governance*, 20(3), 147–158. <https://doi.org/10.1002/ect.539>
- Julnes, G. (2019a). Evaluating Sustainability: Controversies, Challenges, and Opportunities. *New Directions for Evaluation*, 2019(162), 13–28. <https://doi.org/10.1002/ev.20361>
- Julnes, G. (2019b). Supporting Transitions to Sustainability: Evaluation for Managing Processes in the Public Interest. *New Directions for Evaluation*, 2019(162), 119–154. <https://doi.org/10.1002/ev.20366>

- Kangas, O. E., Niemelä, M., & Varjonen, S. (2014). When and why do ideas matter? The influence of framing on opinion formation and policy change. *European Political Science Review*, 6(1), 73–92. <https://doi.org/10.1017/S1755773912000306>
- Kemp, R. (1994). Technology and the transition to environmental sustainability. *Futures*, 26(10), 1023–1046. [https://doi.org/10.1016/0016-3287\(94\)90071-X](https://doi.org/10.1016/0016-3287(94)90071-X)
- Kemp, R., Schot, J., & Hoogma, R. (1998). Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technology Analysis & Strategic Management*, 10(2), 175–198. <https://doi.org/10.1080/09537329808524310>
- King, J. A., & Alkin, M. C. (2019). The Centrality of Use: Theories of Evaluation Use and Influence and Thoughts on the First 50 Years of Use Research. *American Journal of Evaluation*, 40(3), 431–458. <https://doi.org/10.1177/1098214018796328>
- Kivimaa, P., Hildén, M., Huitema, D., Jordan, A., & Newig, J. (2017). Experiments in climate governance – A systematic review of research on energy and built environment transitions. *Journal of Cleaner Production*, 169, 17–29. <https://doi.org/10.1016/j.jclepro.2017.01.027>
- Klein, J. T. (1990). *Interdisciplinarity: History, Theory, and Practice*. Detroit: Wayne State University Press.
- Klein, J. T. (2004). Interdisciplinarity and complexity: An evolving relationship. *Emergence: Complexity & Organization*, 6(1/2), 2–10.
- Klein, J. T. (2017). Typologies of Interdisciplinarity - The Boundary Work of Definition. In R. Frodeman (Ed.), *The Oxford Handbook of Interdisciplinarity* (Second edition) (pp. 21–39). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198733522.013.3>
- Knaap, P. V. D. (2000). Performance Management and Policy Evaluation in the Netherlands: Towards an Integrated Approach. *Evaluation*, 6(3), 335–350. <https://doi.org/10.1177/13563890022209325>
- Langfeldt, L. (2001). The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome. *Social Studies of Science*, 31(6), 820–841. <https://doi.org/10.1177/030631201031006002>
- Leeuw, F. L. (2009). Evaluation policy in the Netherlands. *New Directions for Evaluation*, 2009(123), 87–102. <https://doi.org/10.1002/ev.307>
- Leite, D., & Pinho, I. (2017). *Evaluating Collaboration Networks in Higher Education Research - Drivers of Excellence*. Cham: Springer International Publishing. Retrieved from <http://link.springer.com/10.1007/978-3-319-45225-8>
- Lenton, T. M. (2011). Early warning of climate tipping points. *Nature Climate Change*, 1(4), 201–209. <https://doi.org/10.1038/nclimate1143>
- Leviton, L. C. (2003). Evaluation Use: Advances, Challenges and Applications. *American Journal of Evaluation*, 24(4), 525–535. doi:10.1177/109821400302400410
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic Reviews and Meta-Analysis*. New York: Oxford University Press. Retrieved from <https://doi.org/10.1093/acprof:oso/9780195326543.001.0001>
- Loorbach, D. (2010). Transition Management for Sustainable Development: A Prescriptive, Complexity-Based Governance Framework. *Governance*, 23(1), 161–183. <https://doi.org/10.1111/j.1468-0491.2009.01471.x>

- Loorbach, D., Frantzeskaki, N., Avelino, F. (2017). Sustainability Transitions Research: Transforming Science and Practice for Societal Change. *Annual Review of Environment and Resources*, 42(1), 599–626. <https://doi.org/10.1146/annurev-environ-102014-021340>
- Luederitz, C., Schöpke, N., Wiek, A., Lang, D.J., Bergmann, M., Bos, J.J., ... Westley, F.R. (2017). Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments. *Journal of Cleaner Production, Experimentation for climate change solutions*, 169, 61–76. <https://doi.org/10.1016/j.jclepro.2016.09.005>
- Markard, J., Raven, R., & Truffer, B. (2012). Sustainability transitions: An emerging field of research and its prospects. *Research Policy*, 41(6), 955–967. <https://doi.org/10.1016/j.respol.2012.02.013>
- Meadowcroft, J. (2009). What about the politics? Sustainable development, transition management, and long term energy transitions. *Policy Sciences*, 42(4), 323–340. <https://doi.org/10.1007/s11077-009-9097-z>
- Mertens, D. M., & Hesse-Biber, S. (2013). Mixed Methods and Credibility of Evidence in Evaluation. *New Directions for Evaluation*, 2013(138), 5–13. <https://doi.org/10.1002/ev.20053>
- Meyer, W., & Stockmann, R. (2013). Evaluation Approaches and Their Fundamental Theoretical Principles. In R. Stockmann & W. Meyer (Eds.), *Functions, Methods and Concepts in Evaluation Research* (pp. 108–174). Basingstoke: Palgrave Macmillan.
- Mickwitz, P. (2003). A Framework for Evaluating Environmental Policy Instruments Context and Key Concepts. *Evaluation*, 9(4), 415–436. <https://doi.org/10.1177/1356389003094004>
- Ministry of the Environment, Sweden (2009). *Sweden's fifth National Communication on climate change: under the United Nations Framework Convention on Climate Change*. Ds 2009:63. Stockholm: Fritzes.
- Ministry of the Environment, Sweden (2014). *Sweden's sixth National Communication on climate change: under the United Nations Framework Convention on Climate Change*. Ds 2014:11. Stockholm: Fritzes.
- Mårtensson, P., Fors, U., Wallin, S.-B., Zander, U., & Nilsson, G. H. (2016). Evaluating research: A multidisciplinary approach to assessing research practice and quality. *Research Policy*, 45(3), 593–603. <https://doi.org/10.1016/j.respol.2015.11.009>
- Neij, L. (2001). Methods of evaluating market transformation programmes: Experience in Sweden. *Energy Policy*, 29(1), 67–79. [https://doi.org/10.1016/S0301-4215\(00\)00100-2](https://doi.org/10.1016/S0301-4215(00)00100-2)
- Neij, L., & Åstrand, K. (2006). Outcome indicators for the evaluation of energy policy instruments and technical change. *Energy Policy*, 34(17), 2662–2676. <https://doi.org/10.1016/j.enpol.2005.03.012>
- OECD/DAC (- Network on Development Evaluation) (2019). *Better Criteria for Better Evaluation - Revised Evaluation Criteria Definitions and Principles for Use*. Retrieved from <http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>
- Parsons, D. W. (1995). *Public policy: An introduction to the theory and practice of policy analysis*. Aldershot, UK ; Brookfield, Vt., US: Edward Elgar.
- Patton, M. Q. (1996). A World Larger than Formative and Summative. *Evaluation Practice*, 17(2), 131–144. <https://doi.org/10.1177/109821409601700205>

- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (Third edition). Thousand Oaks, CA: SAGE Publications.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (Fourth edition). Thousand Oaks: SAGE Publications.
- Patton, M. Q. (2019). Transformation to Global Sustainability: Implications for Evaluation and Evaluators. *New Directions for Evaluation*, 2019(162), 103–117. <https://doi.org/10.1002/ev.20362>
- Pollitt, C. (1993). Occasional Excursions: A Brief History of Policy Evaluation in the UK. *Parliamentary Affairs*, 46(3), 353–362. <https://doi.org/10.1093/oxfordjournals.pa.a052425>
- Pye, G. (2018). Interdisciplinary Research and the Early Career Researcher. In M. Kumar & S. Pattanayak, *Positioning Research: Shifting Paradigms, Interdisciplinarity and Indigeneity* (pp. 34–54). Thousand Oaks, CA: SAGE Publications. <https://doi.org/10.4135/9789353282509.n3>
- Rip, A. (2016). The clothes of the emperor. An essay on RRI in and around Brussels. *Journal of Responsible Innovation*, 3(3), 290–304. <https://doi.org/10.1080/23299460.2016.1255701>
- Rip, A., & Kemp, R. (1998). Technological change. In S. Rayner & E. L. Malone (Eds.), *Human Choice and Climate Change, vol. 2* (pp. 327–399). Columbus, Ohio: Battelle Press.
- Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsı, T. A. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation*, 2000(87), 5–13. <https://doi.org/10.1002/ev.1177>
- Rogge, K. S., & Reichardt, K. (2016). Policy mixes for sustainability transitions: An extended concept and framework for analysis. *Research Policy*, 45(8), 1620–1635. <https://doi.org/10.1016/j.respol.2016.04.004>
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (Seventh edition). Thousand Oaks, CA: SAGE Publications.
- Rotmans, J., Kemp, R., & Van Asselt, M. (2001). More evolution than revolution: Transition management in public policy. *Foresight*, 3(1), 15–31.
- Rowe, A. (2019). Sustainability-Ready Evaluation: A Call to Action. *New Directions for Evaluation*, 2019(162), 29–48. <https://doi.org/10.1002/ev.20365>
- Sanderson, I. (2000). Evaluation in Complex Policy Systems. *Evaluation*, 6(4), 433–454. <https://doi.org/10.1177/13563890022209415>
- Sarewitz, D., Foladori, G., Invernizzi, N. & Garfinkel, M. S. (2004). Science Policy in Its Social Context. *Philosophy Today*, 48(Supplement), 67–83. <https://doi.org/10.5840/philtoday200448Supplement8>
- Schalock, R. E. (2001). *Outcome-Based Evaluation*. Dordrecht: Springer US
- Schoenefeld, J. J., & Jordan, A. J. (2019). Environmental policy evaluation in the EU: Between learning, accountability, and political opportunities? *Environmental Politics*, 28(2), 365–384. <https://doi.org/10.1080/09644016.2019.1549782>
- Schot, J., Hoogma, R., & Elzen, B. (1994). Strategies for shifting technological systems. *Futures*, 26(10), 1060–1076. [https://doi.org/10.1016/0016-3287\(94\)90073-6](https://doi.org/10.1016/0016-3287(94)90073-6)
- Scriven, M. (1991). *Evaluation thesaurus* (Fourth edition). Newbury Park, CA: SAGE Publications.

- Scriven, M. (1996). Types of Evaluation and Types of Evaluator. *Evaluation Practice*, 17(2), 151–161. <https://doi.org/10.1177/109821409601700207>
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice* (Reprinted). Newbury Park: SAGE Publications.
- Siedlok, F., & Hibbert, P. (2014). The Organization of Interdisciplinary Research: Modes, Drivers and Barriers. *International Journal of Management Reviews*, 16(2), 194–210. <https://doi.org/10.1111/ijmr.12016>
- Simons, H. (2009). *Case Study Research in Practice*. Los Angeles: SAGE. Retrieved from <https://doi.org/10.4135/9781446268322>
- Smith, A., Stirling, A., & Berkhout, F. (2005). The governance of sustainable socio-technical transitions. *Research Policy*, 34(10), 1491–1510. <https://doi.org/10.1016/j.respol.2005.07.005>
- Smith, A., Voß, J.-P., & Grin, J. (2010). Innovation studies and sustainability transitions: The allure of the multi-level perspective and its challenges. *Research Policy*, 39(4), 435–448. <https://doi.org/10.1016/j.respol.2010.01.023>
- Spicker, P. (2006). *Policy analysis for practice: Applying social policy*. Bristol: Policy.
- Stake, R. (2004). *Standards-Based & Responsive Evaluation*. Thousand Oaks, CA: SAGE Publications.
- Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: SAGE Publications.
- Stame, N. (2008). The European Project, Federalism and Evaluation. *Evaluation*, 14(2), 117–140. <https://doi.org/10.1177/1356389007087535>
- Stern, E. (2009). Evaluation policy in the European Union and its institutions. *New Directions for Evaluation*, 2009(123), 67–85. <https://doi.org/10.1002/ev.306>
- Suchman, E. (1967). *Evaluative Research: Principles and Practice in Public Service and Social Action Programs*. The U.S.A: Russell Sage Foundation.
- Sullivan, H. (2011). “Truth” junkies: Using evaluation in UK public policy. *Policy & Politics*, 39(4), 499–512. <https://doi.org/10.1332/030557311X574216>
- Swanborn, P. (2010). *Case Study Research: What, Why and How?* London: SAGE. <https://doi.org/10.4135/9781526485168>
- Swedish Energy Agency (2019). *Energiläget 2019*. Retrieved from <https://energimyndigheten.a-w2m.se/Home.mvc?resourceId=145396>
- Toulmin, S. (1966). The complexity of scientific choice II: Culture, overheads or tertiary industry? *Minerva*, 4(2), 155–169. <https://doi.org/10.1007/BF01584852>
- Tuominen, P., Klobut, K., Tolman, A., Adjei, A., & de Best-Waldhober, M. (2012). Energy savings potential in buildings and overcoming market barriers in member states of the European Union. *Energy and Buildings*, 51, 48–55. <https://doi.org/10.1016/j.enbuild.2012.04.015>
- United Nations General Assembly (2015). *Transforming our world: The 2030 Agenda for Sustainable Development* (No. A/RES/70/1). Retrieved from <http://www.refworld.org/docid/57b6e3e44.html>
- Vedung, E. (2010). Four Waves of Evaluation Diffusion. *Evaluation*, 16(3), 263–277. <https://doi.org/10.1177/1356389010372452>
- Vedung, E. (1997). *Public policy and program evaluation*. New Brunswick, N.J.: Transaction Publishers.

- Vedung, E. (2015). *Six Uses of Evaluation*. In V. Hennefeld, W. Meyer, S. Silvestrini, & R. Stockmann (Eds.), *Nachhaltige Evaluation? : Auftragsforschung Zwischen Praxis und Wissenschaft*. Festschrift zum 60. Geburtstag von Reinhard Stockmann (pp. 187-210). Münster ; New York: Waxmann Verlag.
- Verschuren, P., Doorewaard, H., & Mellion, M. J. (2010). *Designing a research project* (Second edition, rev. and ed. by M.J. Mellion). The Hague: Eleven International Pub.
- Vine, E., du Pont, P., & Waide, P. (2001). Evaluating the impact of appliance efficiency labeling programs and standards: Process, impact, and market transformation evaluations. *Energy*, 26(11), 1041–1059. [https://doi.org/10.1016/S0360-5442\(01\)00053-6](https://doi.org/10.1016/S0360-5442(01)00053-6)
- Wagenaar, H., & Cook, S. D. N. (2003). Understanding policy practices: Action, dialectic and deliberation in policy analysis. In M. A. Hajer & H. Wagenaar (Eds.), *Deliberative Policy Analysis—Understanding Governance in the Network Society* (pp. 139–171). Cambridge ; New York: Cambridge University Press.
- Weber, L. (1997). Some reflections on barriers to the efficient use of energy. *Energy Policy*, 25(10), 833–835. [https://doi.org/10.1016/S0301-4215\(97\)00084-0](https://doi.org/10.1016/S0301-4215(97)00084-0)
- Weiss, C. H. (1979). The Many Meanings of Research Utilization. *Public Administration Review*, 39(5), 426–431. <https://doi.org/10.2307/3109916>
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 1997(76), 41–55. <https://doi.org/10.1002/ev.1086>
- Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (Second edition). Upper Saddle River, N.J: Prentice Hall.
- White, H. (2013). The Use of Mixed Methods in Randomized Control Trials. *New Directions for Evaluation*, 2013(138), 61–73. <https://doi.org/10.1002/ev.20058>
- Wholey, J. S. (1996). Formative and Summative Evaluation: Related Issues in Performance Measurement. *Evaluation Practice*, 17(2), 145–149. <https://doi.org/10.1177/109821409601700206>
- World Business Council for Sustainable Development. (2009). *Energy efficiency in buildings: Transforming the market*. Retrieved from <http://www3.cec.org/islandora-gb/en/islandora/object/greenbuilding%3A141/datastream/OBJ-EN/view>
- Yin, R. K. (2009). *Case study research: Design and methods* (Fourth edition). Los Angeles, CA: SAGE Publications.

Appendix A – framework paper I

A theory-based interdisciplinary framework used for reviewing and analysing 20 research evaluations, commissioned and conducted in Sweden 2005-2017. The framework combines key insights adopted from transition research, sociology of science, and evaluation theory and is arranged according to three main categories: *operationalization of research fields*, *analysis as performed in evaluations*, and *assessment and use*. The framework is presented here has been adjusted to become a stand-alone list, while being presented as part of a written text in Paper I.

Operationalization of research fields

1. How has energy efficiency as a field been defined by the evaluations?
2. Assert by:
3. boundaries set for the subject, and
4. boundaries and alignment between different governance areas, such as research, policy, markets and regulation in the remits of the evaluations.
5. Assess the framing of the evaluand in relation to transformative efforts by asserting the placement of the evaluand in relation to (political) visions at work; and whether the boundaries of the evaluation allow a systems and multi-actor perspective.

Analysis as performed in evaluations

What is the formal constitution of evaluations with regards to:

6. methods applied for analysis - including types of indicators and measures of goal attainment, and how they conceptualize and measure scientific impact.
7. agency and forms of commissioning – which are the purposes for commissioning evaluations, and is there alignment with other missions and tasks?
8. role and placement of directionality – including if and how the measures of the evaluation are connected to other governance areas.

9. The framing of the evaluand's transformational efforts is assessed by the extent to which the analyses within the evaluations account for multiple governance areas or multiple actors, or heed experimentation.

Assessment and use

10. What are the processes of utilization and dissemination of the evaluations?
11. How has directionality – evaluation scopes and purposes – been operationalized and measured in the communication of evaluations?
12. How are linkages between research and other elements of meta-governance ascertained and assessed?
13. Which value criteria are used? How do assessments made in an evaluation connect to transitions of the energy system? Do selected value criteria support a systems perspective, or acknowledge experimentation and long-term outcomes?

Appendix B – framework paper II

A theory-based interdisciplinary framework used for reviewing and analysing 33 policy evaluations, commissioned and conducted in Sweden 2005-2015. The framework combines key insights adopted from transition research, evaluation theory, and policy analysis and is arranged according to three main categories: *methods*, *value judgements*, and *use of evaluations*.

Methods applied in evaluations

1. What methods are used?
2. How has the impact been assessed?
3. How is the counterfactual constructed?
4. Is the focus on just intended effect or are side-effects (intended and unintended) and rebound effects also considered?
5. If triangulation is used how has the synthesis been produced?
6. Is transition and the potential for transitions analysed? - Does the evaluation have a system perspective?; Does the evaluation have a multi-actor perspective?; Is the evaluation taking a visionary perspective (combined ex-post and ex-ante approach; part of a long-term evaluation approach)?; Is the evaluation capturing experimentation (radical policies and/or potential radical outcomes)?; Is the evaluation considering learning

Value judgments in evaluations

7. How and by whom have the criteria been decided (The organization commissioning the evaluation, the evaluator(s), by stakeholders, general evaluation policy)?
8. Which value criteria were used to judge the intervention? (e.g. relevance, effectiveness, efficiency, flexibility, predictability, persistence, acceptability, transparency, equity).
9. Do the value criteria reflect the interests of different groups?

10. Do the criteria used promote reflexivity and challenge established goals, needs and methods?
11. Is reflexivity part of the value judgement the conclusions are based on?

Use of evaluations

12. Have key stakeholders been identified and involved in the evaluation process?
13. Have there been any specific efforts to engage different groups, including those that are not well organized?
14. What has been the time frame for the use of the results?
15. What particular activities have been undertaken to facilitate use?
16. Have there been efforts to promote use beyond “intended use by intended users” by making the process open and transparent or by making the evaluation results/report freely available and easy to obtain?

Appendix C – Framework paper IV

Based on the characteristics of eleven models of use of evaluations, identified in evaluation theory, three statements were constructed to convey the essence of each model. These statements were scrambled and posed as individual statements to be rated by informants working with evaluation at Swedish state agencies, in a self-administered interview questionnaire.

Table 1. Eleven models of use were identified in evaluation theory. Based on the characteristics of each model, three statements were designed to capture the essence of what the model of use entails.

Model of Use	Characteristics
1. Instrumental	Show how the evaluated program should be changed
	Lead to immediate changes in the program that is being evaluated
	Contribute to knowledge development that creates immediate measures for improvement in the evaluated program
2. Enlightenment / conceptual	Increase the understanding of how a program should be implemented, rather than lead to concrete measures in the program
	Provide general knowledge about how a certain type of program works
	Provide an overview of which aspects (e.g., administration, behaviour, economy, market) are affected by a program
3. Legitimizing / Reinforcing use	Contribute to showing that we are doing things correctly
	Give trust and support for decisions that concern a program
	Confirm what we know about a research program or policy instrument
4. Interactive	Be broad and robust enough to be the only basis for decisions regarding the evaluated program (<i>reverse positive</i>)
	Only be one part of the knowledge basis in a decision process about a program
	Be used in combination with other material in decision-making
5. Ritual / Symbolic use / Mechanical use	Be conducted to show that the program at hand has been followed-up on
	Be performed because it is expected that an evaluation is performed
	Be conducted as per usual, so that actors affected by the evaluation will know what to expect
6. Mobilizing use / Persuasive use	Be used to encourage actors to support a programme or a viewpoint
	Be used to convince opponents
	Be used to create support among others for the evaluated program
7. Overuse	Focus on pre-determined criteria and indicators, regardless of whether the program or the situation has changed
	Be used for decisions about the program that are entirely based on what the evaluation shows
	Not put emphasis on describing why the results show what they show
8. Process use	Spur to change and improvement already during the <u>evaluation process</u> , through dialogues with different actors
	Primarily contribute with learning and knowledge during the evaluations process, through interactions between different actors
	Lead to important insights during the evaluation process
9. Constitutive / Anticipatory use	Realize effects as early as possible, before the evaluation itself is done
	Influence actors to consider what needs to be done to meet the expectations of the planned evaluation
	Spur improvements in a program by communicating that an evaluation is to be done to those whom the evaluation concerns
10. Tactical use	Buy additional time in a decision-making process
	Show that "something is being done"
	Prevent hasty decisions, by allowing the evaluation process to take time
11. Unintended use	Spur further discussions about other programmes
	Be used as a knowledge base to be used for issues that are outside of the evaluated program
	Indicate areas outside of the evaluation boundaries that need further investigation

Appendix D – Overview of reviewed research and policy instrument evaluations

Table 1. The complete sample of the 53 Swedish research and policy instrument evaluations that have informed papers I-III. The table showcases the distribution of evaluation type, commissioner and evaluator. Research institutions include: foundations, collaboration platforms and university departments. The asterisk (*) indicates that the evaluation has been both commissioned and conducted by the same actor. One policy instrument evaluation and two research evaluations were conducted jointly by multiple actors, therefore a larger number of conducted evaluations (n=56) than the number of commissioned evaluations (n=53).

Evaluations		Policy instruments (Total 33)				Research (Total 20)		TOTAL
		Financial (5)	Legislative (2)	Informative (18)	Other (8)	Programmes (15)	Institutions (5)	
Commissioned by	Swedish Energy Agency	2		10	5	13		30
	National Board of Housing, Building and Planning	2	1	1				4
	Ministry of Enterprise and Innovation	1		7	3			11
	Government		1				1	2
	FORMAS					1	2	3
	IQ Samhällsbyggnad					1		1
	MISTRA						2	2
Conducted by	Swedish Energy Agency	1*		2*				3
	National Board of Housing, Building and Planning	2*	1	1*				4
	External consultants	2	1	16	8	10	1	38
	Academia					5	3	8
	IQ Samhällsbyggnad					1*		1
	MISTRA						2*	2

Paper I



Paper II



ORIGINAL ARTICLE

Open Access

Transition governance for energy efficiency - insights from a systematic review of Swedish policy evaluation practices



Sofie Sandin^{1*} , Lena Neij¹ and Per Mickwitz^{1,2}

Abstract

Background: The transition towards a more sustainable energy system is urgent for addressing global environmental and social challenges, and will require transformative changes including improved energy efficiency in the built environment. To reach identified efficiency potentials, various policy instruments have been introduced but their effects are often unclear. In this paper, we argue that the outline of transformative policy strategies will require well-designed evaluations. The objective is to present a theory-based evaluation framework that can be used to assess existing evaluations, in order to support transformative policy strategies. The framework is also applied to provide insights from current Swedish evaluation practices.

Methods: The theory-based evaluation framework presented builds on evaluation theory, policy analysis and transition research and is arranged around *methods*, *value judgements* and *use* of the evaluations. Moreover, key aspects from transition research are included to provide guidance for transformative efforts in the evaluations. The systematic review presented in the paper is qualitative, covering 33 policy evaluations for energy efficiency in buildings in Sweden, commissioned by Swedish governmental authorities over a decade.

Results: The results of the review reveal a wide range of evaluations undertaken, using a sound methodological evaluation base that builds on a variety of methods for analysis, and application of multi-criteria analyses. Commonly, however, a rather narrow scope was applied and we note a missed opportunity for triangulation of findings. Key aspects for capturing transformative efforts, such as system-, scale- and multi-actor approaches, as well as visioning, experimentation and learning, were considered to varying extents, but could be more explicit and elaborate.

Conclusion: In all, we find the proposed theory-based evaluation framework useful for assessing and discussing both robustness and transformative efforts of current policy and evaluation practices. The review of the Swedish policy evaluations further indicates sound evaluation practices, and a foundational structure for identifying and analysing transformative efforts. To fully support transformative changes, we suggest a wider system perspective and a more thorough multi-actor approach and actor involvement in the evaluations. We also stress the need to further link evaluation theory with transition research to design evaluations that can support transformative changes in society.

Keywords: Policy evaluation, Energy efficiency, Buildings, Transition

* Correspondence: sofie.sandin@iiee.lu.se

¹International Institute for Industrial Environmental Economics (IIIEE), Lund University, P.O. Box 196, SE-221 00 Lund, Sweden

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Background

A more sustainable energy system is key for addressing global environmental and social challenges, and an imperative feature for its realization is increased energy efficiency in the built environment. The built environment alone accounts for a large part of global final energy use—32% according to the Intergovernmental Panel on Climate Change (IPCC) [1]—and the energy demand is expected to increase over the coming decades. At the same time, the sector holds an estimated energy efficiency potential worldwide of 50–75% and 50–90% in existing and new buildings respectively [1]. To realize these potentials, and to actualize a transformation towards a more energy-efficient built environment, policy interventions will be necessary to overcome market failures and to accelerate desirable changes in the socio-technical system [1–3].

Over the years, various types of policy instruments for energy efficiency have been introduced, including legislative instruments such as building codes and minimum standards; financial instruments such as subsidies and tax reductions; informative instruments such as campaigns and labelling schemes; but also new types of instruments such as technology procurement programs, voluntary agreements, actor network platforms and white certificate schemes [4–6]. By applying and combining various types of policy instruments, the goal is to achieve not only incremental improvements to a more energy efficient built environment but also to support a transition of the society, i.e. the introduction of new emerging technologies as well as changes in social norms, behaviour and institutional capacity.

Energy efficiency in buildings is a complex field, minted by multiple actors (e.g. designers, developers, contractors, engineers, owners, tenants, capital providers, etc.), institutions (e.g. authorities, regulations and norms) and technological factors shaping it [7]. Building type, age, use and structure varies, as do lifestyles of the tenants, which creates mixed incentives for undertaking of energy efficiency measures, at the same time as buildings intrinsically are slow to change due to long lifespans of measures [1, 7]. Altogether, the design of policy measures to achieve a transition towards a more energy-efficient built environment is a complex process that requires advanced knowledge on how various policy measures affect change.

To understand which policy measures to use and how to design successful policy strategies for transformative changes will require evaluations and a vigorous dialogue on how to evaluate [8–10]. Today, however, the methods and processes underpinning evaluations in the energy and environmental fields are often described as multifarious and largely suboptimal in terms of systematization [9, 11].

The objective of this paper is, thus, to propose a theory-based framework that can be used for assessing existing policy evaluation practices, with a view to enabling

evaluations to further support transformative policy strategies for energy efficiency. This framework is based on evaluation theory, policy analysis and transition research, which we will return to below. The objective is also to assess to which extent existing evaluations already today apply a transformative evaluation approach.

We apply Vedung's [12] definition of evaluation as a 'careful retrospective assessment of the merit, worth and value of administration, output and outcome of government interventions, which is intended to play a role in future, practical action situations'. Moreover, we extend the boundaries of this definition to not only include retrospective ex-post evaluations but also prospective ex-ante assessments, since both types are recognized in the European Commission's [13] guidelines for better regulation, as well as in the European Union's (EU) currently enforced 7th Environment Action Programme as means to 'improve environmental integration and policy coherence' (see also [14]).

In principle, a thorough evaluation approach that supports development of relevant knowledge as well as processes of change in society should build on essential theoretical knowledge-bases such as evaluation theory and policy analysis. In evaluation theory, attention is brought to the need for broad and reflexive methods [15, 16]; value judgements that reflect multiple stakeholders' concerns; and facilitation of use of evaluations through stakeholder involvement [12, 17–20]. Policy analysis brings forth a complementary emphasis on assessing the policy instrument's managing and administration, its role in a societal system (see, e.g. [21]), and the need for policy-mixes to enhance system transformations [22, 23].

Furthermore, if to capture and support more transformative changes in the energy system—as often argued for and described in the processes of motivating and designing new energy policy instruments—transition research could be a very valuable complement to the evaluation design. Some key concepts that are central in most of the transition literature, and that can provide insights to a policy evaluation, include a system-oriented, scale-oriented and multi-actor based approach, along with processes that motivate transitions, such as visioning, experimentation and learning (see, e.g. [24–27]).

At present, however, it is unclear to which extent existing policy evaluations and policy evaluation approaches focus on, and are capable of, capturing and supporting transformative processes for the achievement of an energy efficient society. To which extent do evaluations, either on their own or in combination with other evaluations, support the development of relevant knowledge, and to which extent do they rely on a theoretical knowledge-base, providing a comprehensive and rigorous analysis and assessment? In order to answer this, we review and assess 33 evaluations of policy instruments for energy efficiency in

buildings in Sweden. Based on the review, we discuss the evaluation practices applied today, and seek to identify means to further support the development and application of transformative evaluation strategies for energy policy. Sweden is chosen as a suitable case for the review due to it being one of the forerunners as regards policy evaluation, with an extensive evaluation practice spanning multiple sectors [8, 28].

The outline of the paper is as follows: the ‘Theoretical framework’ section provides a presentation of the proposed theory-based framework along with its theoretical underpinnings. The ‘Methods’ section presents the methods used for data collection and analysis when conducting the review of 33 policy evaluations. In the ‘Results’ section follows results drawn from the review, showcasing actual evaluation practices of policy instruments for energy efficiency in buildings in Sweden. In ‘Discussion’ and ‘Conclusion’ sections, we discuss and summarize our conclusions on how to further support a transformative evaluation approach and the development of relevant knowledge for a more energy efficient society.

Theoretical framework

In this paper, we present a theory-based evaluation framework designed to assess existing evaluations and their abilities to present a transformative and comprehensive evaluation approach for policy instruments. The framework is based on evaluation theory and is, moreover, complemented with insights drawn from the fields of policy analysis and transition research. While evaluation theory and policy analysis provide insights on methodological and contextual conditions of the evaluations, transition research provides conceptual approaches on how to describe the structure and processes that support transformative societal changes (see, e.g. [24, 29–31]). Although transition research is a field consisting of multiple sub-orientations, there are certain key concepts that are transversally pervading throughout the wide span of orientations. For this study, we have identified such key concepts and infused them into the theoretical framework. First, we identify and introduce the need for a system-, scale- and multi-actor approach in the evaluations to support transformative changes. Secondly, we identify the need for evaluations to capture processes of visioning, experimentation and learning. We will return to these concepts in the description of the framework below.

The framework is categorized in accordance with Alkin and Christie’s [17] evaluation tree, arranged around the three main theoretical branches of (I) methods applied in evaluations, (II) value judgments in evaluations and (III) use of evaluations. For each category, a number of sub-categories have been identified, referring to key issues from evaluation theory, policy analysis and transition

research, as described below. A full list of the categories and sub-categories that guide the review is found in Appendix 1.

Methods applied in evaluations

Methods applied in evaluations cover the entire process of evaluation, from theoretical points of departure to data collection and analysis. Methods for assessing interventions and, to some extent, data collection can largely be described as quantitative (e.g. statistical analysis and models) or qualitative (e.g. interviews, document analysis) [10, 32], and a mixed methods approach has been stressed by scholars [10, 32, 33] for the provision of comprehensive results. In relation, triangulation of data, methods, analysts and theories is emphasized for a systematic validation of consistency in findings from various approaches [34, 35], along with a robust counterfactual for attributing effects [36].

In the field of transition research, the application of a system approach is emphasized for assessing transformative change, along with the need to capture the interconnectedness of multiple components of the system [24, 26, 37]. Systems can be described in terms of technology, actors, networks and institutions which are all inter-linked. For the case of energy efficiency and an evaluation point of view, this would require a consideration of norms and institutions, actors and their behaviour and interlinkages and dependencies within the societal system in which the evaluated energy policy is enforced. Moreover, the evaluation should encompass *multiple* actors (e.g. authorities, organizations, businesses and individuals) that hold a stake in or are affected by the energy policy, acknowledging that some actors may be driving transformative changes while others may be counteracting it [27]. It would also require an outlook to any technological factors that may challenge or change the current socio-technical system configuration. Related to this, scale is seen as another key component, often highlighting the importance of niches—hubs where small-scale experimentation can be carried out—and their protection through active shielding, nurturing or empowering efforts [38]. As an example, niches within energy efficiency in buildings may be small test-beds for, e.g. new building materials or appliances, which are supported by policy instruments that facilitate their activities.

Transition research also emphasizes processes that motivate and drive transitions, such as processes of visioning, experimentation and learning. Visioning is emphasized for its guiding of transitions by, for example, providing common goals and inspiring and activating actors [27, 31, 39]. Visions related to energy efficiency may be found in transnational goals, such as the EU energy and climate targets, and in nationally determined goals, thus relating to the institutional components of

the system-perspective laid out above. Capturing of visioning within an evaluation may, however, not only include the investigation of efforts made towards reaching such goals, but may also include the assessment of future outcomes and their contribution to societal changes, calling for a long-term perspective [30, 31, 40].

In addition to visioning, experimentation and learning are recognized as essential transformative drivers [25, 27, 40, 41]. Learning is here viewed as the acquiring of new knowledge, for example derived from experiences and experimentation, which leads to improvement of some kind, through actions and policy decisions, or through alterations of paradigms and ideas [42–44]. Experimentation for energy efficiency in buildings can include physical innovation in terms of, e.g. building materials and technological equipment, but arguably also innovative policy design. Thus, the capturing of experimentation and learning in the evaluation of energy policy instruments requires an investigation of innovative efforts both in terms of policy design, its ability to facilitate or favour experimentation, and its intended outcome. Lastly, to focus on transformative efforts, forces upholding the status quo also need to be acknowledged, along with the efforts geared towards disrupting those [45].

To summarise this category, we review the traditional evaluation methods applied, approaches to assess impact and the construction of counterfactuals when performing impact evaluation, along with the assessment of side-effects, rebound effects and triangulation. Moreover, the potential to capture transformative efforts within the evaluations is assessed by identifying the use of a system-wide perspective, scale, a multi-actor approach, visioning, experimentation and learning (see Appendix 1).

Value judgements in evaluations

The decision regarding which value criteria to be employed may be guided by actors in a descriptive approach, departing from the wide range of values held by stakeholders, or be determined by the evaluator in a prescriptive manner, guided by particular, justified values [46]. Criteria traditionally concern policy outcomes (effects) and goal attainment [10, 12, 33], but may also concern efficiency (cost-effectiveness or cost-benefit), relevance, flexibility, persistence or predictability [47]. They can also relate to democracy and social justice [48], by investigating, e.g. acceptability, transparency and equity [47], or may, as emphasized within policy analysis, concern consistency, coherence, credibility and comprehensiveness [23]. In addition, value judgements should express reflexivity regarding established goals needs and methods—frequently called double-loop learning [49]—which refers to the examination of contributions and implications of a policy program's goals on a societal level, with regard to how they serve to improve societal welfare

(see, e.g. [15]). From a transition research perspective, a particular focus is placed on different interests held by different stakeholders, and their role in the policy process, as captured by a multi-actor approach.

To summarize this category, we review the selected criteria used in the evaluations, how these criteria reflect the interests of different groups, by whom they have been decided and whether the evaluation expresses reflexivity in terms of challenging established goals, needs and methods (see Appendix 1).

Use of evaluations

Making use of evaluation results is paramount in evaluation theory [12, 18, 50]; however, the drivers for conducting an evaluation in terms of its utilization vary, and span accountability, learning and political nature [12, 14]. Ideally, an evaluation should capture and address the concerns of various stakeholders—not only the decision makers—and this calls for an elaborate evaluation design that is able to heed multiple kinds of uses [20].

The making use of evaluation results, moreover, hinges upon the timing of the evaluation, especially so in a policy field like energy efficiency that is rapidly developing [51]. As such, the design of the evaluation should mirror the timeframe of the use, where more rapidly executed evaluations may be favourable over more extensive ones at certain occasions, and vice versa.

From a utilization perspective, the key components adapted from the field of transition research are a multi-actor approach and learning. Learning through its role in contributing to improved policies and processes [42, 44], a multi-actor approach for the capturing of how different groups affect or are affected by a transition and for spurring increased utilization of evaluation results by focusing it on relevant questions and by enabling commitment [12, 18]. Thus, it is important to assess what is done in the evaluation process to increase the accessibility of the evaluation to user groups which are not the obvious participants.

To summarize this category, we review the potential use of evaluations by investigating the identification and involvement of stakeholders in the evaluation process, both in terms of partaking in evaluation design and for data collection, and by other activities undertaken in order to facilitate further use. We also review the time frame for the use of the results (see Appendix 1).

Methods

Assessment and review approach

The assessment approach applied in this paper builds on the theory-based evaluation framework presented in the 'Theoretical framework' section, which gathers core aspects to be considered in order for evaluations to be able to provide essential knowledge in complex fields

such as energy efficiency. By reviewing evaluations conducted for energy efficiency policy instruments in Sweden, following the theory-based evaluation framework, the intention is to identify the extent to which the aspects highlighted in the framework are applied already today in practice.

The review is designed as a qualitative, systematic review of existing policy evaluations. The use of a systematic review for analysing existing evaluations is an established method, but rather than focusing on synthesizing results from the evaluations (see, e.g. [52, 53]), we instead place focus on the conduct of evaluation: focusing on methods, value and use (see ‘Theoretical framework’ section). Each of these three framework categories has a number of sub-categories attached, which creates a protocol to guide the review (see Appendix 1). The review covers 33 Swedish policy evaluations, each of which has been closely read and qualitatively analysed according to the framework. When reviewing, we have identified methods, value judgements and aspects of use, as they have been presented and described in the evaluation reports. Thus, we have avoided interpretation of the evaluation content. Due to the scoping of the study, we have not conducted interviews with stakeholders such as evaluation commissioners, evaluators or other stakeholders involved in the implementation of the evaluated policy instruments.

Evaluations performed in Sweden and evaluations included in the review

Over the last decade, a variety of policy instruments aimed at advancing energy efficiency in buildings have been introduced in Sweden, and many of these policy instruments have been evaluated to form the basis for decision making and reporting. Although policy evaluations may be conducted by a range of actors—governmental, non-governmental and researchers—we include in this study only those evaluations that were directly commissioned by governmental authorities. This is in line with the aim of the study which is to assess the current governmental evaluation practices in Sweden, and the extent to which such evaluations have been used to inform transformative policies for energy efficiency. The evaluations reviewed in this study were selected using the following criteria:

- The evaluated policy instrument was aimed at energy efficiency in buildings
- The evaluation was commissioned by governmental authorities
- The evaluation was commissioned between the years 2005 and 2015

The evaluations were collected from state websites, national reports and interpersonal contacts with Swedish

governmental authorities (e.g. the Swedish Energy Agency; the National Board of Housing, Building and Planning) as well as with external consultants who conducted evaluations. All evaluations collected that met the selection criteria were reviewed. Thus, this study presents the review of 33 evaluations of policy instruments conducted or commissioned by Swedish governmental authorities in the period 2005–2015. The following policy initiatives are included:

- Legislative instruments (two evaluations): revision of building codes and energy requirements.
- Financial instruments (five evaluations): subsidies and tax reductions for energy efficiency measures and subsidies for performing energy audits.
- Informative instruments (18 evaluations): demonstration projects, municipal energy advisory programs, energy performance certificates for buildings, online information portals, facilitating energy services and public authority best practice.
- Other instruments (eight evaluations): technology procurement programs and cooperative network programs for energy efficiency in buildings.

Moreover, a number of evaluations were collected that were deemed to not fully meet the selection criteria. These evaluations concerned policy instruments that were only partly or indirectly connected to energy efficiency in buildings, and were therefore not included in the review:

- Evaluations of multi-sectoral policy instruments: the contribution of the policy instrument to energy efficiency in buildings was not evaluated specifically.
- Assessments of policy mixes using energy system models; economic modelling; or bottom-up/top-down effect calculations: these evaluations contained a number of energy policy instruments that were assessed in combinations, commonly spanning different sectors, of which not all directly referred to energy efficiency in buildings. Examples of such policy instruments include energy taxes, carbon taxes and policies targeting fuel substitution.

Thus, the final selection of policy evaluations that was reviewed within the scope of this study places focus on evaluations conducted of individual policy instruments. An overview of the final sample is found in Table 1, showcasing the distribution of type of evaluand, commissioner and evaluator. A full reference list of the evaluations, both those included in the review and those that were not, is found in Appendix 2.

Results

Below, we present results drawn from the review of evaluations, and assess the extent to which the Swedish

Table 1 Overview of review sample of 33 evaluations, showcasing distribution of policy instrument type, commissioner and evaluator

Total		Financial	Legislative	Informative	Other	Total
		5	2	18	8	33
Commissioned by	Swedish Energy Agency	2		10	5	17
	National Board of Housing, Building and Planning	2	1	1		4
	Ministry of Enterprise and Innovation	1		7	3	11
	Government		1			1
Conducted by	Swedish Energy Agency	1*		2*		3
	National Board of Housing, Building and Planning	2*	1	1*		4
	External consultants	2	1	17	8	28

The asterisk (*) indicates that the evaluation has been both commissioned and conducted by the same governmental authority. Two of these evaluations were, moreover, conducted jointly by both the Swedish Energy agency and external consultants, hence the number of conducted evaluations amounting to 35

evaluations of energy efficiency policies apply the components of transformative evaluations, presented in the theory-based evaluation framework in the ‘Theoretical framework’ section.

Before going into depth in the review, we start with providing some context as to the reviewed evaluations. As outlined above, all 33 evaluations were commissioned by governmental authorities. Two of these authorities—the Swedish Energy Agency and the National Board of Housing, Building and Planning—also conducted evaluations, whereas the majority were conducted by external consultants, as illustrated in Table 1. The external consultants amount to eight different firms, their contribution in the evaluations spanning from data collection, compiling of reports, to the full execution of the evaluation.

The stated purposes for the evaluations within the reviewed sample spanned many objectives: many evaluations focused on outcomes and impacts, other evaluations focused more on process-related inquiries, such as administrative burdens, the organizations involved in the implementation, or views of different stakeholders. In all, the foci of the commissioned evaluations were relatively limited in scope and time, and although the

evaluations provided learning on essential elements, none of them explicitly expressed a purpose to advance knowledge supporting a broad transformative change in society. Within the review sample, certain policy programs were recurring through yearly evaluations, such as the municipal advisory program, or through evaluations conducted mid-term and at the end of an implementation, providing insights over time.

Below, we present the findings from the review following the three main categories: methods applied, value judgements and use.

Methods applied in evaluations

A key challenge in evaluating energy efficiency policies is how to provide credible results and vital learning, and an important aspect of this challenge is the choice of methods to be applied in the evaluations. The evaluations of energy efficiency policies covered in this study show a respectable mix of methods used to assess the results of the policy instruments (Fig. 1). The majority of the evaluations were based on two (17/33) or three methods (12/33); the most frequently used being

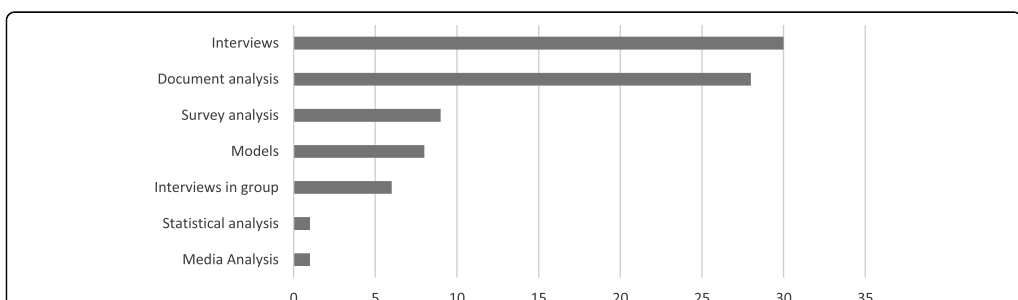


Fig. 1 Methods used for analysis in the reviewed evaluations ('Interviews in groups' is a method employed to gather several different types of stakeholders for interviews and discussions regarding early evaluation results)

interviews (30) and document analysis (covering, e.g. applications, regulations, statutes) (28).

Methods for data collection and analysis overlap in surveys, interviews and document analysis. Surveys, and in some instances document studies, were commonly used to provide data used for both qualitative analysis (e.g. opinions, functions of a policy instrument) and quantitative analysis (e.g. number of measures taken, number of applications). The purely quantitative methods for analysis that were used in the sample were calculations and statistical analysis. Altogether, the reviewed evaluations leaned towards qualitative methods and assessment of effects of a 'softer' nature, geared towards influencing stakeholders in taking energy efficiency measures and facilitating knowledge transfer between actors. Although the review does indicate a use of multiple methods, triangulation as a tool for systematic control of the consistency and validation of results in general was scarce. The combination of different data sources or methods was not commonly used for testing the consistency of findings, but rather as complements to each other, for providing input to specific questions.

Both evaluation theory and the policy analysis literature stress a counterfactual construction for the assessment of the actual impact of any policy intervention (see, e.g. [36]). The review shows that the most prevailing type of counterfactual analysis applied in the evaluations of energy efficiency policies in Sweden was derived from interviews or surveys (14/15), where evaluators sought to determine the additionality by asking stakeholders for their opinion on the extent to which the policy instrument had affected their decisions and actions. The applications of calculated counterfactuals and reference scenarios were less commonly used (3/15). Two of the reviewed evaluations used multiple methods for the construction of counterfactuals, which were synthesized from two and three methods respectively, the latter combining surveys, calculations and a reference group [54, 55]. The use of a reference group was only seen in this one evaluation, which concerned a subsidy for energy audits. Lastly, the choice of methods for constructing the counterfactuals was, however, rarely discussed in the reviewed evaluation reports. Multi-actor involvement occurred in the interviews or surveys, but reflections relating to the selection of actors to be involved, or the number of respondents needed for constructing a robust counterfactual were not apparent in the written reports—but may of course have been undertaken during the evaluation design process.

To provide knowledge on transformative changes in society, and the potential drivers and barriers for energy efficiency, a system-wide evaluation approach is required. Policy analysis specifically brings forward the assessment of aspects such as side-effects and rebound effects. Among the reviewed evaluations, only one incorporated a more thorough consideration of side-effects, by identifying

the environmental impacts and assessing their costs [56]. Eight other evaluations mentioned side-effects, or rather co-benefits, such as marketing and competition advantages, altered value of buildings as an effect of energy efficiency measures and the creation of joint platforms for knowledge exchange between authorities and businesses. Rebound effects were mentioned in one evaluation, a matter of increased energy efficiency leading to increased energy use due to improved comfort [57].

Transition research emphasizes a system and scale-oriented perspective. Such an assessment can be designed in many different ways, including for example the assessment of system components such as actors, institutions and technological factors. The review of the Swedish evaluations shows that many evaluations focused on actors (20/33): on outcomes and effects that influenced different actors in their actions (e.g. the municipal advisory program or the energy audits), or their knowledge acquisition (e.g. within the collaborative network programs). The actors that were considered included authorities, beneficiaries, organizations and companies. Institutional aspects were considered in seven evaluations, of which four referred to evaluation of the building codes [56, 58], the act on energy declarations in buildings [59] and the regulation on support for investment in energy efficiency [60]. It should be mentioned that institutional factors, such as overarching regulations and institutions (both national and on an EU-level), were mentioned in some evaluations with regards to how the evaluand related to or aided in fulfilling overarching initiatives. For example, 11 evaluations—performed by the same external consultant—schematically mapped the evaluand in relation to such institutions, but did not elaborate further on potential effects of their interactions. Technological factors, or the role of new technological innovations, were seen in nine evaluations. These evaluations mainly concerned policy programs for technology procurement, demonstration programs and the cooperative network program LÅGAN. One evaluation concerning a tightening of building codes acknowledged technological factors from the perspective that the current level of technology merited a sharpening of the regulations [56]. In all, the reviewed evaluations assessed actors, institutions and technological factors, but did not elaborate further on their interactions within the system.

Transition research also advocates for a multi-actor approach, to capture the potential and the effects of transformative interventions. For the particular case of energy efficiency in buildings, such actors include private and public house owners and tenants, the construction industry, other businesses and organizations either engaged in providing or adopting energy efficiency measures and authorities and municipalities. In the reviewed sample, 31 evaluations incorporated stakeholders to some extent, and

approximately half of these evaluations (16/31) did take a multi-actor approach, involving two or more groups either targeted or otherwise involved in the implementation of the policy instrument. These were commonly authorities, beneficiaries and representatives from businesses and organizations. Discussions about actor groups that were consequently not involved were uncommon within the reviewed material, but may be a valuable discussion for opening up the evaluation boundaries.

Transition research, moreover, brings forward processes of visioning, experimentation and learning as central in the analysis of transformative changes. Visioning may be captured through a long-term evaluation approach. In the reviewed documents, we did not find such approaches, albeit some recurring evaluations of the same policy instruments, especially in the cases of the municipal energy advisory programs, cooperative network programs and subsidies for energy audits. Visioning may also be captured and supported in evaluations through a combination of ex-ante and ex-post evaluations; one evaluation [55] was constructed as such, creating an initiative for both learning and prediction of future outcomes simultaneously.

In the assessments of experimentation and learning, we looked for the acknowledgement of experimental efforts in terms of innovative policy and potential outcome, as well as consideration of facilitation for experimentation. In all, we found four evaluations that took an experimentation criterion into account. These evaluations did not showcase different foci than did other evaluations in the sample, but differed in that they concerned policy instruments that may be more prone to experimentation. Three of them, concerning cooperative network programs [61], technology procurement [62] and a demonstration program for passive houses [63], stated that the evaluated instruments provided platforms for experimentation that facilitated the development of new energy efficient technologies, including, e.g. new building materials, windows and energy steering systems. A number of examples of successful technologies were presented within particularly the two latter evaluations, to showcase good practices that had been nurtured through the implementations. Another key feature that was mentioned within these evaluations was the need to take risks and perhaps fail in order to learn. Related to experimentation was, moreover, the upscaling of new innovations, which to a certain degree was captured in these evaluations through their assessments of innovative projects: from their inception to making an impact on a larger scale. Although these are important components for capturing transformational efforts, the notion of experimentation was, in these cases, embedded within the policy instruments, and as such a required evaluation angle.

Conversely, the fourth evaluation showed a differing assessment regarding experimentation. This evaluation

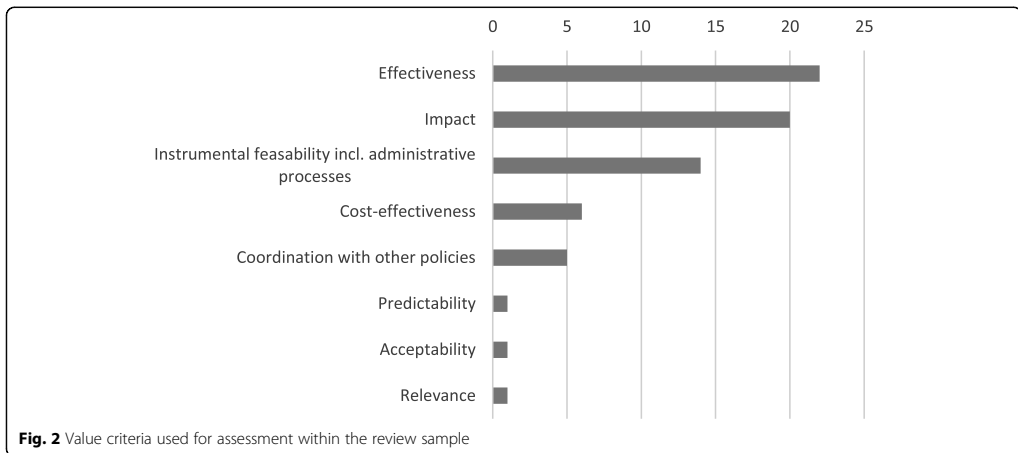
concerned a policy instrument aimed at providing investment support for various measures for energy efficiency [54], and stated that the instrument had been unsuccessful in providing support for experimentation relating to unconventional technology and technology development. This is indeed an important note, since it draws attention to the fact that the policy instrument in question was failing in breaking new ground, which in itself is a valuable insight to be drawn from the evaluation results. It should, however, be emphasized that not all policy instruments allow experimentation to the same extent; a program for technology procurement or a demonstration project holds greater potential to spur experimentation than do, e.g. regulations or energy audits. The lack of acknowledgement of experimentation thus needs to be regarded on the basis of the particular policy instrument's context and objectives.

Another important aspect of learning to be drawn from a policy evaluation is the identification of potential lock-in effects and path dependencies, i.e. to analyse how the current system configuration, with its norms and current technologies, is affecting and potentially hindering stakeholders in pursuing a more energy efficient pathway. It would also be valuable to uncover potential efforts to disrupt such configurations, e.g. to question norms or other factors that have an inhibiting effect on energy efficiency uptake. In the review sample, such discussions were not apparent, apart from the reflection on the (lacking) ability of a financial policy instrument to support development of new technology presented above [54].

Value judgments in evaluations

Evaluation theory states that the nature of evaluation is normative, and that assessments require a value base and criteria for valuation. Valuing is, however, broader than just criteria; it is also about the legitimacy of the value claims, which is closely related to the involvement of multiple actors, social justice and reflexivity. It is crucial to take all these aspects into account in order to evaluate and support transformative processes. The potential success of policy instruments for energy efficiency in buildings, as well as the realization of transformative changes in society, will very much rely on values held by the many involved actors. For this reason, a multi-criteria evaluation approach is often advocated [47], which is able to illuminate the implementation from various perspectives and angles.

In the reviewed evaluation reports, various criteria for valuation have been applied; the number of criteria ranging from one to four. The most frequently used criteria in the reviewed sample were effectiveness and impact (Fig. 2). These criteria partly evolved around elements that could be measured in, e.g. saved amount of kilowatt hours or in monetary terms, but were also of a softer nature,



concerning, e.g. impacts on stakeholders' actions in taking energy efficient measures, etc. The third most commonly applied criterion was instrumental feasibility and assessments concerning administrative processes (13/33).

In order to support transformative changes and efforts, evaluation criteria should not be limited to just impacts and effectiveness, but should also include the drivers for change and their implications. Such criteria may be more process-related and aimed at the mechanisms behind a successful implementation, such as acceptability, relevance and coordination with other policies. While relevance as a criterion was scarcely applied within the reviewed evaluations (1/33), five evaluations explicitly assessed policy coordination with other, similar, policy instruments, thus proving valuable insights on potential synergies or conflicts.

A broader assessment may also aim to capture various viewpoints of stakeholders. Energy efficiency policies may affect different stakeholders in different ways, leading to issues related to who benefits and who does not from the policy instrument. Discussions concerning legitimacy of value claims were not apparent within the written reports, leaving issues concerning the evaluation approach, value constructions or value dissonance—different values held by different stakeholders—in the evaluation process unmentioned. Moreover, it is also important to cover aspects of reflexivity in the evaluations in terms of challenging established goals and needs. Such reflexive elements were found in eight evaluations, which presented discussions concerning the instrument's goals or design, or the intention underlying the instrument. These evaluations covered instruments of the informative (3) and financial (3) kinds, as well as a procurement program and a co-operative network program. Concrete examples include the evaluation of a subsidy for house owners to change

windows [57], which questioned the need for the policy instrument due to lacking additionality, and an evaluation of the municipal advisory program [64], which questioned the focusing of the advisory program to a wide range of actors, instead of focusing it on a more limited range with high saving potential.

Use of evaluations

Evaluation theory emphasizes that evaluations are undertaken in order to be used, and stakeholder involvement in the evaluation process is essential in order to enhance use. Involvement of stakeholders will help gear the evaluation towards issues of importance that can lead to essential learning, and stakeholders should be involved in the evaluation process to focus it, to make it timely, to participate in decisions on methods and data collection, in interpretation of findings and to influence value judgements [12, 19]. From the perspective of transition research, the involvement of multiple actors is, furthermore, emphasized for the understanding of various actors' roles in driving a transition [27]. In this paper, we thus argue that for a realization of an energy transition, that in part is fuelled by strong energy efficiency policies, a reflection on use of results, actor involvement and learning in evaluations is essential.

As illustrated in Table 1, the predominant way for Swedish governmental authorities to proceed in the undertaking of a policy evaluation was to commission external consultants for conducting the evaluation (26/33). There are examples of the authorities conducting evaluations as well: in the role of both commissioner and evaluator (4/33); as commissioned by the government (1/33); and in more extensive collaborations between both authority and consultant (2/33).

In terms of involvement and use, it is not apparent within the written reports to which extent the commissioners were involved in the design of the evaluation. Following the notion that actor involvement may increase and facilitate use of evaluation results, we looked further into actor involvement in the reviewed evaluations. The review showed that the means of actor involvement was dominated by involvement through interviews, with the number of respondents per study ranging from one to 700. Discussion about the scope and size of the group of respondents was uncommon in the reviewed evaluations, and the involvement was largely limited to data collection rather than for facilitating further use. Evaluations conducted by one consultant included seminars where a group of selected stakeholders were invited to discuss preliminary results—the interviews in groups. While this may have had a facilitating effect on future use of the results, it was seemingly not the primary aim. Although a clear strategy for facilitation of further use was not clearly stated in the reviewed evaluations, it may have been part of either the design process or otherwise carried forward after the closing of the report. It should also be mentioned that the majority of the review sample was freely available on state websites, which in itself may increase use.

Lastly, returning to the foci of the reviewed evaluations, a number of evaluations had a stated objective of providing knowledge that could be used for decisions for further financial support of the program, and for improvements of the next phase of a policy program, as such indicating direct areas for use. From a learning perspective, in creating a knowledge base and capturing new knowledge, the review revealed there was an ex-post emphasis in the evaluation practice, as the absolute majority of the reviewed evaluations were conducted in retrospect (30/33). Many evaluations concluded with a set of recommendations or suggestions for further improvements of the policy design, thus summarizing and emphasizing what had been learnt from the evaluation of the implementation. If used and conveyed properly, this may serve as a valuable source of knowledge for how to improve current and new designs of policy instruments for accelerating energy efficiency in buildings.

Discussion

Achieving energy efficiency and transformative changes in the built environment entails addressing multiple components of a complex system, and without proper evaluation of policy instruments, they risk being ineffective. In this paper, we argue that for evaluations to be able to inform about transformative effects, a sound evaluation methodology is essential along with a comprehensive evaluation approach that is able to capture transformative efforts of various scales.

The review and assessment of the Swedish evaluations, first of all, show an impressive range of evaluations undertaken, with multiple evaluations conducted yearly. Moreover, we see a general application of sound evaluation approaches and, to some extents, evaluation aspects supporting transformative processes of change. The sample varies both in terms of evaluation foci and scopes: while some evaluations were rather extensive and evaluated a number of different aspects of the implementations, others were significantly shorter, narrower and were underbuilt by a limited amount of data. On this note, we acknowledge that the aims of the commissioning of evaluations may be based on limited time and resources, and that evaluations therefore differ in scope.

In all, the sample showed a lean towards qualitative methods. Thus, the sample indicates that the individual, mainly qualitative evaluations of this study, provide a good complement to other purely quantitative schemes for monitoring and evaluation, such as indicators and modelling assessments, which commonly focus on policy packages (see, e.g. excluded evaluations in Appendix 2). The frequent combination of methods in the sample, notably interviews and document studies, moreover showed that the evaluations frequently do apply several methods. In these evaluations, the document studies provided insights about, for example, how steering documents, application forms and reports supported or shaped the intervention, while the interviews provided deeper understanding about the workings of the policy, how it had been perceived and received and where there was need for further tuning or development. In combination, they provided valuable insights on the functions and effects of energy policies, and were transparent in their representation of the stakeholders that were involved.

Nevertheless, triangulation of methods for systematic validation of findings was rare, contrasting the predominant application of multiple methods and data sources for the conduct of the same evaluation. Thus, we see an opportunity to apply these methods more flexibly to allow triangulation for validation of findings, and to cross-fertilize and potentially gain additional information when sources are cross-checked. Of course, this is a matter of time and resources as pointed out above, but since the foundation within the evaluation practices largely is in place, the application of triangulation does not necessarily require extensive amounts of additional resources, while presenting a potential to provide added value and robustness to the evaluation results.

The same argument holds for the application of counterfactual constructions for assessing the impact of the policy instruments, where we see that the practices may be strengthened further by combining different methods, by discussing the limitations of the counterfactual and by deliberately including a wide range of actors in its

construction. A good practice to be highlighted from the review was the evaluation combining three different methods for synthesizing a counterfactual, including actor involvement through a reference group [55].

A last note on the methods applied is that, as discussed by Vedung [12] and Weiss [10], the purpose of the evaluation guides the evaluator towards selecting appropriate methods. As we have seen, the methods applied in the reviewed evaluations do, to certain degrees, incorporate key aspects from transition research, even though none of the reviewed evaluations had a stated purpose of investigating the transformative potential of the policy instrument. Thus, by bringing transition to the focal area of the inquiry, the methods applied can also be further strengthened and developed to investigate transformative contributions of an energy policy instrument.

As regards value judgements within the evaluations, the majority of evaluations in the review applied two or more criteria, as such taking a wider approach. However, the criteria were often limited to effectiveness and impact, whereas criteria for investigating mechanisms of the implementation that aided or counteracted the processes or outcomes of a policy instrument, such as acceptability and predictability, were underrepresented. As was assessments of relevance, a criterion that has been recognized as a key criterion for environmental policy evaluation [13, 65]. Consequently, evaluations that opened up for a broader scope and set of criteria, often combining assessments of outcomes and impacts with assessments regarding instrumental and administrative feasibility or coordination with other policy instruments, gave rise to deeper understanding of the workings of the implementation. These evaluations are, moreover, also potentially more prone to be informing about transformative contributions of the evaluand, since they to a larger degree could engage in discussions regarding the design and processes of a policy instrument, and how they link to effects on energy efficiency in the built environment.

Turning then to the evaluations' application and acknowledgement of key aspects for capturing contributions to an energy transition, the reviewed evaluations showed that assessments of changes on a system level were limited, and predominantly characterised by scattered evaluations of side-effects and rebound effects. An evaluation focus on the potential transformation of the current institutions, technological factors or the participation and behavioural aspects among actors was scarce, if treated at all, neither was evaluation of the interaction between these different system components. As for actor involvement, good practices were seen, e.g. in six evaluations that used interviews in groups as a method for discussing results and issues, which allowed for a deeper participation and discussions among actors. Otherwise, actor involvement was largely limited to data collection,

where selected groups were asked to partake in surveys, interviews or workshops. Generally, the evaluations did not explicitly explain to which extent stakeholders such as beneficiaries, businesses, organizations and authorities had been invited to partake in decisions regarding evaluation design and the selection of criteria or methods to be used in the evaluation processes, but we acknowledge that such discussions may have been undertaken outside of the final evaluation report. Nevertheless, transitions are not isolated events, but are complex processes involving and affecting multiple actors and system components (see, e.g. [24, 27], and without accounting for these factors, the evaluations risk becoming isolated snapshots of effects that are not put in relation to other efforts or policy initiatives. Thus, if evaluations take a deliberate systems and a multi-actor approach, they may be better equipped for a deeper analysis of whether the policy instrument is in fact delivering crucial contributions to a transition.

As for the process-related aspects of a transition, such as the acknowledgement of visioning, experimentation, upscaling and learning, our review showed that these were featured in some evaluations, through combined ex-post and ex-ante assessments, or through discussions about how the evaluand had promoted experimentation or otherwise capitalized on learning. In large, however, such assessments were subordinated, which is possibly explained by the fact that the evaluations' foci were not aimed at them. As was previously mentioned, we thus see great potential in deliberately gearing the evaluations' foci towards capturing transformative efforts, and thereby prompting an adaptation of methods to include key concepts for a transition.

Furthermore, many policy instruments for energy efficiency in buildings are aimed at encouraging a change of behaviour among citizens, organizations and authorities alike. Energy efficiency in buildings is, thus, ultimately not only a matter of new technological innovation and improvement of building materials and equipment, but certainly also a matter of altering our behaviour and how we live in and interact with our buildings—something that was frequently overlooked in the reviewed evaluations. This calls to be addressed if evaluations are to fully provide multi-faceted information that can support the highly multi-faceted challenge of realizing an energy transition.

Since the study at hand has been scoped by close reading of existing evaluation reports, a further investigation of the evaluation design processes is suggested for future research. This could include interviews with commissioners, evaluators and other stakeholders involved in the implementation, and would allow an extended analysis of the current evaluation practices with regards to underlying decisions and commission requirements. An extended analysis could also include evaluations commissioned and performed by a broader span of actors, such as academia

and stakeholders in the private sector. While we in this study focused on evaluations commissioned by governmental authorities in order to facilitate an assessment of how these evaluations in particular are performed with regards to assessing transformative contributions, an extended analysis could further provide insights about how policy evaluations from different sectors can strengthen each other in order to support transformative processes of change.

Conclusion

Evaluation holds great potential to support the realization of energy efficiency potentials in the built environment, by informing policy makers in designing and deciding upon strong and effective policy instruments. If evaluations are equipped to address the complexity of the energy system, and the transformative processes of change in the system, essential insights will be gained to accelerate energy efficiency. Supported by the theory-based framework, our review shows that the current evaluation practices have the structures in place to certain degrees for capturing transformative efforts and effects on a system level. In order to fully harness the potential of evaluations, however, some areas are outlined that may be further strengthened.

First, this study shows that numerous policy evaluations are performed, but their use is fragmented and lacking in coordination building up to an overarching transformative evaluation approach. To address this, we propose the application of an evaluation framework that can guide the structuring of evaluation designs so that they, in combination, can provide a more comprehensive overview of effects from current implementations at work.

Secondly, in order to capture transformative efforts, evaluations should more actively gear the focus towards effects on the system level, acknowledge the scale of change, stress a more thorough multi-actor approach and include actor involvement in the evaluation designs. Based on the review, we conclude that such aspects are present in various combinations and extents in the evaluations, but do not yet effectively assess interlinkages and effects on a grander scale. A further development of these evaluation components can enable a more transformative approach of the evaluation. In addition to this, evaluations targeting transformative changes should also seek to address drivers for change, such as visioning, experimentation and learning. The review suggests that these aspects were acknowledged in certain policy evaluations; visioning through a long-term evaluation approach; experimentation and learning seemingly dependent on the type of evaluation and its ability to be experimental. Nonetheless, these drivers are key components if to inform about whether the energy policy is contributing to new technologies or insights about policy designs, and to ensure that learning is capitalized and conveyed into new and stronger policy initiatives.

Lastly, concerning the methodologies for conducting evaluations, current evaluation practices show strengths in the predominant application of multiple methods and criteria in the same evaluation. We also notice some good practices in how counterfactuals are constructed, building on different methods and sources. In order to further advance good practices, an emphasis should be placed on extending the use of multiple criteria, and especially criteria promoting reflexivity, to further capture processes and effects of a policy instruments. The frequent application of multiple methods, furthermore, paves way for triangulation that can validate findings and thus provide added robustness to the results.

Altogether, we find the proposed theory-based evaluation framework useful for assessing and discussing both robustness and transformative efforts of current policy instruments and evaluation practices. In order to further promote an evaluation approach that can support transformative changes in society, we stress the need to further link evaluation theory and policy analysis with transition research to design evaluations that can provide a more systematic picture of the progress towards a transition to a more sustainable energy system.

Appendix 1

This appendix contains the theory-based framework used for the systematic review of evaluations. The framework is based on evaluation theory, policy analysis and transition research, and comes together under the three main categories of *methods*, *value judgements* and *use of evaluations*.

- A. Methods applied in evaluations
 - A.1. What methods are used?
 - A.2. How has the impact been assessed?
 - A.3. How is the counterfactual constructed?
 - A.4. Is the focus on just intended effect or are side-effects (intended and unintended) and rebound effects also considered?
 - A.5. If triangulation is used how has the synthesis been produced?
 - A.6. Is transition and the potential for transitions analysed? - Does the evaluation have a system perspective?; Does the evaluation have a multi-actor perspective?; Is the evaluation taking a visionary perspective (combined ex-post and ex-ante approach; part of a long-term evaluation approach)?; Is the evaluation capturing experimentation (radical policies and/or potential radical outcomes)?; Is the evaluation considering learning
- B. Value judgments in evaluations
 - B.1. How and by whom have the criteria been decided (The organization commissioning the

evaluation, the evaluator(s), by stakeholders, general evaluation policy)?

- B.2. Which value criteria were used to judge the intervention? (e.g. relevance, effectiveness, efficiency, flexibility, predictability, persistence, acceptability, transparency, equity).
 - B.3. Do the value criteria reflect the interests of different groups?
 - B.4. Do the criteria used promote reflexivity and challenge established goals, needs and methods?
 - B.5. Is reflexivity part of the value judgement the conclusions are based on?
- C. Use of evaluations
- C.1. Have key stakeholders been identified and involved in the evaluation process?
 - C.2. Have there been any specific efforts to engage different groups, including those that are not well organized?
 - C.3. What has been the time frame for the use of the results?
 - C.4. What particular activities have been undertaken to facilitate use?
 - C.5. Have there been efforts to promote use beyond "intended use by intended users" by making the process open and transparent or by making the evaluation results/report freely available and easy to obtain?

Appendix 2

This appendix contains two parts: (I) a complete reference list of evaluations included in the review, and (II) a list of evaluations that were not included in the review.

(I) Evaluations included in the review

The sample of reviewed policy evaluations amounts to a total of 33 evaluations. The complete list of reviewed evaluations is presented below, grouped according to the policy instrument types.

Legislative instruments

1. Helmersdotter Eriksson, A., Sandén, P., Jacobsson, R. (2015) Kartläggning och analys avseende aktörers syn på en lämplig skärpning av energikraven i nu varande byggregler [Mapping and analysis of actors' views on an appropriate tightening of the energy requirements of current building codes]. Sweco Strategy.
2. National Board of Housing, Building and Planning (2014) Skärpta energihushållningskrav – redovisning av regeringens uppdrag att se över och skärpa energireglerna i Boverkets byggregler, rapport 2014:19 [More stringent energy conservation requirements - the accounting of the government's instructions of reviewing and tightening energy regulations in building codes,

report 2014:19]. National Board of Housing, Building and Planning.

Financial instruments

3. Grontmij (2007) Utvärdering av stöd för genomförande av energikartläggningar inom OFFROT [Evaluation of the support for implementation of energy audits within OFFROT]. Swedish Energy Agency.
4. National Board of Housing, Building and Planning (2009a) Utformningen reducerade effekterna – Boverkets utvärdering av OFFrotstödet. Redovisning av stöd till investeringar i energieffektivisering och konvertering till förnybara energikällor i lokaler som används för offentlig verksamhet 2005:205, enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A [The design reduced the impact - the Board's evaluation of the OFFrot financial support. Reporting of support for investment in energy efficiency and conversion to renewable energy sources in premises used for public activities 2005:205, in accordance with appropriation directions for the financial year 2009 concerning the National Board of Housing, Building and Planning. M2008/4791/A]. National Board of Housing, Building and Planning.
5. National Board of Housing, Building and Planning (2009b) Utvärdering av stödet för installation av energieffektiva fönster eller biobränsleanordningar [Evaluation of the financial support for installation of energy efficient windows or biofuel devices]. National Board of Housing, Building and Planning.
6. Sweco (2014) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.6 – Energikartläggningsstöd [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.6 – Financial support for energy audits]. Sweco Energuide AB.
7. Swedish Energy Agency (2013) Energikartläggningscheckar – En samhällsekonomisk utvärdering [Financial support for energy audits - A socio-economic evaluation]. Swedish Energy Agency.

Informative instruments

8. Ekander, F., Olin, M., Ågren, R-M. (2015) Slutrapport: effekter av energi- och klimatrådgivningen

- 2013 till privatpersoner, företag och organisationer [Final report: the effects of the energy and climate advisory in 2013 to private individuals, companies and organizations]. Swedish Energy Agency.
9. Grontmij (2009) Utvärdering av. Energimyndighetens program för passivhus och lågenergihus 2005–2008 [Evaluation of the Swedish Energy Agency's program for passive houses and low-energy houses 2005–2008]. Swedish Energy Agency.
 10. Holm, D., Martinsson, F. (2012a) Utvärdering av. externa projekt – Bygg och Industri – För Energimyndigheten. 5 - Stoppsladd – Etapp 2 [Evaluation of external projects - Construction and Industry – For the Swedish Energy Agency. 5 - Stoppsladd - Stage 2]. IVL, Swedish Environmental Research Institute.
 11. Holm, D., Martinsson, F. (2012b) Utvärdering av. externa projekt – Bygg och Industri – För Energimyndigheten. 8 – BeAware [Evaluation of external projects - Construction and Industry – For the Swedish Energy Agency. 8 – BeAware]. IVL, Swedish Environmental Research Institute.
 12. Holm, D., Martinsson, F. (2012c) Utvärdering av. externa projekt – Bygg och Industri – För Energimyndigheten. 10 – Energieffektivisering och bevarande i vårt kulturarv [Evaluation of external projects - Construction and Industry – For the Swedish Energy Agency. 10 - Energy efficiency and conservation of our cultural heritage]. IVL, Swedish Environmental Research Institute.
 13. Holm, D., Martinsson, F. (2012d) Utvärdering av. externa projekt – Bygg och Industri – För Energimyndigheten. 12 – Radikal energieffektivisering och upprustning av. miljonprogrammet [Evaluation of external projects - Construction and Industry – For the Swedish Energy Agency. 12 - Radical energy efficiency and renovation of the Million Program]. IVL, Swedish Environmental Research Institute.
 14. Holm, D., Martinsson, F. (2012e) Utvärdering av. externa projekt – Bygg och Industri – För Energimyndigheten. 13 – Herrgårdsängen [Evaluation of external projects - Construction and Industry – For the Swedish Energy Agency. 13 – Herrgårdsängen]. IVL, Swedish Environmental Research Institute.
 15. National Board of Housing, Building and Planning (2009) Utvärdering av. systemet med energideklarationer. Uppdrag nr 12 Uppföljning av. energideklarationer enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A [Evaluation of the system of energy performance certificates. Assignment No. 12 Follow-up of energy performance certificates under the appropriation directions for the financial year 2009 concerning the National Board of Housing, Building and Planning. M2008/4791/A].
 16. Sweco (2014a) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av. 24 styrmedel – En rapport till Näringsdepartementet. 2.1 - Energieffektiva myndigheter [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.1 – Energy Efficient Authorities]. Sweco Energuide AB.
 17. Sweco (2014b) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av. 24 styrmedel – En rapport till Näringsdepartementet. 2.2 Energieffektiviseringsrådet [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.2 – The council for energy efficiency]. Sweco Energuide AB.
 18. Sweco (2014c) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av. 24 styrmedel – En rapport till Näringsdepartementet. 2.3 - Energieffektiviseringsstöd till kommuner och landsting [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.3 – Support for energy efficiency for municipalities and county councils]. Sweco Energuide AB.
 19. Sweco (2014d) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av. 24 styrmedel – En rapport till Näringsdepartementet. 2.8 – Energi- och Klimatrådgivning [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.8 – Energy and climate advisory program]. Sweco Energuide AB.
 20. Sweco (2014e) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av. 24 styrmedel – En rapport till Näringsdepartementet. 2.17 Energitjänster [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.17 - Energy services]. Sweco Energuide AB.
 21. Sweco (2014f) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av.

- 24 styrmedel – En rapport till Näringsdepartementet. 2.21 – Energiaktiv.se [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.21 – Energiaktiv.se]. Sweco Energuide AB.
22. Sweco (2014 g) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.22 – Byggnaders energiprestanda – energideklarationer [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.22 - Energy performance of buildings - energy performance certificates]. Sweco Energuide AB.
23. Swedish Energy Agency (2011) Långtidsuppföljning av energi- och klimatrådgivningen 2008 och 2009, ER 2011:20 [Long-term follow-up of the energy and climate advisory in 2008 and 2009, ER 2011:20]. Swedish Energy Agency.
24. Swedish Energy Agency (2013a) Effekter av energi- och klimatrådgivningen 2011 [Effects of the energy and climate advisory in 2011]. Swedish Energy Agency.
25. Swedish Energy Agency (2013b) Effekter av energi- och klimatrådgivningen 2012 till privatpersoner, företag och organisationer [Effects of the energy and climate advisory in 2012 to private individuals, companies and organizations]. Swedish Energy Agency.
- Other instruments**
26. Jansson, T., Terrell, M. (2012) Halvtidsutvärdering av HyLok, statliga Hyresgäster för energieffektivisering av Lokaler [Mid-term evaluation of HyLok, state tenants for energy efficiency in premises]. Faugert & Co Utvärdering AB.
27. Jansson, T., Terrell, M. (2013) Halvtidsutvärdering av program för byggnader med mycket låg energianvändning – LÅGAN [Mid-term evaluation of the program for buildings with very low energy use – LÅGAN]. Faugert & Co Utvärdering AB.
28. Jansson, T., Enberg, J., Stålfors, S., Terrell, M., Karlström, S. (2014) Utvärdering av HyLok, statliga Hyresgäster för energieffektivisering av Lokaler [Evaluation of HyLok, state tenants for energy efficiency in premises]. Faugert & Co Utvärdering AB.
29. Jansson, T., Grudin, M., Henningsson, K. (2016) Slututvärdering av Programmet för byggnader med mycket låg energianvändning (Lågan) [Final evaluation of the program for buildings with very low energy use (Lågan)]. Faugert & Co Utvärdering AB.
30. Stern, P. Enberg, J., Fridholm, T., Årenman, E. (2013) Evaluation of BeLivs, the Swedish Energy Agency's network program for energy efficiency in food premises (Utvärdering av. BeLivs, Energimyndighetens Beställargrupp Livsmedelslokaler). Faugert & Co Utvärdering AB.
31. Sweco (2014a) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.18 – Beställargrupper formerade för att driva utvecklingen [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.18 – Cooperative network programs for driving development]. Sweco Energuide AB.
32. Sweco (2014b) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.19 – Programmet för mycket låg energianvändning i byggnader – LÅGAN [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.19 – The program for buildings with very low energy use – LÅGAN]. Sweco Energuide AB.
33. Sweco (2014c) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.20 – Projektstöd för marknadsintroduktion och teknikupphandling [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.20 - Project support for market introduction and technology procurement]. Sweco Energuide AB.
- (II) Evaluations excluded from the review
- A number of evaluations of policy instruments that were only partially or indirectly related to energy efficiency in buildings were excluded from the review. These are presented according to the general type of evaluand: multi-sectoral policy instruments and policy mixes.
- Evaluations of multi-sectoral policy instruments
- National Board of Housing, Building and Planning (2014) Utvärdering av stödet till hållbara städer, rapport 2014:26 [Evaluation of the support to the Sustainable Cities program, report 2014:26].

- Samakovlis, E., Vredin Johansson M. (2007) En utvärdering av kostnadseffektiviteten i klimatinvesteringsprogrammen [An evaluation of the cost-effectiveness of climate investment programs]. National Institute of Economic Research.
- Swedish Environmental Protection Agency (2009) Effekter av investeringsprogrammen LIP och Klimp, rapport 5991 [Effects from Local Investmentprograms LIP and KLIMP, report 5991].
- Swedish Environmental Protection Agency (2013) Klimatinvesteringsprogrammen Klimp 2003–2012, Slutrapport, rapport 6517 [Climate Investmentprograms Klimp 2003–2012, Final report, report 6517].
- Vredin Johansson, M. (2010) En utvärdering av det ekonomiska stödet till åtgärder för att främja hållbara städer [An evaluation of the financial support for measures to promote Sustainable Cities]. Specialstudier Nr 23, National Institute of Economic Research (NIER).

Assessments of policy mixes using energy system models, economic modelling or bottom-up/top-down effect calculations

- Hammes, K., Högberg, M., Forsberg, A., Karlsson, T., Lopes, C., Waluszewski, D., Widerström, G. (2007) Effektivare energianvändning [More efficient energy use]. ER 2007:21, Swedish Energy Agency.
- National Institute of Economic Research (2009) Samhällsekonomiska effekter av energi- och koldioxidskatteförändringar som beslutades av riksdagen 2009 [Socio-economic effects of energy and carbon taxes decided by the Riksdag 2009]. Fördjupnings-PM Nr 10.2012.
- National Institute of Economic Research (2013) Miljö, ekonomi och politik 2013 [Environment, economy and politics 2013].
- Profu (2008) Konsekvenser av den svenska energi- och klimatpolitiken sedan 1990 – Modellberäkningar [Consequences of Swedish energy and climate policy since 1990 - Model calculations]. Profu i Göteborg AB.
- Profu (2011) Utvecklad modellutvärdering av CO₂-utsläppen från bostäder & service – Slutrapport [Developed model evaluation of CO₂ emissions from housing & service - Final report]. Profu i Göteborg AB.
- Profu (2013) Beräkningar med MARKAL-NORDIC inför Sveriges klimatrapportering (NC6) [Calculations with MARKAL-NORDIC for Sweden's climate reporting (NC6)]. Profu i Göteborg AB.

Abbreviations

EU: European Union; IPCC: the Intergovernmental Panel on Climate Change

Acknowledgements

The authors are thankful to fellow researchers in the research project Transition Governance: Professor Mats Benner and Professor Maria Johansson, for their comments and input on this work; as well as to members of the reference group: Peter Stern, Lisa Eriksson, Hanne Foss-Hansen and Tobias Persson, for their valued feedback. Additional valued feedback has been received from the highly appreciated audiences during presentations of this work at scientific conferences, including the European Council for an Energy Efficient Economy (ECEEE) Summer Study (2017), the International Sustainability Transitions Conference (2017) and the European Environmental Evaluators Network Forum (2017). The authors would also like to acknowledge the very valuable comments received from the peer reviewers, which strengthened the manuscript into its final version.

Funding

This work was supported by the Swedish Energy Agency under Grant number 39938-1, the research project Transition Governance.

Availability of data and materials

All 33 Swedish policy evaluations that have been reviewed within the scope of this study are included in Appendix 2 of this published article. At the time of inquiry, the evaluations were available upon request to relevant authorities, or on relevant authorities' webpages.

Authors' contributions

LN and PM developed research design and the theoretical framework used to guide the review. SS performed the review and assessment of the evaluations. SS and LN were major contributors in writing the manuscript. All authors have regularly read, refined and lastly approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹International Institute for Industrial Environmental Economics (IIIEE), Lund University, P.O. Box 196, SE-221 00 Lund, Sweden. ²Finnish Environment Institute SYKE, Helsinki, Finland.

Received: 8 January 2019 Accepted: 12 May 2019

Published online: 29 May 2019

References

1. Lucon O, Urge-Vorsatz D, Zain Ahmed A, Akbari H, Bertoldi P, Cabeza LF et al (2014) Buildings. In: Edenhofer O, Pichs-Madruga R, Sokona Y, Farahani E, Kadner S, Seyboth K, Adler A, Baum I, Brunner S, Eickemeier P, Kriemann B, Savolainen J, Schlömer C, von Stechow C, Zwickel T, Minx JC (eds) Climate Change 2014: mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York
2. International Energy Agency [IEA] (2017) Energy technology perspectives 2017—catalyzing energy technology transformations
3. OECD/IEA & IRENA (2017) Perspectives for the energy transition—investment needs for a low-carbon energy system
4. International Energy Agency [IEA] (2018) Energy efficiency 2018 - Analysis and outlooks to 2040. Market Report Series
5. International Energy Agency [IEA] (2019) Building energy efficiency policies (BEEP) database. <http://www.iea.org/beep/>. Accessed 7 Jan 2019.
6. MURE (2018) Policies by Topics. <http://www.measures-odyssey-mure.eu/topics-energy-efficiency-policy.asp>. Accessed 30 Nov 2018.

7. World Business Council for Sustainable Development [WBCSD] (2016) A handbook on creating dynamic local markets for energy efficiency in buildings
8. Leeuw FL, Vaessen J (2010) Introduction: Mind the Gap. In: Vaessen J, Leeuw FL (eds) *Mind the gap : perspectives on policy evaluation and the social sciences*. Transaction Publishers, New Brunswick
9. Mela H, Hildén M (2012) Evaluation of climate policies and measures in EU member states—examples and experiences from four sectors. *The Finnish Environment*, 19/2012.
10. Weiss CH (1998) *Evaluation: methods for studying programs and policies*, 2nd edn. Prentice Hall, Upper Saddle River
11. Haug C, Rayner T, Jordan A, Hildingsson R, Stripple J, Monni S et al (2010) Navigating the dilemmas of climate policy in Europe: evidence from policy evaluation studies. *Clim Chang* 101(3–4):427–445. <https://doi.org/10.1007/s10584-009-9682-3>
12. Vedung E (1997) *Public policy and program evaluation*. Transaction Publishers, New Brunswick
13. European Commission (2017) *Better Regulation Guidelines (SWD (2017) 350 - Staff Working Document)*. Brussels: European Commission.
14. Schoenefeld JJ, Jordan AJ (2019) Environmental policy evaluation in the EU: between learning, accountability, and political opportunities? *Environ Politics* 28(2):365–384. <https://doi.org/10.1080/09644016.2019.1549782>
15. Fischer F (2006) *Evaluating public policy* (paperback ed). Wadsworth Publ, Boston
16. Vedung E (2006) Evaluation Research. In: Peters BG, Pierre J (eds) *Handbook of public policy*. SAGE Publications, Thousand Oaks, pp 397–416
17. Alkin MC, Christie CA (2013) An evaluation theory tree. In: Alkin MC (ed) *Evaluation roots: a wider perspective of theorists' views and influences*, 2nd edn. SAGE Publications, Thousand Oaks
18. Patton MQ (2008) *Utilization-focused evaluation*, 4th edn. Sage Publications, Thousand Oaks
19. Preskill H (2004) The transformational power of evaluation—passion, purpose, and practice. In: Alkin MC (ed) *Evaluation roots - tracing theorists' views and influences*. SAGE Publications, Inc, Thousand Oaks
20. Weiss CH (1998) Have we learned anything new about the use of evaluation? *Am J Eval* 19(1):21–33
21. Knoepfel P, Larrue C, Hill M, Varone F (2011) *Public policy analysis*. Bristol: Policy Press
22. Kern F, Kivimaa P, Martiskainen M (2017) Policy packaging or policy patching? The development of complex energy efficiency policy mixes. *Energy Res Soc Sci* 23:11–25
23. Rogge K, Reichardt K (2016) Policy mixes for sustainability transitions: an extended concept and framework for analysis. *Res Policy* 45(8):1620–1635
24. Geels FW (2002) Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Res Policy* 31(8–9): 1257–1274. [https://doi.org/10.1016/S0048-7333\(02\)00062-8](https://doi.org/10.1016/S0048-7333(02)00062-8)
25. Kemp R (1994) Technology and the transition to environmental sustainability. *Futures* 26(10):1023–1046. [https://doi.org/10.1016/0016-3287\(94\)90071-X](https://doi.org/10.1016/0016-3287(94)90071-X)
26. Rip A, Kemp R (1998) Technological change. In: Rayner S, Malone EL (eds) *Human choice and climate change*, vol 2. Battelle Press, Columbus, pp 327–399
27. Rotmans J, Kemp R, Van Asselt M (2001) More evolution than revolution: transition management in public policy. *Foresight* 3(1):15–31
28. Furubo J-E, Sandahl R (2002) Coordinated pluralism—The Swedish case. In: Furubo J-E, Rist RC, Sandahl R (eds) *International Atlas of Evaluation*. Transaction Publishers, New Brunswick, pp 115–128
29. Geels FW, Sovacool BK, Schwanen T, Sorrell S (2017) Sociotechnical transitions for deep decarbonization. *Science* 357(6357):1242–1244. <https://doi.org/10.1126/science.aao3760>
30. Kemp R, Schot J, Hoogma R (1998) Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management. *Tech Anal Strat Manag* 10(2):175–198. <https://doi.org/10.1080/09537329808524310>
31. Loorbach D (2010) Transition management for sustainable development: a prescriptive, complexity-based governance framework. *Governance* 23(1): 161–183. <https://doi.org/10.1111/j.1468-0491.2009.01471.x>
32. Chelmsky E, Shadish WR (eds) (1997) *Evaluation for the 21st century: a handbook*. Sage Publications, Thousand Oaks
33. Scriven M (1991) *Evaluation thesaurus*, 4th edn. Sage Publications, Newbury Park
34. Denzin NK (2009) *The research act: a theoretical introduction to sociological methods*. AldineTransaction, New Brunswick
35. Patton MQ (2002) *Qualitative research and evaluation methods*, 3rd edn. Sage Publications, Thousand Oaks
36. Ferraro PJ (2009) Counterfactual thinking and impact evaluation in environmental policy. In: Birnbaum M, Mickwitz P (eds) *Environmental program and policy evaluation: addressing methodological challenges*, vol 122, pp 75–84
37. Raven R, Schot J, Berkhout F (2012) Space and scale in socio-technical transitions. *Environ Innov Soc Trans* 463–78. <https://doi.org/10.1016/j.jeist.2012.08.001>
38. Smith A, Raven R (2012) What is protective space? Reconsidering niches in transitions to sustainability. *Res Policy* 41(6):1025–1036. <https://doi.org/10.1016/j.respol.2011.12.012>
39. Smith A, Stirling A, Berkhout F (2005) The governance of sustainable socio-technical transitions. *Res Policy* 34(10):1491–1510. <https://doi.org/10.1016/j.respol.2005.07.005>
40. Loorbach D (2007) Transition management: new mode of governance for sustainable development = Transitie management ; nieuwe vorm van governance voor duurzame ontwikkeling. Internat. Books, Utrecht
41. Schot J, Geels FW (2008) Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Tech Anal Strat Manag* 20(5):537–554. <https://doi.org/10.1080/09537320802292651>
42. Bennett CJ, Howlett M (1992) The lessons of learning: reconciling theories of policy learning and policy change. *Policy Sci* 25(3):275–294. <https://doi.org/10.1007/BF00138786>
43. Loorbach D, Rotmans J (2006) Managing transitions for sustainable development. In: Olsthoorn X, Wiecek AJ (eds) *Understanding industrial transformation*, vol 44. Kluwer Academic Publishers, Dordrecht, pp 187–206. https://doi.org/10.1007/1-4020-4418-6_10
44. van de Kerkhof M, Wiecek AJ (2005) Learning and stakeholder participation in transition processes towards sustainability: methodological considerations. *Technol Forecast Soc Chang* 72(6):733–747. <https://doi.org/10.1016/j.techfore.2004.10.002>
45. Kivimaa P, Kern F (2016) Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions. *Res Policy* 45(1):205–217. <https://doi.org/10.1016/j.respol.2015.09.008>
46. Shadish WR, Cook TD, Leviton LC (eds) (1991) *Foundations of program evaluation: theories of practice* (reprinted). Sage Publications, Newbury Park
47. Mickwitz P (2003) A framework for evaluating environmental policy instruments context and key concepts. *Evaluation* 9(4):415–436. <https://doi.org/10.1177/1356389003094004>
48. House E, Howe KR (1999) *Values in evaluation and social research*. Thousand Oaks, CA: SAGE Publications
49. Argyris C (1976) Single-loop and double-loop models in research on decision making. *Adm Sci Q* 21(3):363. <https://doi.org/10.2307/2391848>
50. Chelmsky E (1985) Comparing and contrasting auditing and evaluation: some notes on their relationship. *Eval Rev* 9(4):483–503. <https://doi.org/10.1177/0193841X8500900406>
51. Cooper ACG (2018) Evaluating energy efficiency policy: understanding the 'energy policy epistemology' may explain the lack of demand for randomised controlled trials. *Energy Effic* 11(4):997–1008. <https://doi.org/10.1007/s12053-018-9618-8>
52. Hansen HF, Rieper O (2009) The evidence movement: the development and consequences of methodologies in review practices. *Evaluation* 15(2):141–163. <https://doi.org/10.1177/1356389008101968>
53. Vedung E (2010) Four waves of evaluation diffusion. *Evaluation* 16(3):263–277. <https://doi.org/10.1177/1356389010372452>
54. National Board of Housing, Building and Planning (2009) *Urfomningarna reducerade effekterna – Boverkets utvärdering av Offrotstødet. Redovisning av stöd till investeringar i energieffektivisering och konvertering till förnybara energikällor i lokaler som används för offentlig verksamhet 2005:205, enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A* [The design reduced the impact - the Board's evaluation of the Offrot financial support. Reporting of support for investment in energy efficiency and conversion to renewable energy sources in premises used for public activities 2005: 205, in accordance with appropriation directions for the financial year 2009 concerning the National Board of Housing, Building and Planning. M2008/4791/A]. National Board of Housing, Building and Planning
55. Swedish Energy Agency (2013) *Energikartläggningscheckar – En samhällsekonomisk utvärdering* [Financial support for energy audits - A socio-economic evaluation]. Swedish Energy Agency

56. National Board of Housing, Building and Planning (2014) Skärpta energihushållningskrav – redovisning av regeringens uppdrag att se över och skärpa energireglerna i Boverkets byggregler, rapport 2014:19 [More stringent energy conservation requirements - the accounting of the government's instructions of reviewing and tightening energy regulations in building codes, report 2014:19]. National Board of Housing, Building and Planning
57. National Board of Housing, Building and Planning (2009) Utvärdering av stödet för installation av energieffektiva fönster eller biobränsleanordningar [Evaluation of the financial support for installation of energy efficient windows or biofuel devices]. National Board of Housing, Building and Planning
58. Helmersdotter Eriksson, A., Sandén, P., Jacobsson, R. (2015) Kartläggning och analys avseende aktörers syn på en lämplig skärpning av energikraven i nu varande byggregler [Mapping and analysis of actors' views on an appropriate tightening of the energy requirements of current building codes]. Sweco strategy
59. National Board of Housing, Building and Planning (2009) Utvärdering av systemet med energideklarationer. Uppdrag nr 12 Uppföljning av energideklarationer enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/a [Evaluation of the system of energy performance certificates. Assignment no. 12 follow-up of energy performance certificates under the appropriation directions for the financial year 2009 concerning the National Board of housing, building and planning. M2008/4791/A]
60. Grontmij (2007) Utvärdering av stöd för genomförande av energikartläggningar inom OFFROT [Evaluation of the support for implementation of energy audits within OFFROT]. Swedish Energy Agency
61. Sweco (2014) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.18 – Beställargrupper formerade för att driva utvecklingen [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments—a report to the Ministry of Enterprise and Innovation. 2.18—cooperative network programs for driving development]. Sweco Energuide AB
62. Sweco (2014) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.20 – Projektstöd för marknadsintroduktion och teknikupphandling [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments—a report to the Ministry of Enterprise and Innovation. 2.20—project support for market introduction and technology procurement]. Sweco Energuide AB
63. Grontmij (2009) Utvärdering av Energimyndighetens program för passivhus och lågenergihus 2005–2008 [Evaluation of the Swedish Energy Agency's program for passive houses and low-energy houses 2005–2008]. Swedish Energy Agency
64. Sweco (2014) Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.8 – Energi- och Klimatrådgivning [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments—a report to the Ministry of Enterprise and Innovation. 2.8—energy and climate advisory program]. Sweco Energuide AB
65. European Environment Agency (2016) Environment and climate policy evaluation. EEA Report No 18/2016. European Environment Agency, Copenhagen

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



Paper III





Bolstering a transition for a more sustainable energy system: A transformative approach to evaluations of energy efficiency in buildings

Lena Neij^{a,*}, Sofie Sandin^a, Mats Benner^b, Maria Johansson^c, Per Mickwitz^a

^a International Institute for Industrial Environmental Economics (IIIEE), Lund University, P.O. Box 196, SE-221 00 Lund, Sweden

^b Department of Business Administration, Lund University, P.O. Box 196, SE-221 00 Lund, Sweden

^c Environmental Psychology, Department of Architecture and the Built Environment, Lund University, P.O. Box 118, SE-221 00 Lund, Sweden

ARTICLE INFO

Keywords:
Evaluation
Transition
Energy efficiency
Buildings

ABSTRACT

A transition to a more sustainable energy system is called for, and to realise such a transition governments around the world have introduced various research programmes and policy instruments. In this paper, we elaborate on how the transition of the energy system could be reinforced through evaluation of research programmes and policy instruments, focusing on energy efficiency in buildings. We argue that a transition will require transformative evaluations that provide necessary learning from actions taken, and we propose an evaluation approach that complements, and goes beyond, the evaluation practice of today; an approach framed in line with the transition literature as well with evaluation research. We present a conceptual review of research literature, and apply the conceptual findings of this review to assess current evaluations of research programmes and policy instruments, focusing on energy efficiency in buildings in Sweden. The results show that the weakness of current evaluation practice is the fragmented approach to, and limited scope of, evaluations, which does not support transformative change. Despite this, there is potential for modification of the current evaluation approach to better capture transformative change by, for example, strengthening coordination between evaluations, providing better vision documents for the envisioned transition, and by providing better evaluations of innovative efforts by evaluating the outcome and learning in relation to upscaling processes. Moreover, the use of key evaluation concepts for supporting knowledge-based learning, such as the use of multiple methods, multiple criteria, and counterfactual analysis, can be further extended and improved to cater for transformative knowledge.

1. Introduction

Improving the energy efficiency of buildings can make a significant contribution to a more sustainable energy economy. This will not only contribute to the more efficient use of resources, but will also be important in meeting goals related to climate change, air pollution, energy security, energy poverty, health and environmental degradation [1–3]. The IPCC [2] scenarios show that improving energy efficiency is by far the most important measure in stabilizing atmospheric concentrations of greenhouse gases. The IPCC also projects that over the next two decades annual investments in energy efficiency will have to increase by about 336 billion USD to stabilize atmospheric concentrations in the range of 430 to 530 ppm CO₂ by 2100. The European Commission [4] has estimated that the annual investment required in residential buildings within the EU between 2021 and 2030 to achieve the EU

climate and energy targets by 2030, is approximately 250 billion euros. Buildings account for approximately 32% of the global final energy use and 19% of energy-related greenhouse gas emissions, and it has been estimated that the potential for improvements in energy efficiency is 50–75% in existing buildings and 50–90%, in new buildings, based on changes in design practices, technology and behaviour [5].

Improving the energy efficiency in buildings is, however, extremely complex, comprising a variety of technologies, actors and institutions. The sector is highly fragmented, and the number of actors embedded in the building value chain is high, including e.g. designers and architects, developers and capital providers, contractors, engineers, building owners and tenants [6]. Moreover, buildings can be of various types and ages, with tenants who have diverse values, lifestyles and ways of using buildings. This complexity leads to mixed motivations for energy efficiency measures, resulting in barriers to improve energy efficiency. As a

* Corresponding author.

E-mail addresses: lena.neij@iiiee.lu.se (L. Neij), sofie.sandin@iiiee.lu.se (S. Sandin), mats.benner@fek.lu.se (M. Benner), per.mickwitz@iiiee.lu.se (P. Mickwitz).

result of such barriers, traditional energy efficiency measures most often only provide limited and incremental energy efficiency adjustments (including the implementation of technological measures such as insulation, energy-efficient windows, and the installation of heat exchangers, energy-efficient lamps and appliances). More recent governmental and private initiatives have, however, been designed with the aim of bringing about a more profound shift in the building sector, and more radical changes have been envisaged in terms of zero-energy houses, plus-energy houses and smart buildings, as well as stricter approaches to deep renovations. More radical changes can be accomplished, for example through new types of building materials, shared and flexible building space, and an extended definition of “buildings” that integrates services such as local energy supply and storage.

To support incremental, and also more radical, shifts towards greater energy efficiency in buildings, governments have implemented and financed research programmes and policy instruments. Research has been carried out in various fields, form new technology to human behaviour. Policy programmes include informative instruments, regulatory frameworks and fiscal and economic frameworks [5,7] designed to overcome market and behavioural failures and accelerate change [8–10]. Studies have shown that the most effective policy instruments for energy efficiency in buildings in the past have been appliance standards, building codes, tax exemptions and voluntary energy labelling [11,12]. Over time, more progressive and alternative policy instruments have been advocated [13–15], and new types of incentives have been introduced, such as collaborative platforms, innovation procurement programmes and tradable permits [5,16]. Moreover, a variety of voluntary measures and standards, such as the passive-house concept, LEEDS, BREEAM, etc., have been introduced mainly by the private sector to profile energy efficiency in buildings [7]. Many cities around the world have also adopted progressive building requirements, that are more ambitious than the national building codes [5], and various forms of dialogues have been established in an effort to facilitate the transition to more energy efficient buildings [17,18].

Despite numerous research and policy programmes, and a significant potential for improvements, progress in energy efficiency in buildings has been moderate [3] and could not be described in terms of a transition, i.e. a transformative change in society.¹ A transition will require that technical, institutional and behavioural changes go hand in hand, and change must permeate the entire value chain of buildings; from construction, renovation and maintenance to their use. Such a transformative change has not been realised thus far.

Our ambition in this study is to elaborate on how a transition to more energy efficient buildings could be reinforced. We argue that governance will play a key role in shaping the transition towards improved energy efficiency in buildings, and that the realisation of this transition will require not only innovative measures, research and policy programmes, but also evaluation frameworks that ensure learning from the actions taken [24,25] and guidance for future initiatives. Thus, there is a need for a continuous dialogue on ways of evaluating [26–28] and supporting this kind of transition through knowledge provided by the evaluation of both research programmes and policy instruments. Similar to regarding the evaluand as a case of a bigger context, where an evaluation may provide learning that affects other implementations, there is a need to approach evaluands from a broader perspective [29]. However, several review studies have shown that current evaluation practices for energy efficiency policy often have had a narrow focus on the impact of measures and on compliance with targets [24,30,31], although the requirements for evaluations have become more stringent [25]. Whereas the evaluation literature provides a wide variety of

theories, methodologies and tools for the evaluation of research and policy programmes (see Section 3 below), such approaches seem to be less common in the evaluations commissioned by governments. Nonetheless, the application of rigorous evaluations of individual research and policy programmes may vary from country to country, and the results are often only available as grey literature, in native languages, not readily accessible for international reviews.² *It is therefore unclear to what extent current evaluation practice is capable of embracing and supporting the transition towards energy efficiency in buildings.*

In all, we see an urgent need for an evaluation approach that complements, and goes beyond, the evaluation practice focusing on impact and compliance with targets; that promotes transformative evaluations that ensures learning and provides knowledge supporting a transition. Such transformative evaluations will require not only rigorous evaluation of individual research and policy programmes, but also an overarching evaluation strategy that ensures that evaluations of individual programmes altogether capture and assess contributions to the envisaged transition. The transformative evaluations should support long-term, radical and systemic change, i.e. the integration of technical, institutional and behavioural changes, and be based on the synthesis of the outcomes of evaluations at the programme level, as well as their combined transformative efforts and effects at the societal level.³

In this paper, we propose transformative evaluations that can be used to facilitate the transition towards greater energy efficiency in buildings. We have identified two important lines in such a transformative evaluation approach: 1) it should be framed in line with the transition literature, and 2) it should build on a combination of relevant evaluation-oriented disciplines (below referred to as a knowledge-based evaluation approach). The evaluation of research programmes and policy instruments to improve energy efficiency in buildings can be based on many areas of knowledge; in this study we have identified and synthesized key aspects from evaluation theory, the sociology of science, policy analysis and behavioural science perspectives on energy use. Nevertheless, there are other disciplines that may be relevant as well, and a broader evaluation approach could also include insights from e.g. organisational theory, institutional theory, ethics, politics, planning theory and literature on energy poverty and social justice. In Section 3, we present a conceptual review of the literature relevant to a transformative evaluation approach focusing on energy efficiency in buildings.

We also analyse the extent to which transformative evaluations can be applied on current evaluation practice in Sweden – i.e. the extent to which current evaluation practice includes transformative concepts that resonate with the transition literature, and the extent to which evaluations rely on knowledge-based evaluation approaches combining relevant scientific disciplines. Thus, in Section 4, we apply the key concepts found in the literature review presented in Section 3, to perform a review of evaluations conducted in Sweden focusing on energy efficiency in buildings (residential, public and commercial buildings, both existing and new). This second review will include evaluations commissioned by governmental authorities and funding agencies. However, as behavioural aspects were treated in a rudimentary way or were completely

¹ By transitions we here refer to large-scale, non-linear, radical and structural changes in a complex system [19]. The research literature on transitions in general is rich (see for example [20,21], and includes literature on energy efficiency in buildings [22,23].

² Examples of reports on energy efficiency in buildings (easily accessible in English) providing broad evaluation approaches of research and policy programmes include e.g. [32,33].

³ The suggested transformative evaluation approach may be associated to intervention theories; however, intervention theories are designed to analyse how interventions are intended to operate in relation to a given outcome, covering the links and causal mechanisms needed for the instrument to perform and function properly [34,35]. Whereas intervention theories are oriented towards the evaluation of policy programmes and a given outcome, the transformative evaluations suggested in this paper implies a much broader focus on learning based on synthesized outcomes of evaluations at the programme level, as well as their combined transformative efforts and effects at the societal level.

overlooked in the commissioned evaluations, academic research papers have been reviewed broadly in the case of energy use behaviour. The Swedish (policy) evaluation practice is well-established and covers multiple sectors, of which energy efficiency in buildings is one. As a country regarded as one of the pioneers of evaluation, with an extensive array of evaluations available [26,36], Sweden was deemed an appropriate case for reviewing evaluation practices. The reviews of the research and policy programmes, and the review on behavioural aspects are presented in more detail in Appendix A.

Based on our findings from the reviews of research literature and evaluation reports, we discuss the foundation for transformative evaluations for energy efficiency in buildings in Sweden in Section 5. We present an assessment of the extent to which transformative evaluations are applied in current evaluation practice, and how the current evaluation practice can be improved towards a more transformative evaluation approach.

2. Methodology

The study presented in this paper is based on two review studies. In Section 3, we present a conceptual review (Literature review 1) of literature relevant for forming transformative evaluations focusing on energy efficiency in buildings, see Fig. 1. This conceptual review is then used in a second review (Literature review 2), to analyse the extent to which current evaluation practice, focusing on research and policy programmes for energy efficiency in buildings, includes transformative concepts consistent with the findings of the first review, see Fig. 1. The sample of reviewed evaluations (in Literature review 2) amounts to a total of 79 evaluations, of which 20 are research evaluations, 33 are policy instrument evaluations, and 26 are evaluations targeting behavioural aspects. The sample includes evaluations of Swedish research and policy instruments performed between the years 2005–2017, and research evaluations of behavioural aspects performed between 1997 and 2018; for more details see Appendix A.

3. Conceptualisation of evaluations for transformative change

In this section we present a conceptual review of the literature relevant for transformative evaluations (Literature review 1, see Fig. 1). As outlined previously, we identify two main lines when structuring transformative evaluations: first, to ensure that they are framed according to the transition literature, and secondly, that they are *knowledge-based* and build on key concepts from relevant evaluation-oriented research.

The point of departure is transition research, and in Section 3.1 we present the key concepts identified in this literature that can provide guidance for the evaluation of transformative change. In general, the transition literature has focused on conditions for transformative change in society and the capacity of the state to engage, while less attention has been paid to evaluations commissioned by governments, and to key concepts emphasized in the evaluation literature, such as counterfactual assessments for determining effects of research and policy interventions in society. Transformative assessment approaches have mainly provided insights into opportunities for, and barriers to, systemic change, and the literature emphasizes assessments in terms of feedback to strategic development practice [37,38] or in terms of process evaluation [39], rather than focusing on traditional objectives of evaluations such as relevance, impact and efficiency.

In order to strengthen the focus on learning, we complemented the review of transition literature with a review of relevant evaluation literature (to support a knowledge-based evaluation approach, or knowledge-based learning). In Section 3.2 we synthesize key concepts from evaluation theory and in Section 3.3 we synthesize key concepts from the sociology of science, policy analysis and energy use behaviour, which are research areas we identify as important when evaluating research and policy instruments related to energy efficiency in buildings. In Section 3.4 we summarise the key concepts identified in the literature review of different disciplines, i.e. concepts of relevance and importance when designing a systemic framework for the evaluation of interventions for transformative change. These key concepts are then used in Section 4 to assess evaluations of research and policy instruments aimed at improving energy efficiency in buildings in Sweden.

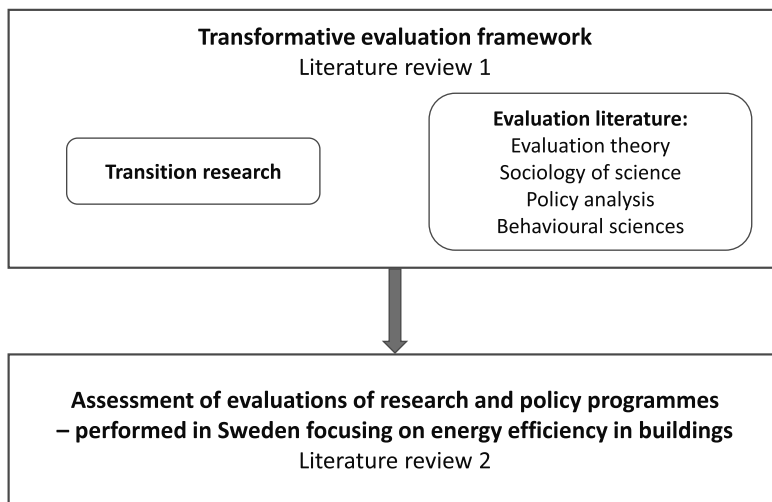


Fig. 1. Schematic overview of the two reviews presented in this paper. The first review concerns selected relevant literature for forming transformative evaluation (see Section 3), the second review covers 79 existing Swedish evaluations (see Section 4).

3.1. Guidance towards transitions

Transition research can provide valuable inspiration regarding the design of how to form a transformative evaluation approach, and three notable concepts are prevalent throughout the wide range of transition literature: visioning, experimentation and learning (see e.g. [40,41]). These focus on processes that are vital for the realisation of transitions, and may very well be used in a transformative evaluation approach. The transition literature also encompasses the evaluation of path dependencies and lock-in effects, and how to disrupt and destabilize them [42].

Visioning, or visions, which can be translated into roadmaps, missions and goals, inspire and guide efforts made towards transformative change [43–45]. Importantly, visioning underlines the long-term perspectives of any transition process, which should also be considered in evaluations; notably so for evaluations related to the building sector bearing in mind the long lifespan of buildings. A vision, in combination with roadmaps, missions and goals, will provide a useful map or plan for possible interventions, or combinations of interventions, that can be used to guide evaluations. The need for reflexivity, i.e. the need to reflect on and challenge the goals set, the interventions and efforts used to achieve them, and the evaluation approaches used to assess them, are also underlined in transition research in relation to visioning.

Experimentation is highlighted in radical change through risk-taking and acknowledging the challenges in the transformation of conventional practices and informal power structures [46–48]. Experimentation can be supported, for example, through research, pilot projects, demonstration projects, living labs and policy labs. The acknowledgement of innovation and experimentation in transition processes further highlights the need to evaluate outcome, in addition to impact, as well as parameters such as relevance and legitimacy. The evaluation of outcome is essential for the analysis and assessment of transformative change, and refers to the responses to research and policy interventions by actors, organisations and institutions in the socio-technical system. These responses may affect changes in technology, e.g. technological improvements and cost reductions, as well as changes in knowledge, behaviour and awareness [49].

Learning emphasizes the need for evaluations to improve the practice of research and policy instruments. In this paper we argue for a rigorous evaluation approach that goes beyond transition research [37,38]. Evaluation theory could be the starting point, but depending on the transition process in focus, additional disciplinary approaches could be applied (see text below).

Transition research also highlights the importance of a systems perspective in the analysis [20,40,50,51], which is also relevant for a transformative evaluation approach. This includes defining the system, analysing combinations of actions at different levels (global, national, regional, local, niche/hubs), and analysing side-effects, time aspects of interventions, uncertainties and complexities. A systems approach, moreover, calls for a thorough multi-actor approach and analysis that can support assessments of the involvement and interests of different stakeholder groups in the transition process.

3.2. Knowledge-based learning – A theoretical starting point

The evaluation of any intervention requires a rigorous foundation in evaluation theory. Traditional evaluation theory originated from experimental and quasi-experimental approaches [35,52,53], but has become a wide field involving various approaches such as rationalistic, constructivist, and realist evaluation. These fields promote different methods and areas of focus, for example the need for complete stakeholder participation and interactive processes as suggested by constructivist approaches [54,55], or the need to consider the context and social interventions in evaluations as promoted by realist evaluators [56]. More recently, efforts have been made to integrate various methods and approaches in evaluation theory. One such approach was

proposed by Alkin [52], who highlights three overarching dimensions of evaluations: (I) the methods applied in the evaluation, (II) the valuing or the making of judgment, and (III) the use of evaluations.

The methods applied in an evaluation may be used separately, but many scholars stress the importance of triangulation and the use of interdisciplinary approaches, such as mixed methods and combining theories, empirics and complementary datasets [57]. Methods provide tools to deliver credible results, accountability and learning; at the same time, there is a need to demonstrate the impact and value of evaluation activities. A counterfactual construction has, moreover, been emphasized for the accurate determination of the impacts and effects that can be ascribed to an intervention [58].

Valuing is often context-specific and refers to the extent to which a programme is able to meet its needs. Although it may not be possible to quantify the value, it will always be possible to identify and articulate it. Valuing may include criteria concerning relevance, flexibility, adequacy, feasibility and predictability, as well as efficiency (cost-effectiveness or cost-benefit) [35,59,60], and aspects related to democracy and social justice [61] such as legitimacy, transparency and equity [60].

The use of evaluations has become an essential element of evaluation practice, and in order to secure use, theorists stress stakeholder identification and involvement [35,62,63]. Moreover, many scholars emphasize the need to involve key stakeholders in the entire evaluation process in order to ensure an outcome that is relevant, and results that are useful to users [35,63]. It should also be noted that there are numerous approaches to making use of (or misusing) evaluation results, including instrumental, conceptual and political to mention a few (see for example [35,28,64,65]).

In addition to the overarching dimensions of evaluations presented above (I, II, and III), evaluation theory stresses the importance of evaluating the process of the intervention, i.e., the preparation of the intervention, the involvement in the intervention process and the timing of the intervention [35,63].

3.3. Knowledge-based learning – An interdisciplinary approach focusing on energy efficiency in buildings

Transition research and evaluation theory provide a solid starting point for the evaluation of transformative change in society. Nevertheless, the evaluation of transitions in relation to energy efficiency in buildings will require an interdisciplinary approach. In this study, we recognise the need to consider research areas related to the evaluation of research (sociology of science) and the evaluation of policy instruments (policy analysis); we have also chosen to specifically focus on the evaluation of behaviour.

3.3.1. Evaluating research – An introduction based on the sociology of science

When evaluating research in the field of energy efficiency, the sociology of science provides a relevant theoretical starting point. In general, research evaluation has been based on the methods described by Bush [66] and Merton [67], in terms of peer review (collegial analysis) and quantitative indicators (e.g. bibliometry). Over time, increasingly sophisticated versions of coupled models have been developed in an attempt to capture the relationships between demand, knowledge production, social interests and other intervening factors. The main aim has been to assess the mechanisms behind the interaction and engagement between stakeholders, the dynamics in problem formulation and agenda setting, the level of reflexivity in assessments, and openness to variation and change in multi-actor and multi-purpose constellations.

Today, the practice of research evaluation is broad and multifaceted [68,69], with evaluations focusing on, for example, outcome and impact, networks, mobility, partner collaboration, and quality, as well as on processes and administration [69–71]. The methods and approaches used in research evaluation are varied, and include e.g. bibliometry and peer review [69,70], but also methods based on metrics

concerning patents and knowledge dissemination. The evaluation of processes and administration is approached from the perspectives of funding policies, from assessments of priority-setting etc. [68,72], or from assessing networking and support of knowledge production [73,74].

3.3.2. Evaluating policy instruments – An introduction to policy analysis

The complexity of evaluating policy instruments, especially in the field of energy efficiency in buildings, cannot be overlooked and calls for evaluation theory to be complemented with policy analysis. Policy analysis provides insights into policy processes with regards to how they are managed, and by placing the policy in relation to the societal expectations and systems that they are part of [75]. Such a system-oriented focus is important in capturing the inherent complexity of transformative change [31], and also calls for the recognition of the development and enforcement of policy mixes, and their role in a transformative change process [76,77]. Some key characteristics of the instruments and the processes are consistency, coherence, credibility and comprehensiveness [77]. Other critical aspects that are highlighted within policy analysis are issues related to policy instruments based on incorrect assumptions [78], implementation failures [79], side-effects [35], rebound effects [80], legitimacy [81] and ensuring learning through the application of policy evaluations [82]. Policy analysis also stresses the analysis of policies in terms of multi-layered policy, i.e. multi-level governance [83], stressing the interdependence of governmental and non-governmental actors at different levels.

The literature in policy analysis that focuses on market transformation provides tools and methods for the analysis of changes in markets (see, for example [84–86]). Neij and Åstrand [49] specifically stress the evaluation of outcome to support the evaluation of long-term transformative change when evaluating policy instruments on innovation; the use of outcome indicators is suggested in terms of e.g. technology performance, price development, changes in manufacturers' assortment, sales data, changes in knowledge, awareness and behaviour of important actors.

3.3.3. Evaluating behaviour

Energy efficiency behaviour is studied using a wide variety of disciplinary approaches, including technology, environmental sciences, architecture, economics, behavioural economics, psychology, geography, sociology and science-technology studies [87]; all of which can be relevant when evaluating transitions in energy efficiency in buildings to advance our understanding of behavioural change. The social science research on behaviour may, however, be of special importance as it provides an understanding of behaviour per se at both the individual and collective level.

Traditionally, energy efficiency behaviour has been addressed from the perspective of rational choice, departing from attitude theory with a focus on link information, knowledge, attitudes and individual choice, combined with the expectations that monetary savings and moral benefits would be sufficient to motivate behavioural change (e.g. [88], see also the debate between Shove [89] and Whitmarsh [90]). Several scholars argue for the need of additional examinations of processes, meaning, understandings and motivations [91]. This is argued, for example, in social practice theory, which has been proposed as a way of understanding the social context of people's behaviour [88,89,92,93]. Environmental psychology also emphasizes the importance of context, focusing on the individual as an integrated part of its social and physical environmental setting [94,95], and the need to define the target of behaviour and the drivers (antecedents) of behavioural change as well the alignment of focus and the tools employed. The importance of evaluating change in terms of overcoming individual behavioural failures is also emphasized in behavioural economics [96–98].

In all, different research perspectives on human behaviour complement each other in their use of theory to formulate the research question, in their methodological approach to provide the answers, and in

their integrative perspectives on the role of human behaviour in energy efficiency [99,100]. Behavioural research further emphasizes elements of transformative change, closely linked to the transition literature in terms of visioning. Jasanoff and Kim [101] highlight the importance of imaginary futures, stating that social and political life are highly dependent on the capacity to imagine what the future might be like, and engage with the future as a sociotechnical imaginary (visions, symbols and the feelings people have about energy efficiency). This imagination is found not only in individuals but more widely in society. These imaginaries influence behaviour, individual and collective identity, and the development of narratives, policy and institutions.

3.4. Key concepts for a transformative evaluation approach

In Section 3.1, we identified key concepts in the transition literature that can be applied in the evaluation of transformative change: *visioning*, *experimentation* and *learning*. In relation to these concepts, *path dependencies* and *lock-in effects* should be applied in the assessment to capture any potential hurdles to transitions. Moreover, we identified the importance of designing the evaluation using a *systems perspective* assessing aspects such as *side-effects* and *rebound effects* in the transition literature, as well as in other literature. The need for *multi-actor approaches* and *analyses* to advance relevant and essential transformative knowledge is also stressed in the literature.

In terms of providing learning and knowledge-based guidance for transformative change in the area of energy efficiency in buildings, we identified the following key concepts for the evaluation of research and policy interventions: the application of *multiple methods and multiple criteria*, and the use of *counterfactuals*. These concepts, emphasized in all the reviewed disciplinary fields reviewed, will be crucial for supporting transformative change. In addition, all the disciplinary fields considered contributed a wide selection of evaluation concepts that can enrich the evaluation and ensure that it is thoroughly interdisciplinary.

4. Evaluations of energy efficiency programmes in Sweden

In this section, we present a review of existing evaluations commissioned by Swedish governmental authorities and funding agencies focusing on energy efficiency in buildings (Literature review 2, see Fig. 1) (see also [102,103]). Our aim is to assess the extent to which existing evaluations already today provide experience in addressing transformative change and knowledge-based learning. The evaluations were reviewed based on the conceptual review presented in Section 3, which revealed a number of aspects relevant for advancing knowledge on how transformative change can be supported in general, and in buildings specifically. In Section 4.1 we present the scope of the evaluations reviewed (evaluations included are presented in Appendix A), in Section 4.2 we discuss the links to key concepts in the transition literature, and in Section 4.3 we elaborate further on the nature of knowledge-based learning for transformative change. As mentioned in the Introduction, behavioural aspects were largely overlooked in the reviewed evaluations commissioned by the Swedish governmental authorities, and in order to obtain an indication of the way in which behavioural aspects have been evaluated, we have included a review of research papers on behavioural aspects in Section 4.4. Our primary ambition is to identify strengths and weaknesses in current evaluation practice, to obtain an idea of what has been done and what is missing, and how to design an overarching transformative evaluation approach based on what has already been done.

4.1. Scope of research and policy programmes

Our review of the Swedish evaluations revealed a wide scope of research programmes and policy instruments applied to support energy efficient buildings. The research programmes covered a variety of aspects related to energy efficiency in buildings (see Appendix 1),

examples of which are given below:

- The design of energy efficient buildings, e.g. energy-plus houses, near-zero-energy buildings, passive houses, innovative solutions for renovating the existing building stock, as well as targeted efforts for energy efficiency solutions adapted for cultural heritage buildings
- Energy efficient technological systems, e.g. energy efficient lighting systems, energy efficient ventilation, cooling and heat pump systems, new materials and new types of insulation
- Control and steering devices for energy efficiency, smart meters, wireless sensors, etc.
- Information devices for visualising energy efficiency through design and gaming
- Efforts directed towards behaviour and sustainable lifestyles
- Policy-related aspects, e.g. knowledge development through collaborative processes, informative efforts, design and experience of policy instruments, involvement of local actors in energy efficiency measures, interaction between stakeholders and energy efficiency processes, locally, nationally and internationally
- Business models for market introduction of energy efficiency measures

In terms of policy instrument evaluations, the evaluations covered different types of policy instruments (see Appendix 1), such as:

- Regulatory instruments: building codes and energy requirements
- Financial instruments: subsidies and tax reductions for energy efficiency measures and subsidies for performing energy audits
- Informative instruments: demonstration projects, municipal energy advisory programmes, energy performance certificates for buildings, online information portals, facilitating energy services, and public authority best practice
- Other instruments: technology procurement programmes and cooperative network programmes for energy efficiency in buildings

These policy instruments involved various technological solutions (e.g. near-zero energy buildings, passive houses, building envelope, insulation, windows, heating, cooling and ventilation systems, control devices, ICT, appliances and office equipment), as well as behavioural aspects.

4.2. Evaluating transitions in practice

The broad outline and application of research programmes and policy instruments illustrated in Section 4.1 provides valuable support for transformative change in energy efficiency in buildings. Such a transformative process was not captured in the evaluations reviewed (see Table 1) – however, this was not the objective of the evaluations. In order to provide a more transformative evaluation approach, which may complement the current evaluation practice, we highlight the transformative strengths and weaknesses in the current evaluation practice below.

The review of the evaluations indicated a tendency for rather fragmented evaluations in which the focus was on separate programmes, with only limited attention to systemic transformative change in society (see also [104]). While it was noted that the need for a transition in the building sector was outlined as the ultimate rationales for some evaluations (see for example [105,106]), it was more common to present the evaluands in terms of their role in reaching or supporting national and international (political) goals.

As mentioned in Section 3, transition research stresses the need for visions, experiments and learning to support transitions. In the reviewed evaluations, *visions* were often expressed in terms of national and international goals, for example Sweden's energy goals for the years 2020 and 2050, or the vision of all new construction being near-zero buildings. The evaluands were then presented in their capacity to support

Table 1

Addressing of selected concepts that support evaluation of a transition towards energy efficiency in buildings, in reviewed Swedish evaluations of research programmes/institutions and policy instruments.

	Evaluations of research programmes	Evaluations of policy instruments
Guiding visions	13 of 20 evaluations present the evaluand in relation to its supporting of fulfilling overarching (political) goals or its role in supporting energy efficiency through a clearly stated programme vision.	28 of 33 evaluations present the evaluand in relation to its supporting of fulfilling overarching (political) goals or its role in supporting energy efficiency through a clearly stated programme vision.
Evaluation of experimentation	Often implicit; research is seen as in nature experimental.	12% of the evaluations considered experimenting efforts of the evaluand.
System aspects: side-effects	1 of 20 evaluations outlined potential side-effects.	9 of 33 evaluations outlined potential side-effects.
System aspects: rebound effects	No assessments of rebound effects have been performed.	1 of 33 evaluations acknowledged rebound effects, but did not make any quantitative assessment of them.
Multi-actor analysis	None of the evaluations presented a multi-actor analysis.	None of the evaluations presented a multi-actor analysis.

reaching these goals, but the evaluations did not thoroughly address the extent to which the evaluand contributed to fulfilling a broader vision (and roadmap) of transformative change in society. As for capturing transformative contributions in terms of *experimentation* and testing, e.g. through assessing innovative efforts, this was commonly implied as a starting point in the evaluations of research programmes aimed at energy efficiency. Examples of such programme evaluations included research aimed at developing new, more energy efficient technologies, such as lighting products [107] or heat pumps [108]. The evaluations of policy instruments, on the other hand, focused less on transformative progression in terms of innovative efforts and contributions, i.e., they focused less on transformative outcome (providing support for long-term transitions) and more on programme impact. While impact, for example expressed in terms of kilowatt hours saved, provides information on the efficiency of the (final) energy use in buildings, it does not provide any information on the transformative processes in terms of advancements in e.g. technological performance, price development, changes in manufacturers' assortment, changes in knowledge, or the awareness and behaviour of important actors. Nevertheless, transformative efforts were analysed and assessed in terms of outcomes in some evaluations of policy instruments, in terms of, for example, access to energy efficient products on the market, knowledge- and competence-building among stakeholders, as well as changes in actors' interest and motivation in taking energy efficiency measures (see for example [109]). Moreover, certain policy instruments more prone to experimentation, such as cooperative network programmes, technology procurement and demonstration programmes, also elaborated on policy design for transformation, its ability to facilitate experimentation and its intended outcome (see for example [106,110,111]). These evaluations elaborated, to some extent, on outcome, risks and potential upscaling. *Learning* on a societal level was seldom discussed in the evaluations. Instead, attention was mainly directed towards the programme level, and how the programme itself could be improved or modified to support further goal achievement or more efficient implementation, while overarching discussions on how such improvements could support large-scale changes were largely lacking. The reviewed evaluations did not include any analysis of any potential *lock-in effects* or *path dependencies*, meaning that structures in the current systems, with regards to technologies and infrastructures, as well as preferences that may hinder new, more energy efficient practices to be up-scaled, were not regarded or

uncovered.

The need for a systems perspective in order to understand and support transformative change was underlined in both the transition research literature and the policy analysis literature. Our review shows that the evaluations of policy instruments did only cover side-effects and rebound effects to a limited extent, which suggests that there is a shortfall in the extent to which focus was deliberately placed on effects outside of the programme or institution boundaries; only one evaluation of a policy instrument included an assessment of side-effects [112]. Evaluations of research often connected the research programme to other governance areas, often in conjunction with discussions on funding interests, but in a few instances also from the perspectives of how the research programme had spurred effects on the energy system at large, or how knowledge produced within a programme could become generally useful in society (see for example [113]).

Another key issue that is important to consider in a systems perspective is the involvement of actors in an evaluation. The majority of the reviewed evaluations involved stakeholders on some level, but mainly only for data collection. Some evaluations involved stakeholders from different stakeholder groups, thus moving towards a multi-actor approach. However, the involvement of actors from different actor groups merely for data collection does not mean a multi-actor analysis, which should seek to shed light on different actors' roles, responsibilities and expectations in a transition process. Thus, while we did see involvement of actors from various stakeholder groups in the same evaluations, the focus was on their perceptions of the evaluand, and its effects, rather than an analysis of their interconnections and dependences on each other (in a transformative process).

The reviewed evaluations were conducted at various stages of the implementation process, from *ex ante* to *ex post*, and some research programmes and policy instruments were evaluated several times, for example the energy advisory programme [114–118]. In the cases of recurring evaluations, some alignment was seen between the evaluations, in that they built on each other. However, assessments of alignment between evaluations of the research and policy instruments were generally scarce. In particular, there were no thorough attempts to assess any possible alignment between research and policy instruments, whereupon the evaluations in general did not provide a systemic overview of research and policy as essential parts of transformative change. Neither did the evaluations commissioned by authorities consider the alignment between governmental and non-governmental actions (e.g. municipal initiatives and measures taken by actors in the private sector).

4.3. Evaluating knowledge-based learning in practice

Learning, through a reflexive and flexible approach to processes of change, is central to the transition towards improved energy efficiency in buildings; and such learning can be advanced through knowledge-based evaluations. The literature review presented in Section 3, identified key concepts such as *multiple methods*, *multiple criteria* and *counterfactuals* as crucial for supporting knowledge-based learning and transformative change. Below we use these concepts, and others identified in the review presented in Section 3, in the review of the current evaluation practice.

In all, the results revealed that the reviewed evaluations were, to a large extent, based on theoretical approaches to evaluation, providing learning from the interventions evaluated (see also [102,103]). We identified the application of core concepts based on evaluation theory, the sociology of science and policy analysis in the evaluations, see Table 2 (as outlined in Sections 3.2 and 3.3). In general, the knowledge-based evaluation approaches identified in the reviewed material must be regarded as robust, however, we identified both strengths and weaknesses in current evaluation practice.

Our review of the evaluations of *research programmes* showed that the evaluations included assessments of outcome and impact, as well as processes in individual programmes. Outcome and impact have been

Table 2

Addressing of selected concepts that support knowledge-based learning in reviewed Swedish evaluations of research programmes/institutions and policy instruments.

	Evaluations of research programmes	Evaluations of policy instruments
Multiple value criteria	20 of 20 evaluations (100%)	21 of 33 evaluations (64%)
Multiple methods	13 of 20 evaluations (65%)	32 of 33 evaluations (97%)
Counterfactual construction	8 of 20 evaluations (40%)	15 of 33 evaluations (45%)

evaluated using various methods to provide insights into e.g. research dissemination and the effects and implications for different stakeholder groups (academia, society, companies, etc.) as well as collaboration between stakeholders (see for example [119]). In relation to processes, some evaluations covered networks and partner collaboration, involving industry, market actors, authorities and other societal agents (see for example [120]). Processes were also evaluated in terms of research organisation, administrative processes and funding mechanisms (see for example [121]). Up to four different methods were combined in some evaluations, including documents, interviews, survey analysis and interviews in groups, to support learning [108,113,121,122].

The reviewed evaluations of *policy instruments* for energy efficiency in buildings showed a strong focus on assessing the impact (e.g. the number of kWh saved by the realization of energy efficiency measures) and the effectiveness, while other evaluations also included cost-effectiveness [123], feasibility/administrative burden, and coordination with other policy instruments [124]. Few, or none, of the evaluations applied value criteria such as relevance, predictability, acceptability, legitimacy, transparency, appropriateness, adequacy, or responsiveness, i.e. criteria that can provide deeper knowledge about transformative contributions of the evaluand. Such criteria could, for example, advance knowledge regarding risks and potential upscaling of processes that are essential for a transition.

Evaluations of policy instruments also showed a frequent use of multiple methods to support learning, often with an emphasis on qualitative approaches, some even combining four methods (document analysis, interviews, survey analyses, and interviews in groups [125,126]). In spite of the general application of multiple methods, analysts or data sets, these were largely not combined in order to triangulate and validate findings, but rather to provide data or assessments on certain issues.

Both evaluation theory and the policy analysis literature stress the importance of a counterfactual construction for the assessment of the actual impact of any policy instrument and for supporting processes of learning. The review shows that additionality was assessed in approximately half of the evaluations and that counterfactuals were sometimes even derived from multiple methods [120,127,128].

4.4. Evaluating behaviour in practice

As mentioned above, the evaluations commissioned by Swedish authorities did not include analyses of behavioural aspects; the evaluations that did, had been performed by researchers as part of diverse research programmes. Thus, the review presented in this section is based on scientific peer-reviewed papers presenting evaluations of energy efficiency behaviour. The review includes empirical studies carried out in Sweden, across different disciplines, including two types of interventions, namely information and feedback and building design⁴ (see

⁴ This means that the review does not cover “nudging” interventions or energy policy interventions in terms of e.g. regulation and economic incentives, for example.

Appendix A).

Our review of the relevant research studies shows that energy efficiency behaviour has been studied from a wide variety of disciplinary approaches. However, technological studies dominate (10 out of 26), followed by sociology (science-technology-society studies) (5 out of 26), psychological studies (5) and economic studies (1). The approaches employed differ with regard to the identification, definition and assessment of behaviour, and there is a clear difference in the outcomes considered when using different disciplinary approaches.

The results of this review indicate that most of the evaluations developed from technical approaches are lacking specifications of important sociological and psychological variables. For example, the evaluations performed within the discipline of technology do not consider the characteristics of the target group or the role of their social and environmental context, or the fit between their current behaviour/practices and choice of intervention. A “one size fits all” approach risks the loss of important nuances in the effect of an intervention. An intervention may, for example, have a different outcome depending on whether it is introduced to a person, or group of people, with an established habit or not. A certain practice or behaviour may be facilitated or hindered by group norms. In many cases, evaluations based on technical approaches have weak theoretical support for the analysis of behaviour, and thereby disregards for example how likely it is that the expected change will occur in a certain target group in response to an intervention [129]. Nevertheless, evaluations based on technical, or architectural, approaches provide detailed information regarding building-related factors, and, to some extent, other physical contextual factors. Such information is important to drive the technical development. These approaches also provide actual measurements of energy use.

The results also show that evaluations based on sociology, including science and technology-studies (e.g. [130–133]), and psychology, including environmental psychology (e.g. [134–138]), are in general based on different but well-established solid theoretical frameworks such as social practice theory, goal framing, group identity, theory of affordances providing a rich understanding of behaviour/practice and the drivers of behavioural change. In the very few studies employing an economic approach, the understanding of the drivers departing from social norm theory, for example, is linked with theory on bounded rationality (e.g. [139]). In these evaluations behaviour is interwoven with its context, and by directing attention to the role of the physical and social environmental contexts of the behaviour, a better understanding can be gained of how energy efficiency behaviour may be promoted or hindered by certain environmental attributes. Sociology in particular (science-technology-studies) directs attention to the role of the social context of energy efficiency behaviour using the household as the unit of analysis and addressing the household’s interaction with energy advisers and other energy experts or managers, architects and construction companies (e.g. [140–142]). However, the individual’s behavioural antecedents beyond socio-demographics, as well as behavioural spillover and long-term effects, are often neglected in this approach. Evaluations based on environmental psychology focus on the behaviour of an individual (in a given environment) and provide a framework for the analysis of the alignment in the individual’s motivation for behavioural change, the intervention introduced, the target behaviour and the effects of the target behaviour, other behaviour (spillover) and the effects over time (e.g. [135,137]). Nevertheless, the interdependence between the behaviour of people, such as interactions between family members in a household, is of less concern in this approach.

Our review thus showed that evaluations performed by researchers are fragmented and strongly linked to different disciplines. However, the different approaches used to analyse energy efficiency behaviour complement each other. Whereas evaluations based on a technological approach provide measurements of energy use, for example, evaluation based on sociology and science-technology studies provide a complementary nuanced contextual understanding of energy efficiency,

elaborating on the role of people’s daily practices. Environmental psychology provides additional insights at the individual level. The results of this review indicates that greater effort is needed in coordinating evaluations based on different disciplinary approaches to achieve a more holistic understanding of what interventions work in reducing energy use, in what target group and in what social and environmental context, in order to support a transformative change in society. The integration of theoretical approaches in the evaluation of interventions would provide specific data on energy use together with and a nuanced understanding of behaviour/practice allowing for further development of interventions. Some attempts have been made to do this by researchers in environmental sciences. As an example, Nilsson et al. [143] combined the technical aspects of energy use measures with a systematic quantitative investigation of psycho-social variables.

5. Discussion

The transition towards more energy efficient buildings will be shaped by complex processes involving a range of actors in the reconfiguration of socio-technical systems. If evaluations are to capture such change processes, they need to provide an overarching perspective that advances transformative knowledge. If they are merely operationalized and conducted as individual attempts at assessing programme effects, they will not be able to depict the potential of such systemic change, and provide genuine support for a transition.

In this paper we argue that the realization of a transition towards energy efficiency in buildings will require a vision and a deliberate plan, realized through the implementation of a range of research programmes and policy incentives – which in turn need to be evaluated using an interdisciplinary approach that allows for gauging of transformative effects, that provides knowledge-based learning, and that goes beyond the evaluation discourse of today. Below, we discuss how current evaluation practice can be improved towards a transformative evaluation approach in the Swedish context.

5.1. Key messages from current evaluation practice

Our review of evaluation practices in Sweden included 53 evaluations of a variety of research and policy programmes aimed at improving energy efficiency in buildings; the additional review of 26 behavioural studies provides an even richer foundation for the assessment of current evaluation practice. Together, these evaluations provided important information on the achievements made towards energy efficiency; they cover a broad range of research programmes and policy instruments, and they rely on a robust foundation consisting of various evaluation approaches and methods. In general, the review revealed a rigorous evaluation practice based on a combination of rationalistic and constructivist evaluation approaches, highlighting the need to evaluate predefined goals, as well as the need to advance relevant knowledge (learning).

The weakness of the current evaluation practice is the fragmentation of approaches, which hampers transformative change. Several evaluations have been commissioned by Swedish authorities and funding agencies, but we found no overarching plans for how the evaluations are to complement each other and contribute to learning beyond the individual programmes, i.e. learning in relation to any overarching transformative vision or plan of the evaluated research and policy interventions. The scope of the evaluations tends to be rather narrow, focusing on the tangible outcomes of specific measures and programmes, rather than the complex interplay between technologies, institutions and actors. Evaluations of research programmes and policy instruments are not aligned, despite the fact that the evaluations were commissioned, or even performed, by the same actors. This, however, provides good conditions for improvements.

The study further shows that in evaluations commissioned by agencies, behavioural aspects were treated in a rudimentary way.

Nevertheless, such evaluations have been designed and developed by individual researchers in various research programmes. These evaluations rely on a number of different methodological approaches, however, most of the evaluations reviewed have been performed by researchers with a background in engineering, and we found limited attempts to broaden the evaluation approach across engineering and behavioural sciences or to combine different social science approaches, such as sociology and psychology, to understand behaviour. Learning could be improved by broadening and strengthening the methodological approaches in behavioural evaluations, including not only existing theoretical approaches, but also new theoretical approaches such as behavioural economics. The different approaches have strengths and weaknesses and complement each other to provide a holistic understanding of the potential of different interventions, as discussed in the text above.

In general, the reviewed evaluations did not explicitly focus on transformative change, however, many evaluations included aspects relevant for the assessment of transformative change. A systems perspective was often adopted, but this could be improved, for example by greater involvement of various stakeholders or others that are affected by the implementation of a programme or policy. Evaluations focusing on transformative change should preferably also seek to address drivers of change, such as visioning, experimentation and learning. The evaluation of innovative efforts, through experimentation, was common in the evaluation of research programmes but weaker in the case of policy instruments. Evaluations of the outcome of policy instruments did not address the advancement of knowledge relevant for transformative change, and little attention was paid to potential upscaling. Whereas evaluations of behaviour in theory provide tools for the evaluation of experiments and knowledge relevant for upscaling, the strength of the research design applied in the reviewed evaluations varied, and the strength of current evaluation practice lies in the nuanced understanding obtained from case studies.

In relation to learning, our review indicates a sound and robust evaluation practice that do apply key concepts for providing knowledge-based learning, however, we also identified some weaknesses. In general, several methods were applied in the evaluations, but triangulation was rare. Given that current evaluation practice already is applying multiple methods and data sources, we see a potential to move for an increased robustness in the evaluation results by validating results using triangulation. The review also revealed some good practices in the undertaking of counterfactual analysis when performing impact assessments, particularly in a few evaluations where results were synthesised from different methods when constructing the counterfactual. However, the overall application of counterfactual analyses can be further improved, regarding both frequency and methodological robustness.

On the whole, attention is mainly focused on assessments of impact and effectiveness, whereas other criteria that could support assessments of transformative contributions, or could provide insights into processes related to achieving transformative contributions, were underrepresented in the material reviewed. Example of such criteria include relevance, to ensure societally relevant implementations, acceptability or legitimacy to ensure the mandate among policy makers to support energy efficiency measures in buildings, and predictability, to ensure administratively effective implementations and to prepare stakeholders for the implications of the implementations. Neither did the evaluations include any assessment of path-dependency or potential lock-in effects that the implementation may lead to, or may challenge.

5.2. A transformative evaluation approach

Based on the need to strengthen the support for a transformative change, we propose the application of a transformative and interdisciplinary evaluation approach. This approach provides a comprehensive overview of the effects of various interventions altogether. Building on coordination and combination of individual evaluations of relevant

research programmes and policy instruments, the proposed approach provides systemic and structured learning. Individual evaluations should be part of an assessment approach that synthesizes results and sheds lights on the momentum of multiple programmes to stimulate and support transformative change. In the design of such an evaluation approach, we underline the need to link evaluation research to transition research so that evaluations can be designed to support transformative change in society. The transition literature could be useful in developing an overarching framework based on the concepts of visioning, experimentation and learning. Such a framework would provide a direction (visioning) and emphasize the need for novelty and innovative efforts (experimentation), as well as interdisciplinary learning.⁵ The framework could be translated into an evaluation approach including the following elements:

1. A vision (including roadmaps and goals) describing the intended direction in the building sector and envisioned effects of the interventions⁶. The results of the evaluations can then be discussed in relation to the vision in terms of whether, and to what extent, the interventions have contributed to the intended transformative change.
2. A plan for the evaluation of experimental initiatives, for example in the form of research, demonstration projects, and policy labs, to provide an understanding of progress in terms of novelty and innovative effects, and the potential for upscaling.
3. A plan for the advancement of learning, i.e. the identification of different types of evaluations in different disciplinary fields, such as evaluation theory, the sociology of science, policy analysis, and behavioural related research, to structure the review of energy efficiency in buildings.

Individual evaluations should complement each other, by being performed by relevant evaluators, using insights and methods from various disciplines, and answering different questions. When combined, the results of these individual evaluations can support a holistic perspective and a structured learning. The evaluations included in such an evaluation plan must, thus, be linked to ensure overall learning and knowledge transfer. Barriers hindering a transition must be revealed and addressed. The overall evaluation plan should be developed by actors from different parts of society and should be reviewed at regular intervals.

We argue that the current evaluation practice currently applied in Sweden (focusing on energy efficiency in buildings) has the potential to be modified to provide an interdisciplinary evaluation approach for transformative change. This can be done, for example, by strengthening coordination between evaluations, providing better vision documents (visions, roadmaps, missions and goals) of the envisioned transformative effects of various research programmes and policy instruments, and by providing stronger evaluations of innovative efforts and contributions. Evaluations could also strengthen learning further by improving the use of triangulation, counterfactual analysis and a wider range of value criteria. Several actors are already included in the evaluation process, and by structuring the inclusion of these actors, essential learning and alignment could be improved. Behavioural aspects should be captured in the evaluations, using complementary theoretical approaches.

Clearly, the proposed evaluation approach does call for an extension and revision of current evaluation practice, and additional research is

⁵ In this paper, we introduce a few key concepts from transition research which can be useful when evaluating research and policy instruments. However, we recommend evaluators to be inspired by this literature and include more transformative aspects in their evaluations.

⁶ The evaluation literature (programme theory) emphasizes the need to identify the supposed effects of the interventions and to analyse alternative instruments by means of *ex ante* evaluations.

needed for understanding the implications and requirements that come with an implementation of such a transformative evaluation approach. We identify two key areas to discuss further: the design and implementation of transformative evaluations, and the use of transformative evaluations. The design and implementation processes relate to the practices of commissioning and conducting evaluations, which may be fragmented between multiple institutions and actors with different mandates and jurisdictions. Such institutional boundaries may hinder coordination within and between agencies and actors, which may obscure the necessary birds-eye view on the need of knowledge as well as existing knowledge. It may also hinder timely exchange and use of knowledge created in evaluations. Further understanding is also needed concerning the processes of using, and for increasing use of, evaluations in order to support the adoption of a transformative evaluation approach among commissioners and evaluators. As for the conduct of evaluation, it should also be noted that transformative evaluations in large build on the application of interdisciplinary approaches, which in itself may give rise to challenges related to acquiring evaluation expertise, and overcoming barriers related to disciplinary expectations of both those commissioning and conducting the evaluations.

Another challenge that may need attention is the potentially limiting effects that policy designs have on evaluation designs. If a policy instrument is in fact promoting incremental change and short-term goals, then its evaluation may be bound to mirror this and may as such be less prone to support transformative efforts. This, however, merits discussions on the politics and processes of policy-making, e.g. with regards to the norms guiding decisions, the actors involved and the negotiations that may take place between them, and the envisioned societal goals, which promotes a screening of policy designs and their role for supporting a transition (see for example [35,144]).

Lastly, in the case of Sweden with an extensive amount of evaluations undertaken in the sector of building energy efficiency, conducted by authorities, by external actors and academia together, it should be emphasized that while there are challenges to address for adopting transformative evaluation, it is not a matter of starting over completely new, but rather creating infrastructures that allow coordination and a systemic overview of what is done where, to make sure that the knowledge created is taken stock off, and that it is being conveyed into a larger picture of contributions on a transformative scale.

6. Conclusions

This paper has argued that energy efficiency in buildings is a core aspect of the transformation of energy systems. Energy efficiency in buildings, however, is fragmented in both the policy and research areas, and few efforts have been made to form a coherent approach to assessing and evaluating a transformative change. In this paper, we have outlined and assessed the current state of affairs and outlined an alternative transformative evaluation approach in which we suggest complements to current evaluation practice as well as adjustments in the design of individual evaluations; this assessment of the current evaluation practice is framed in line with the results of a review of literature relevant for transformative evaluations of energy efficiency in buildings. The review of transition literature highlights key concepts such as *visioning*, *experimentation*, *learning*, *path dependencies*, *lock-in effects* and stress the importance of designing the evaluation using a *systems perspective* and *multi-actor approaches*; whereas the review of evaluation-oriented literature, providing learning and knowledge-based guidance, underlines key concepts such as *multiple methods (triangulation)*, *multiple criteria*, and the use of *counterfactuals*.

In the analysis of the current evaluation practice in the field of energy efficiency in buildings in Sweden we identified many, although fragmented, aspects characterising good evaluation practice. However, there is a need for an overarching, transformative, knowledge-based evaluation approach to support transformative change. This can be realized by modifying current evaluation practice, improving the

coordination of evaluations, and by providing a plan for an overarching and interdisciplinary evaluation approach for transformative change. Evaluations have traditionally been designed to ensure efficient use of resources in delimited areas of interventions. This cost-efficiency approach is however too myopic to meet ambitious, system-transforming goals. A transformative evaluation approach needs therefore to include visioning documents of the envisioned transformative effects of various research programmes and policy instruments, and strengthen the evaluations in terms of assessing and capturing innovative efforts, e.g. by evaluating outcomes. Moreover, learning from evaluations calls for individual evaluations to be aligned, but also requires further development of current evaluation practice, which can be strengthened, for example, in terms of triangulation for the validation of results, the application of counterfactual analyses, and a more diverse and conscious application of value criteria. Evaluations commissioned by governmental authorities and agencies should also include behavioural change to a greater degree than at present to further increase the applicability of evaluations. When evaluating behavioural aspects, it is important to take into account the fact that the different evaluation approaches have different strengths and weaknesses although they complement each other. In conclusion, the results of this study underline the need to link evaluation research with transition research in the design of an evaluation approach to support transformative change in society.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the reviewers for their thorough efforts in providing constructive and inspiring feedback on this manuscript. We would also like to thank the Swedish Energy Agency for supporting this work through the project Transition Governance (grant 39938-1).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.erss.2020.101864>.

References

- [1] GEA, Global Energy Assessment - Toward a Sustainable Future, Cambridge University Press, Cambridge, UK and New York, NY, USA and the International Institute for Applied Systems Analysis, Laxenburg, Austria, 2012.
- [2] IPCC, Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.
- [3] IEA (International Energy Agency), Energy Efficiency 2018 - Analysis and Outlook 2040, 2018.
- [4] European Commission, Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions - United in delivering the Energy Union and Climate Action - Setting the foundations for a successful clean energy transition, COM (2019) 285 final, 2019.
- [5] O. Lucon, D. Ürge-Vorsatz, A. Zain Ahmed, H. Akbari, P. Bertoldi, L.F. Cabeza, Buildings. In Climate Change (2014). Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.
- [6] WBCSD (World Business Council for Sustainable Development), A handbook on creating dynamic local markets for Energy Efficiency in Buildings, 2016.
- [7] M. Levine, D. Ürge-Vorsatz, K. Blok, L. Geng, D. Harvey, S. Lang, G. Levermore, A. Mongameli Mehlwana, S. Mirasgedis, A. Novikova, J. Rilling, H. Yoshino,

- Residential and commercial buildings. In Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.
- [8] K. Gillingham, K. Palmer, Bridging the energy efficiency gap: policy insights from economic theory and empirical evidence. *Rev. Environ. Econ. Policy* 8 (1) (2014) 13–38, <https://doi.org/10.1093/reep/ret021>.
- [9] Sutherland, “No Cost” efforts to reduce carbene emissions in the U.S: an economic perspective, *Energy J.* 21 No 3 (2000).
- [10] G.C. Unruh, Understanding carbon lock-in. *Energy policy* 28 (12) (2000) 817–830.
- [11] B. Boza-Kiss, S. Moles-Grueso, D. Urge-Vorsatz, Evaluating policy instruments to foster energy efficiency for the sustainable transformation of buildings, *Curr. Opin. Environ. Sustainability* 5 (2) (2013) 163–176, <https://doi.org/10.1016/j.coesust.2013.04.002>.
- [12] D. Urge-Vorsatz, L.D.D. Harvey, S. Mirasgedis, M.D. Levine, Mitigating CO2 emissions from energy use in the world’s buildings, *Build. Res. Inf.* 35 (4) (2007) 379–398, <https://doi.org/10.1080/09613210701325883>.
- [13] P.G. Boonekamp, W. Eichhammer, Applied policy measures for energy savings in end-use sectors in EU countries: what is best? Proceedings of the ECEEE Summer Study 2007 (2007) 263–273.
- [14] R. Lowe, Policy and strategy challenges for climate change and building stocks, *Build. Res. Inf.* 37 (2) (2009) 206–212.
- [15] L. Murphy, F. Meijer, H. Visscher, Effective national energy performance instruments for existing dwellings? Lessons from Front-runners. In: proceedings of Retrofit 2012 Conference, Salford Quays 24–26 January, 2012.
- [16] P. Bertoldi, Assessment and experience of white certificate schemes in the European Union, JRC European Commission. (2011).
- [17] M.J. Hoffman, Climate Governance at the Crossroads: Experimenting with a Global Response, Oxford University Press, New York, 2011.
- [18] N. Smedby, L. Neij, Experiences in urban governance for sustainability: the Constructive Dialogue in Swedish municipalities, *J. Cleaner Prod.* 50 (2013) 148–158, <https://doi.org/10.1016/j.jclepro.2012.11.044>.
- [19] K. Hölischer, J.M. Mittmaver, D. Loorbach, Transition versus transformation: What’s the difference? *Environ. Innov. Soc. Trans.* 27 (2018) 1–3.
- [20] F.W. Geels, Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study, *Res. Policy* 31 (8–9) (2002) 1257–1274, [https://doi.org/10.1016/S0048-7333\(02\)00662-8](https://doi.org/10.1016/S0048-7333(02)00662-8).
- [21] J. Köhler, F.W. Geels, F. Kern, J. Markard, A. Wieczorek, F. Alkemade, P. Wells, An agenda for sustainability transitions research: state of the art and future directions, *Environ. Innov. Soc. Trans.* 31 (2019) 1–32, <https://doi.org/10.1016/j.eist.2019.01.004>.
- [22] P. Kivimaa, P. Kangas, D. Lazarevic, J. Lukkariinen, M. Åkerman, M. Halonen, M. Nieminen, Transition towards zero energy buildings: Insights on emerging business ecosystems, new business models and energy efficiency policy in Finland, *SYKE Publications*, No, 2019, p. 5.
- [23] D. Lazarevic, P. Kivimaa, J. Lukkariinen, H.-L. Kangas, Understanding integrated solution innovations in sustainability transitions: Reconfigurative building-energy services in Finland, *Energy Res. Social Sci.* 56 (2019), 101209, <https://doi.org/10.1016/j.erss.2019.05.019>.
- [24] D. Hultema, A. Jordan, E. Massey, T. Rayner, H. van Asselt, C. Haug, The evaluation of climate policy: theory and emerging practice in Europe, *Policy Sci.* 44 (2) (2011) 179–198, <https://doi.org/10.1007/s11077-011-9125-7>.
- [25] M. Nilsson, A. Jordan, J. Turnpenny, J. Hertin, B. Nykvist, D. Russel, The use and non-use of policy appraisal tools in public policy making: an analysis of three European countries and the European Union, *Pol. Sci.* 41 (2008) 335–355.
- [26] F.L. Leeuw, J. Vaessen, Introduction: Mind the Gap; in: J. Vaessen, F.L. Leeuw (Eds.), *Mind the gap: perspectives on policy evaluation and the social sciences*, Transaction Publishers, New Brunswick, 2010.
- [27] H. Mela, M. Hildén, Evaluation of climate policies and measures in EU member states—examples and experiences from four sectors. *Finn. Environ.* 19 (2012).
- [28] C.H. Weiss, Evaluation: methods for studying programs and policies, 2nd ed., Prentice Hall, Upper Saddle River, N.J., 1998.
- [29] E. Vedung, 2009, *Utvärdering i politik och förvaltning*, Studentlitteratur, Lund, 2009.
- [30] C. Haug, T. Rayner, A. Jordan, R. Hildingsson, J. Strippel, S. Monni, Navigating the dilemmas of climate policy in Europe: evidence from policy evaluation studies, *Clim. Change* 101 (3–4) (2010) 427–445, <https://doi.org/10.1007/s10584-009-9682-3>.
- [31] M. Hildén, A. Jordan, T. Rayner, Climate policy innovation: developing an evaluation perspective, *Environ. Polit.* 23 (5) (2014) 884–905, <https://doi.org/10.1080/09644016.2014.924205>.
- [32] J. Kahn, M. Harmelink, R. Harmsen, W. Irrek, N. Labanca, From Theory Based Policy Evaluation to Smart Policy Design – Summary report of the AID, ECOFYS Netherlands.
- [33] J. Wade, N. Eyre, Energy efficiency evaluation: the evidence for real energy savings from energy efficiency programmes in the household sector, UK Energy Research Centre, London, 2015.
- [34] H. Chen, *Theory-driven evaluations*, Sage Publications, Newbury Park, Calif, 1990.
- [35] E. Vedung, *Public policy and program evaluation*, Transaction Publishers, New Brunswick, N.J., 1997.
- [36] E. Furubo, R. Sandahl, *Coordinated Pluralism - The Swedish Case*, in: J.-E. Furubo, R.C. Rist, R. Sandahl (Eds.), *International Atlas of Evaluation*, Transaction Publishers, New Brunswick, N.J., 2002, pp. 115–128.
- [37] D. Loorbach, Transition management for sustainable development: a prescriptive complexity-based governance framework, *Governance* 23 (1) (2010) 161–183, <https://doi.org/10.1111/j.1468-0491.2009.01471.x>.
- [38] J.-P. Voss, B. Bornemann, The politics of reflexive governance: challenges for designing adaptive management and transition management, *Ecol. Soc.* 16 (2) (2011) 9.
- [39] M. Bussels, S. Happaerts, H. Bruininxck, Evaluating and Monitoring Transition Initiatives: lessons from a field scan, *Research Paper 5*, HIVA, KU Leuven, Leuven, 2013.
- [40] A. Rip, R. Kemp, Technological change, in: S. Rayner, E.L. Malone (Eds.), *Human Choice and Climate Change*, 2, Battelle Press, Columbus, Ohio, 1998, pp. 327–399.
- [41] J. Rotmans, R. Kemp, M. Van Asselt, More evolution than revolution: transition management in public policy, *Technol. Anal. Strateg. Manage.* 18 (3) (2006) 285–298.
- [42] P. Kivimaa, F. Kern, Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions, *Res. Policy* 45 (1) (2016) 205–217, <https://doi.org/10.1016/j.respol.2015.09.008>.
- [43] F. Berkhout, Normative expectations in systems innovation, *Technol. Anal. Strateg. Manage.* 18 (3) (2006) 299–311.
- [44] M. Bourap, N. Brown, K. Konrad, H.V. Lente, The sociology of expectations in science and technology, *Technol. Anal. Strateg. Manage.* 18 (3) (2006) 285–298.
- [45] D. Loorbach, J. Rotmans, Managing Transitions for Sustainable Development, in: X. Olsthoorn, A.J. Wiecek (Eds.), *Understanding Industrial Transformation*, Springer, 2006.
- [46] F.W. Geels, R.P.J.M. Raven, Non-linearity and expectations in niche-development trajectories: ups and downs in Dutch biogas development (1973–2003), *Technol. Anal. Strateg. Manage.* 18 (3/4) (2006) 375–392.
- [47] R. Kemp, J. Schot, R. Hoogma, Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management, *Technol. Anal. Strateg. Manage.* 10 (2) (1998) 175–198, <https://doi.org/10.1080/09537329808524310>.
- [48] F. Nevens, N. Franzeskaki, L. Gorissen, D. Loorbach, Urban Transition Labs: co-creating transformative action for sustainable cities, *J. Cleaner Prod.* 50 (2013) 111–122, <https://doi.org/10.1016/j.jclepro.2012.12.001>.
- [49] L. Neij, K. Åstrand, Outcome indicators for the evaluation of energy policy instruments and technical change, *Energy Policy* 34 (17) (2006) 2662–2676, <https://doi.org/10.1016/j.enpol.2005.03.012>.
- [50] R. Raven, J. Schot, F. Berkhout, Space and scale in socio-technical transitions, *Environ. Innov. Soc. Trans.* 4 (2012) 63–78, <https://doi.org/10.1016/j.eist.2012.08.001>.
- [51] A. Smith, R. Raven, What is protective space? Reconsidering niches in transitions to sustainability, *Res. Policy* 41 (6) (2012) 1025–1036, <https://doi.org/10.1016/j.respol.2011.12.012>.
- [52] M.C. Alkin (Ed.), *Evaluation roots: a wider perspective of theorists’ views and influences*, 2nd ed., SAGE Publications, Los Angeles, 2013.
- [53] W.R. Shadish, T.D. Cook, L.C. Leviton (Eds.), *Foundations of program evaluation: theories of practice* (Reprinted), Sage Publications, Newbury Park, 1991.
- [54] E.G. Guba, Y.S. Lincoln, *Effective evaluation*, 1st ed., Jossey-Bass Publishers, San Francisco, 1981.
- [55] E.G. Guba, Y.S. Lincoln, *Fourth generation evaluation*, Sage Publications, Newbury Park, Calif, 1989.
- [56] R. Pawson, N. Tilley, *Realistic evaluation*, Sage, London; Thousand Oaks, Calif, 1997.
- [57] M. Scriven, *Evaluation thesaurus*, 4th ed., Sage Publications, Newbury Park, Calif, 1991.
- [58] P.J. Ferraro, Counterfactual thinking and impact evaluation in environmental policy, *New Directions Eval.* 122 (2009) 75–84, <https://doi.org/10.1002/ev.297>.
- [59] W.L. Dunn, *Public policy analysis: an integrated approach*, Sixth Edition. ed. Routledge, Taylor & Francis Group, New York, 2018.
- [60] P. Mickwitz, A framework for evaluating environmental policy instruments context and key concepts, *Evaluation* 9 (2003) 415–436, <https://doi.org/10.1177/1356389003094004>.
- [61] E. House, K.R. Howe, *Values in Evaluation and Social Research*, SAGE, 1999.
- [62] L.J. Cronbach, *Course improvement through evaluation*, Teachers College Record 64 (1963) 672–683.
- [63] M.Q. Patton, *Utilization-focused evaluation*, 4th ed., Sage Publications, Thousand Oaks, 2008.
- [64] M.C. Alkin, J.A. King, Definitions of evaluation use and misuse, evaluation influence, and factors affecting use, *Am. J. Eval.* 38 (3) (2017) 434–450, <https://doi.org/10.1177/1098214017717015>.
- [65] L.M. Shulha, J.B. Cousins, *Evaluation use: theory, research, and practice since 1986*, *Eval. Pract.* 18 (3) (1997) 195–208, [https://doi.org/10.1016/S0886-1633\(97\)90027-1](https://doi.org/10.1016/S0886-1633(97)90027-1).
- [66] V. Bush, *Science, the Endless Frontier: A Report to the President on a program for postwar scientific research*, United States Government Printing Office, Washington, 1945.
- [67] R.K. Merton, *The Sociology of Science: Theoretical and Empirical Investigations*, University of Chicago Press, Chicago, 1973.
- [68] A. Geuna, B.R. Martin, *University Research Evaluation and Funding: An International Comparison*, *Minerva* 41 (4) (2003) 277–304.
- [69] H.F. Moed, *Citation analysis in research evaluation. Information science and knowledge management*, Springer, Dordrecht, 2005.
- [70] T. van Leeuwen, *Descriptive Versus Evaluative Bibliometrics*, in: H.F. Moed, W. Glänzel, U. Schmoch (Eds.), *Handbook of quantitative science and technology research: the use of publications and patent statistics in studies of S&T systems*, Kluwer Academic Publishers, Dordrecht, 2004, pp. 373–388.

- [71] P. Mårtensson, U. Fors, S.-B. Wallin, U. Zander, G.H. Nilsson, Evaluating research: A multidisciplinary approach to assessing research practice and quality, *Res. Policy* 45 (3) (2016) 593–603.
- [72] B. Lepori, P. van den Besselaar, M. Dinges, B. van der Meulen, B. Poji, E. Reale, S. Silbersæter, Indicators for comparative analysis of public project funding: concepts, implementation and evaluation, *Res. Eval.* 16 (4) (2007) 243–255.
- [73] W. Glänzel, A. Schubert, Analyzing Scientific Networks Through Co-authorship, in: H.F. Moed, W. Glänzel, U. Schmoch (Eds.), *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems*, Kluwer Academic Publishers, Dordrecht, 2004, pp. 257–276.
- [74] D. Leite, L. Pinho, Evaluating Collaboration Networks in Higher Education Research, Springer International Publishing, Cham, 2017.
- [75] P. Knoepfel, C. Larrue, M. Hill, F. Varone, *Public Policy Analysis*, Policy Press, 2011.
- [76] F. Kern, P. Kivimaa, M. Martiskainen, Policy packaging or policy patching? The development of complex energy efficiency policy mixes, *Energy Res. Social Sci.* 23 (2017) 11–25, <https://doi.org/10.1016/j.erss.2016.11.002>.
- [77] K. Rogge, K. Reichardt, Policy mixes for sustainability transitions: An extended concept and framework for analysis, *Res. Policy* 45 (8) (2016) 1620–1635.
- [78] A. Hoogerwerf, Reconstructing policy theory, *Eval. Prog. Plann.* 13 (3) (1990) 285–291, [https://doi.org/10.1016/0149-7189\(90\)90059-6](https://doi.org/10.1016/0149-7189(90)90059-6).
- [79] J.L. Pressman, A. Wildavsky, Implementation: How Great Expectations in Washington Are Dashed in Oakland; Or, Why It's Amazing that Federal Programs Work at All, This Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation, University of California Press, 1984.
- [80] K. Gillingham, D. Rapson, G. Wagner, The rebound effect and energy efficiency policy, *Rev. Environ. Econ. Policy* 10 (1) (2015) 68–88.
- [81] G.D. Skogstad (Ed.), *Policy Paradigms, Transnationalism, and Domestic Politics*, University of Toronto Press, Toronto, 2011.
- [82] E. Vine, Strategies and policies for improving energy efficiency programs: Closing the loop between evaluation and implementation, *Energy Policy* 36 (10) (2008) 3872–3881.
- [83] I. Bache, M. Flinders, Multi-level governance and the study of the British State, *Public Policy Admin.* 19 (1) (2004) 31–51, <https://doi.org/10.1177/095207670401900103>.
- [84] C. Blumstein, S. Goldstone, L. Lutzenhiser, A theory-based approach to market transformation, *Energy Policy* 28 (2000) 137–144.
- [85] H. Geller, S. Nadel, Market transformation strategies to promote end-use efficiency, *Ann. Rev. Energy Environ.* 19 (1994) 301–346.
- [86] e. Vine, P. de Pont, P. Waide, Evaluating the impact of appliance labelling programs and standards: process, impact and market transformation evaluations, *Energy*, 26 (11) (2001) 1041–1059.
- [87] M. Keller, B. Halkier, T. Wilksa, Policy and governance for sustainable consumption at the crossroads of theories and concepts, *Environ. Policy Governance* 26 (2) (2016) 75–88, <https://doi.org/10.1002/eet.1702>.
- [88] M. Moezzi, K.B. Janda, From “if only” to “social potential” in schemes to reduce building energy use, *Energy Res. Social Sci.* 1 (2014) 30–40, <https://doi.org/10.1016/j.erss.2014.03.014>.
- [89] E. Shove, On the Difference between Chalk and Cheese—A Response to Whitmarsh et al.’s Comments on “beyond the ABC: Climate Change Policy and Theories of Social Change”, *Environ. Plann. A*, Econ. Space 43 (2) (2011) 262–264, <https://doi.org/10.1068/a43484>.
- [90] L. Whitmarsh, S. O’Neill, I. Lorenzoni, Commentary. Climate change or social change? Debate within, amongst, and beyond disciplines”, *Environ. Plann. A* 43 (2010) 258–261.
- [91] L. Schweber, R. Leiring, Beyond the technical: A snapshot of energy and buildings research, *Build. Res. Inf.* 40 (4) (2012) 481–492, <https://doi.org/10.1080/09613218.2012.675713>.
- [92] E. Shove, Beyond the ABC: Climate change policy and theories of social change, *Environ. Plann. A* 42 (6) (2010) 1273–1285, <https://doi.org/10.1068/a42282>.
- [93] T. Hargreaves, Practicing behaviour change: Applying social practice theory to proenvironmental behaviour change, *J. Consumer Culture* 11 (1) (2011) 79–99, <https://doi.org/10.1177/1469540510390500>.
- [94] R. Gifford, *Environmental Psychology*, 5th ed., Optimal Books, Colville, U.S., 2014.
- [95] R. Lenoir-Improta, P. Devine-Wright, J.Q. Pinheiro, P. Schweizer-Ries, Energy Issues: Psychological Aspects, in: G. Fleury-Bahi, E. Pol, O. Navarro (Eds.), *Handbook of Environmental Psychology and Quality of Life Research*, Springer, Switzerland, 2017.
- [96] D. Kahneman, Maps of bounded rationality: Psychology for behavioral economics, *Am. Econ. Rev.* 93 (2003) 1449–1475.
- [97] C. Camerer, Behavioral economics: Reunifying psychology and economics, *Proc. Natl. Acad. Sci.* 96 (1999) 10575–10577, <https://doi.org/10.1073/pnas.96.19.10575>.
- [98] R. Thaler, *Misbehaving: The Making of Behavioral Economics*, Norton, Grand Haven, MI, 2015.
- [99] T. Chatterton, C. Wilson, Characterising domestic energy related behaviours, Excess summer study proceedings (2013) 289–300.
- [100] P.M. Bögel, P. Upham, Role of psychology in sociotechnical transitions studies: Review in relation to consumption and technology acceptance, *Environ. Innov. Societal Trans.* 28 (2018) 122–136, <https://doi.org/10.1016/j.eist.2018.01.002>.
- [101] S. Jasanoff, S.-H. Kim, Containing the atom: sociotechnical imaginaries and nuclear power in the United States and South Korea, *Minerva: A Rev. Sci. Learn. Policy* 47 (2) (2009) 119–146, <https://doi.org/10.1007/s11024-009-9124-4>.
- [102] S. Sandin, L. Neij, P. Mickwitz, Transition governance for energy efficiency - insights from a systematic review of Swedish policy evaluation practices, *Energy, Sustainability Soc.* 9 (1) (2019), <https://doi.org/10.1186/s13705-019-0203-6>.
- [103] S. Sandin, M. Benner, unpublished results. Research evaluations for an energy transition? Insights from a review of Swedish research evaluation reports, Under peer review, 2020.
- [104] P. Stern, J. Enberg, M. Jondell Assbring, A.-K. Swenning, M. Terrell, Metautvärdering av utvärderingar genomförda på Energimyndighetens uppdrag 2000–2012 [Meta-evaluation of evaluations for the Swedish Energy Agency 2000–2012], Faugert & Co Utvärdering AB, 2013.
- [105] T. Jansson, K. Henningsson, M. Terrell, Halvtidsutvärdering av Samverksprogrammet Forskning och innovation för energieffektivt byggande och boende (E2B2) [Evaluation of the research programme E2B2, in the field of energy-efficient building and living (E2B2)], Faugert & Co Utvärdering AB, 2016.
- [106] Sweco, Styrmedels förutsättningar att styra mot ökad energieffektivisering - en utvärdering av 24 styrmedel - En rapport till Näringsdepartementet. 2.18 - Beställargrupper formade för att driva utvecklingen [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.18 - Cooperative network programs for driving development], Sweco Energuide AB, 2014.
- [107] M.-C. Dubois, K. Johnsen, Evaluation of the Swedish Energy Agency R&D - programme ENERGY EFFICIENCY OF LIGHTING - Innovation, research, development, demonstration and market introduction 2008-01-01 - 2011-12-31. Swedish Energy Agency, 2011.
- [108] P. Stern, E. Arenman, M. Terrell, Utvärdering av EFFSYS+, FoU-program för Resurseffektiva Kyl- och Värmepumpssystem [Evaluation of EFFSYS+, R&D programme for efficient cooling and heat pump systems], Faugert & Co Utvärdering AB, (2014).
- [109] T. Jansson, J. Enberg, S. Stålfors, M. Terrell, S. Karlström, Utvärdering av HyLok, statliga Hyresgäster för energieffektivisering av Lokaler [Evaluation of HyLok, state tenants for energy efficiency in premises], Faugert & Co Utvärdering AB, 2014.
- [110] Grontmij, Utvärdering av Energimyndighetens program för passivhus och Lågenergihus 2005–2008 [Evaluation of the Swedish Energy Agency’s program for passive houses and low-energy houses 2005–2008], Swedish Energy Agency, 2009.
- [111] Sweco, Styrmedels förutsättningar att styra mot ökad energieffektivisering - en utvärdering av 24 styrmedel - En rapport till Näringsdepartementet. 2.20 - Projektstöd för marknadsintroduktion och teknikupphandling [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.20 - Project support for market introduction and technology procurement], Sweco Energuide AB, 2014.
- [112] National Board of Housing, Building and Planning, Skärpta energihushållningskrav - redovisning av regeringens uppdrag att se över och skärpa energireglerna i Boverkets byggregler, rapport 2014:19 [More stringent energy conservation requirements - the accounting of the government’s instructions of reviewing and tightening energy regulations in building codes, report 2014:19], National Board of Housing, Building and Planning, 2014.
- [113] T. Jansson A.-K. Swenning, M. Lindström E. Berglund Utvärdering av programmet Energi, IT och Design [Evaluation of the programme Energi IT and Design], Faugert & Co Utvärdering AB, 2017.
- [114] F. Ekander, M. Olin, R.-M. Ågren, Slutrapport: effekter av energi- och klimatrådgivningen 2013 till privatpersoner, företag och organisationer [Final report: the effects of the energy and climate advisory in 2013 to private individuals, companies and organizations], Swedish Energy Agency, (2015).
- [115] Sweco, Styrmedels förutsättningar att styra mot ökad energieffektivisering - en utvärdering av 24 styrmedel - En rapport till Näringsdepartementet. 2.8 - Energi- och Klimatrådgivning [The ability of policy instruments to steer towards increased energy efficiency - an evaluation of 24 policy instruments - A report to the Ministry of Enterprise and Innovation. 2.8 - Energy and climate advisory program], Sweco Energuide AB, 2014.
- [116] Swedish Energy Agency, Långtidsuppföljning av energi- och klimatrådgivningen 2008 och 2009, ER 2011:20 [Long-term follow-up of the energy and climate advisory in 2008 and 2009, ER 2011:20], Swedish Energy Agency, 2011.
- [117] Swedish Energy Agency, Effekter av energi- och klimatrådgivningen 2011 [Effects of the energy and climate advisory in 2011], Swedish Energy Agency, 2013.
- [118] Swedish Energy Agency, Effekter av energi- och klimatrådgivningen 2012 till privatpersoner, företag och organisationer [Effects of the energy and climate advisory in 2012 to private individuals, companies and organizations], Swedish Energy Agency, 2013.
- [119] Profu, Utvärdering av CERBOF-programmet [Evaluation of the CERBOF programme], Profu, 2011.
- [120] S. Faugert, E. Arnold, T. Jansson, H. Segerpalm, N. Brown, Utvärdering av samverkan Formas-Bic [Evaluation of collaboration Formas-Bic], Faugert & Co Utvärdering AB, 2006.
- [121] T. Jansson, A.-K. Swenning, M. Terrell, Utvärdering av Spara och bevara [Evaluation of a research programme for energy efficiency in cultural heritage building], Faugert & Co Utvärdering AB, (2014).
- [122] P. Stern, T. Fridholm, A. Håkansson, M. Terrell, Utvärdering av Forskningsprogrammet Uthållig kommun [Evaluation of the research programme Forskningsprogrammet Uthållig kommun], Faugert & Co Utvärdering AB, (2013).
- [123] National Board, of Housing, Building and Planning., Utvärdering av stödet för installation av energieffektiva fönster eller biobränsleändringar [Evaluation of

- the financial support for installation of energy efficient windows or biofuel devices], National Board of Housing, Building and Planning, 2009.
- [124] T. Jansson, M. Terrell, Halvtidsutvärdering av HyLok, statliga Hyresgäster för energieffektivisering av Lokaler [Mid-term evaluation of HyLok, state tenants for energy efficiency in premises]. Faugert & Co Utvärdering AB, 2012.
- [125] T. Jansson, M. Grudin, K. Henningson, Slututvärdering av Programmet för byggnader med mycket låg energianvändning (Lågan) [Final evaluation of the program for buildings with very low energy use (Lågan)]. Faugert & Co Utvärdering AB, 2016.
- [126] T. Jansson, M. Terrell, Halvtidsutvärdering av program för byggnader med mycket låg energianvändning – LÅGAN [Mid-term evaluation of the program for buildings with very low energy use – LÅGAN]. Faugert & Co Utvärdering AB, 2013.
- [127] Swedish Energy Agency, Energikartläggningscheckar – En samhällsekonomisk utvärdering [Financial support for energy audits - A socio-economic evaluation]. Swedish Energy Agency, 2013.
- [128] National Board of Housing, Building and Planning, Utformningen reducerade effekterna – Boverkets utvärdering av Offrotstödet. Redovisning av stöd till investeringar i energieffektivisering och konvertering till förnybara energikällor i lokaler som används för offentlig verksamhet 2005:205, enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A [The design reduced the impact - the Board's evaluation of the Offrot financial support. Reporting of support for investment in energy efficiency and conversion to renewable energy sources in premises used for public activities 2005:205, in accordance with appropriation directions for the financial year 2009 concerning the National Board of Housing, Building and Planning. M2008/4791/A]. National Board of Housing, Building and Planning, 2009.
- [129] T. Dietz, P.C. Stern, E.U. Weber, Reducing carbon-based energy consumption through changes in household behavior, *Daedalus* 142 (1) (2013) 78–89, https://doi.org/10.11162/DAED_a.00186.
- [130] W. Glad, Housing renovation and energy systems: the need for social learning, *Build. Res. Inf.* 40 (3) (2012) 274–289, <https://doi.org/10.1080/09613218.2012.690955>.
- [131] L. Vlasova, K. Gram-Hanssen, Incorporating inhabitants everyday practices into domestic retrofits, *Build. Res. Inf.* 42 (4) (2014) 512–524, <https://doi.org/10.1080/09613218.2014.907682>.
- [132] J. Palm, K. Reindl, Understanding energy efficiency in Swedish residential building renovation: A practice theory approach, *Energy Res. Social Sci.* 11 (2016) 247–255, <https://doi.org/10.1016/j.erss.2015.11.006>.
- [133] T.M. Skjølsvold, M. Ryghaug, Embedding smart energy technology in built environments: A comparative study of four smart grid demonstration projects, *Indoor Built Environ.* 24 (7) (2015) 878–890, <https://doi.org/10.1177/1420326X15596210>.
- [134] A. Nilsson, C.J. Bergstad, L. Thuvander, D. Andersson, K. Andersson, P. Meiling, Effects of continuous feedback on households' electricity consumption: Potentials and barriers, *Appl. Energy* 122 (2014) 17–23, <https://doi.org/10.1016/j.apenergy.2014.01.060>.
- [135] A. Nilsson, K. Andersson, C.J. Bergstad, Energy behaviors at the office: An intervention study on the use of equipment, *Appl. Energy* 146 (2015) 434–441, <https://doi.org/10.1016/j.apenergy.2015.02.045>.
- [136] M. Bergquist, A. Nilsson, I saw the sign: Promoting energy conservation via normative prompts, *J. Environ. Psychol.* 46 (2016) 23–31, <https://doi.org/10.1016/j.jenvp.2016.03.005>.
- [137] P. Maleetipwan-Mattsson, T. Laike, M. Johansson, Factors affecting optimal lighting use in shared hospital environments: A case-study, *Build. Environ.* 96 (2016) 260–269, <https://doi.org/10.1016/j.buildenv.2015.11.026>.
- [138] P. Maleetipwan-Mattsson, T. Laike, M. Johansson, The effects of user interface designs on lighting use, *J. Eng. Des. Technol.* 15 (1) (2017) 58–78, <https://doi.org/10.1108/JEDT-06-2015-0040>.
- [139] K. Ek, P. Söderholm Patrik, The devil is in the details: Household electricity saving behavior and the role of information, *Energy Policy* 38 (3) (2010) 1578–1587, <https://doi.org/10.1016/j.enpol.2009.11.041>.
- [140] K. Gram-Hanssen, Residential heat comfort practices: Understanding users, *Build. Res. Inf.* 38 (2) (2010) 175–186, <https://doi.org/10.1080/09613210903541527>.
- [141] J. Palm, The public-private divide in household behavior: How far into home can energy guidance reach? *Energy Policy* 38 (6) (2010) 2858–2864, <https://doi.org/10.1016/j.enpol.2010.01.018>.
- [142] P. Hagbert, P. Femenas, Sustainable homes, or simply energy-efficient buildings? *J. Hous. Built Environ.* 31 (1) (2016) 1–17, <https://doi.org/10.1007/s10901-015-9440-y>.
- [143] A. Nilsson, D. Lazarevic, N. Brandt, O. Kordas, Household responsiveness to residential demand response strategies: Results and policy implications from a Swedish field study, *Energy Policy* 122 (2018) 273–286, <https://doi.org/10.1016/j.enpol.2018.07.044>.
- [144] B. Turnheim, F. Berkhout, F. Geels, A. Hof, A. McMeekin, B. Nykvist, D. van Vuuren, Evaluating sustainability transitions pathways: Bridging analytical approaches to address governance challenges, *Global Environ. Change* 35 (2015) 239–253.

Paper IV



Article

Making Use of Evaluations to Support a Transition towards a More Sustainable Energy System and Society—An Assessment of Current and Potential Use among Swedish State Agencies

Sofie Sandin

The International Institute for Industrial Environmental Economics (IIIEE), Lund University, P.O. Box 196, 22100 Lund, Sweden; sofie.sandin@iiiee.lu.se

Received: 20 August 2020; Accepted: 3 October 2020; Published: 7 October 2020



Abstract: Evaluations hold the potential to support decision-making so that current global challenges related to climate and energy can be addressed; however, as the challenges are becoming increasingly large and complex, new and transformative evaluation approaches are called for. Such transformative evaluation in turn builds on an extended and more deliberate use of evaluations. This study focuses on the current evaluation use practices among Swedish state agencies who are commissioning and/or conducting evaluations within climate and energy-related areas. Building on focus group sessions with four agencies and a structured interview questionnaire answered by representatives at five state agencies, the results shed light on how informants perceive the current practices of using evaluations, following the models of use presented in the evaluation literature. These results show perceived use as mainly instrumental or conceptual, along with showing an overall emphasis on models of use that are deemed constructive for moving towards transformative evaluations. The results also outline key benefits and challenges related to the adoption of a transformative evaluation approach. Such benefits include a more structured planning and use of evaluations, while challenges relate to institutional barriers and mandates to coordinate evaluations on a transformative scale.

Keywords: evaluation use; models of use; transition; climate; energy

1. Introduction

The projections and trajectories of the current development process show that advancements towards addressing global climate challenges are too slow [1–3]. Thus, calls are becoming urgent for changes at larger scales [3] and for sustainability transitions that shift entire socio-technical configurations [4] towards a society that is built on more sustainable practices [5–7]. In order to achieve such sustainability transitions, deliberate and well-designed research initiatives and policy instruments are needed [1,2,8].

This paper focuses on facilitating a transition towards a more sustainable energy system and society. To ensure that time and effort are indeed invested in policy instruments that will support such a transition, there is a need to equip decision-makers with reliable and informed knowledge about how different instruments perform. Evaluations of both research efforts and policy instruments are abundant, as ex-post evaluation of past experiences as well as prospective ex-ante assessments are integrated into the seventh EU Environment Action Program in the EU's Better Regulation Agenda [9], and previous studies have shown that multiple evaluations are undertaken in the fields of climate and energy policy, as well as other sectors [10–14].

Thus, it is key to ensure that evaluations of research and policy instruments are used properly. The use of evaluations is a frequently discussed and studied area among evaluation scholars [15–21].

Although the theory is rich in describing reasons for impeded use and how to address these [16,18,21–24], there are still concerns among evaluation scholars that evaluation findings are misused or are not used to their full potential [19,25,26]. The purposes for evaluating are closely related to the different types of uses [21], which have been conceptualized as models of use. These models represent a vast range of uses that take place at various times of an implementation: spanning, for example, instrumental improvement, tactical or political use, enlightenment and process-related learning [27–30].

Previous studies on evaluation use focus on the drivers and requirements for an enhanced use of evaluation findings [10,31–33], but attempts have also been made at assessing which models of use are prevalent in practice [17]. The results of these studies show that many models of (perceived) use can exist in parallel [10,17], and that a high-quality and clear evaluation is a prerequisite for usefulness [10], while not being a guarantee for actual use [17,31]. Different models of use are noted in these studies, including instrumental, conceptual, legitimizing and ritual models [17,31], but it is also noted that the context in which an evaluation is performed matters for its use [10,33].

However, evaluation use studies tend to focus on the use of individual policy evaluations and not on the use of evaluations to acquire systemic insights about evaluands' support of systemic change and transition. Likewise, evaluations often omit a deeper analysis of the interplay between instruments [34], and as a consequence, the usefulness of evaluations of individual policies has been questioned as the effects from one single instrument may be difficult to discern from the impacts of the policy instrument mix it is embedded into [35]. Thus, a more holistic approach to evaluating—and using evaluations—has been called for [34,36]. Neij et al. [37] propose that such a transformative evaluation approach needs to build on insights drawn from various evaluations to deliberately synthesize a broad, holistic, transformative perspective of how different research initiatives and policy instruments are building momentum towards a sustainability transition together. A framework that accounts for the key aspects to consider when performing such transformative evaluations includes drivers for a transition, such as visioning, experimentation and learning, and accounts for a systems and multi-actor perspective, as well as considering reflexivity and the construction of value judgements [37,38]. How such evaluations can actually be facilitated and integrated into current evaluation commissioning and use practices is, however, largely unknown, particularly in the context of facilitating a transition towards a more efficient energy system and society.

The objective of this study is thus to provide an exploratory assessment of how evaluations of research and policy instruments aimed at climate and energy-related issues are used currently, how enhanced use is supported today, or can be supported, and which benefits and challenges that stakeholders perceive as resulting from the adoption of a transformative evaluation approach. For this, the case of Sweden is used, focusing on the readiness among Swedish state agencies to adopt transformative evaluation and integrate it into their current evaluation practices, and how the current reported modes of using evaluations fare in relation to the outlined need for supporting a transition.

The study is conducted by engaging four Swedish state agencies in focus groups and is complemented by a structured interview questionnaire sent out to seven Swedish state agencies. The selection of agencies is based on their working with the commissioning and/or evaluation of instruments related to climate and energy. The theoretical framework that guides the study and analysis builds on evaluation theory, particularly concerning different models of evaluation use proposed by evaluation scholars, in order to categorize different kinds of use and means to promote enhanced use.

This paper is organized as follows: the case used for the study—i.e., the proposed transformative evaluation approach—is presented in Section 1.1. This is followed by a description of the theoretical framework of the study, which is founded in evaluation theory, mainly focusing on the use of evaluations and different models of use, in Section 2. Section 3 presents the material and methods, comprising of focus group sessions with four Swedish state agencies and a structured interview questionnaire answered by representatives from five Swedish agencies. Results from the focus group sessions and questionnaire are presented in Section 4, followed by a discussion in Section 5 and some concluding remarks in Section 6.

1.1. Case Description—Transformative Evaluation

In order to address the need for a more holistic approach to using evaluations to generate insights about contributions on a transformative scale, a broader transformative evaluation approach was proposed by Neij et al. [37]. In general, there is no generic method to apply in an evaluation to secure its reliability or utility. Instead, different methods are called for to address different aspects and purposes and should thus be selected in order to appropriately address the needs that the evaluation is intended to meet [24]. Next, syntheses of evaluations, or meta-analyses, are promoted by evaluation scholars in order to provide broader evaluation results that span beyond the individual evaluand. Such syntheses are not promoted only for the sake of broadening the application of findings but also to increase the credibility of the findings to users [27,39]. Then, to address environmental challenges in particular, there is not only a need for strong knowledge production and evaluation practices but also for collaboration between knowledge producers and policy makers [40]. Together, these insights support the adoption of a transformative evaluation approach that builds on a variety of methods, that promotes the deliberate synthesizing of findings to support the credibility and utility of the findings among users and that has an articulated purpose of supporting a transition towards a more sustainable energy system and society.

The transformative evaluation approach was developed in an interdisciplinary research project between 2016 and 2019, intended to describe how evaluations of different individual research and policy programs and initiatives can be thoughtfully planned and conducted to become part of a more holistic perspective of transformative change. The proposed approach for transformative evaluation, hereafter called “transformative evaluation”, targets both research initiatives and policy instruments. It promotes making individual evaluations build on key insights from different but relevant disciplines, taking different areas of focus that are of different scales. Together, when properly coordinated and synthesized, these evaluations can provide an overview of how different instruments support change at a larger scale [37].

A cornerstone of the transformative evaluation approach is that it should be an extension of current evaluation practices, which gears evaluations towards a more deliberate integration of relevant viewpoints and concepts from evaluation theory and transition research, for example. Another cornerstone is that the approach should build on an extended use and synthesis of evaluation findings that go beyond the evaluand itself. Thus, the applicability of the proposed approach depends on current practices for conducting and using evaluations. In order to shed light on such current practices, as well as the overall attitudes towards transformative evaluation, this study focuses on Swedish state agencies that work with commissioning and/or conducting evaluations within areas related to climate and energy to assess their readiness to adopt a transformative evaluation approach.

2. Theoretical Framework

The point of departure for this study is that, to succeed in piecing together a more holistic picture of how research and policy instruments support a transition towards a more sustainable energy system and society, the use of their evaluations needs to be facilitated and extended. This requires that evaluations are conducted in a way that allows them to be accepted as carriers of information that can support a transition, and it requires an extended and clear strategy for securing the actual and deliberate use of evaluations beyond the program level—a key part of transformative evaluation (see Section 1.1). As such, this contrasts the more traditional perception of the use of evaluations as improving a program’s effectiveness and implementation [18]. Instead, evaluations are proposed to make a broader promise to users: to be applicable beyond the evaluated program in terms of providing pieces of a more holistic perspective. In order to address these issues, this study draws upon the evaluation literature, mainly surrounding the use and utility of evaluations to understand why use is impeded and how it can be supported, and surrounding the conceptualization of different models of use.

2.1. Use of Evaluations

The use patterns and identification of intended users depend on the purposes of the evaluation [18,27], which can be diverse, multiple, interrelated, parallel or even simultaneous. In general, two main domains of use are outlined—process use and findings use—with the former referring to use or actions taken prior to the finalization of the evaluation report and the latter concerning the use of actual evaluation findings [41].

The multitude of works within evaluation theory that refer to use bears witness to the fact that the actual use of evaluations may be substandard [17,18,22,23,32]. In order to support a broader and more active use of evaluations—a cornerstone of transformative evaluation—it is essential to understand why use is impeded. Evaluation theory scholars have identified various causes for this, including political resistance, where findings do not align with actors' agendas [22,27]; scepticism or cynicism towards findings as users, for example, do not see the value or benefits of an evaluation; fear of judgement; a lack of resources to act on the evaluation; organizational cultures of non-use [18]; and other organizational tensions such as a lack of coordination, uneven power balances and competing interests between different actors [42]. Other barriers may stem from the friction between institutions and systems, which may occur when an evaluation topic crosses over different areas, subjects and agencies. Other systemic barriers include how the acceptability of evaluation findings is affected due to user preferences in terms of the methodology and modes of communication and knowledge when spanning different systems and actors [22]. Thus, the use of evaluation findings hinges on the users, which is a heterogeneous group with different interests and agendas [18,22,29]. Circumstances of non-use can, however, be deemed appropriate, for example, when an evaluation is not delivered on time or is poorly executed [18,26].

What are then the countermeasures for non-use? Two main areas for addressing non-use concern (I) the intended use of the evaluation and (II) the conduct and processes of the evaluation. Starting with the former, scholars have emphasized the need for a clear and realistic expectation of the intended use [18,27,28]. Next, it is paramount that users are defined and are involved in the evaluation process to support the intended use. Thus, receivers of the evaluation need to be integrated from the start in the determination of the questions of inquiry to make the evaluation relevant to them and to increase their sense of ownership and investment in the results [18,28,29,32]. The evaluation process should also be shaped according to input from users and key stakeholders, and should cater for the needs of users by being timely and performed at different stages of an implementation [23,28,29].

To connect this with transformative evaluation, the purpose and use of an evaluation may be twofold: in part, the purpose of evaluation is to gain knowledge that can improve or guide the implementation, but importantly, the purpose should also be to assist in the assessment of whether and how the implementation is supporting a transition. This calls for a framing of transition in terms of the purpose of the evaluation but also a consideration of who the users of such knowledge are; this may be both decision-makers on a state level and on program levels.

The second key area for addressing non-use focuses on the actual conduct of the evaluation and the reports that are delivered to users. Apart from the prerequisites of a good evaluation, such as being delivered on time, clarity and accessibility to users [10,22,28], evaluations also need to be performed using sound and robust methodologies [22,27], and they should seek to consider the political context and the decision-making context in which the evaluation is intended to be used [10,16,24,27,33,43]. Relatedly, an evaluation also gains credibility by being performed by a broad but deliberate evaluation team, whose methods and competencies gain trust among users, and through clearly communicated methods and competencies [22,23].

Lastly, a key for increasing the use of an evaluation is to carefully consider how it is disseminated. Audiences need to be aware of the existence of an evaluation report. This can be done by securing a dialogue between the evaluator and users, which should be maintained over the duration of the evaluation process, including the use phase [18,22,27–29].

2.2. Models of Use

As has been outlined in the previous section, there are numerous ways to use evaluations and numerous approaches to increase their use and utility. Table 1 presents an overview of 11 different models of use of evaluations—a conceptualization that builds on the work of scholars in the evaluation theory field. While this list may not be exhaustive, it provides an overview of the main approaches to the use of evaluations, which is essential to understand how the purpose of an evaluation is translated into some kind of desired action or effect, or how evaluations are received and acted upon by users.

Table 1. A list of 11 models of use for evaluations. The names of the models are adopted from evaluation scholars, as indicated in the table; the characteristics briefly describe the essence of the model and build on the work of evaluation scholars.

Model of use	Characteristics
1. Instrumental [27–30]	Evaluation findings are applied immediately for specific actions concerning the evaluand.
2. Enlightenment/conceptual [27–30]	Evaluation findings are not used for the evaluand per se, but rather to provide insights about issues or implementations in general.
3. Legitimizing/reinforcing use [27,29,30]	Evaluation findings are used to legitimize and justify decisions already made; e.g., to confirm current knowledge and beliefs, or support confidence in a standpoint.
4. Interactive [20,29]	Evaluation findings form part of a larger decision-making process, where other sources of information and actors influence the decisions.
5. Ritual/symbolic use/mechanical use [18,30]	Evaluations are conducted symbolically, because this is expected by current customs and practice, or evaluations are performed mechanically; evaluands seek only to fulfil requirements and get a good “score”.
6. Mobilizing use/persuasive use [27,28]	Evaluation findings are used to create support for a particular standpoint, for mobilizing actors.
7. Overuse [18]	Evaluation findings are put to use as definitive facts, without considering the contextual factors, or without exploring why certain outcomes have emerged.
8. Process use [18,30]	The evaluation process in itself contributes with insights and learning and spurs action.
9. Constitutive/anticipatory use [30]	An evaluation process spurs action and has impacts already in its initial phases by making evaluands and stakeholders aware of its coming into force.
10. Tactical use [29,30]	Evaluations are conducted in order to postpone decisions, by showcasing that there is an ongoing evaluation. It is the process rather than the findings that are at the focal point of use.
11. Unintended use [18]	Evaluation findings are used outside of an evaluated program; e.g., by spawning other investigations concerning issues separate from the program.

3. Materials and Methods

The data collection for this study was conducted following two steps: the arranging of focus groups with representatives from selected state agencies, and a structured interview questionnaire sent out to state agencies to be filled in anonymously. As such, the focus groups allowed an initial scoping of general perceptions among agencies about taking on a transformative evaluation approach (as described in Section 1.1), while the structured and anonymous interview questionnaire allowed a deeper probing about current evaluation use practices and the possibilities and challenges that may come with moving current evaluation practices towards transformative evaluation.

3.1. Focus Groups

Using a focus group as a method for data collection allows an engaging and dynamic way of acquiring insights about participants’ opinions and experiences about a topic, which allows a comparison and discussion of various and conflicting viewpoints to be considered [44,45]. Apart from

allowing a more inclusive and diverse discussion among participants compared to individual interviews [45], focus groups were also deemed appropriate for this initial state of the research in that the interactive discussions about evaluation practices among colleagues of a department were creating interest within departments to join in.

Four focus groups were arranged with Swedish state agencies in April and May 2019. The agencies were selected based on their working with evaluation by commissioning and/or conducting evaluations of research initiatives and/or policy instruments within energy and climate-related areas, with overarching mandates and responsibilities to support a more sustainable energy system and society. Requests to visit were made to five agencies, of which four agreed to partake in a focus group session. These were the Swedish Energy Agency, the Swedish Environment Protection Agency, Vinnova and the Swedish Agency for Growth Policy Analysis (the Swedish Energy Agency (Energimyndigheten) is overseeing the work towards an energy transition for sustainability; the Swedish Environment Protection Agency (Naturvårdsverket) is responsible for climate-related issues and goals; Vinnova is responsible for innovation for sustainable growth; and the Swedish Agency for Growth Policy Analysis (Tillväxtanalys) analyses and evaluates political effects on growth). The attendees in the focus groups were invited by the agencies, based on the description that the session targeted employees that were working with tasks related to evaluation. The focus group sessions were carried out as workshops, held individually at each agency, which typically lasted for two hours; the number of participants ranged from three to approximately 15 (some participants were unable to join for the full duration of the workshop due to other duties, hence the approximate number of attendants). When the groups were large, they were split into sub-groups of up to eight participants. The aim of these workshops was to present the proposed evaluation approach and to gain an understanding of how different agencies perceived the possibilities and challenges of applying such an evaluation approach.

The format of the workshops was the same for each: the researchers started by presenting the proposed evaluation approach, then discussions followed related to the application of such an evaluation approach in their respective departments. At the last phase of the workshops, researchers particularly asked for the participants' perceptions of the benefits and challenges that they foresaw when applying such an evaluation approach in their departments. Notes were taken by two researchers, which were compiled and organized into sub-categories according to content.

3.2. Structured Interview Questionnaire

In order to address the limitations of focus groups, particularly in relation to participants' adjustments to any perceived expectations of views at the work place or any self-censorship due to social relationships [44], the focus group sessions were complemented with a structured interview questionnaire that was sent out to seven state agencies, four of which were represented in the focus groups. The interview questionnaire was filled in online individually by informants. The purpose of the structured interview questionnaire was thus twofold: to triangulate (i.e. crosscheck results using various methods [46]) and validate the initial scoping of perceptions among agencies towards transformative evaluation, and to acquire deeper insights about the current evaluation practices within state agencies, based on individual and anonymous responses uninfluenced by peer expectations

Sending out the questionnaire online was deemed a convenient means to reach the informants on their terms, giving the flexibility to fill in the questionnaire whenever it suited the informants. Moreover, the design of the interview guide contained multiple blocks of similar questions and multiple statements to rate, and as such it was deemed to perform better in a self-administrated format [44,47]. The structured interview questionnaire was primarily founded in a theoretical framework (as presented in Section 2) but also built on the categories of benefits and challenges that were drawn from the initial scoping of perceptions from the focus groups sessions. The questionnaire was originally in Swedish (see Appendix A for an English translation).

The selection of state agencies to receive the interview questionnaire was, like the focus groups, delimited to informants at state agencies working with evaluation either by commissioning

and/or conducting evaluations of research initiatives and/or policy instruments within energy and climate-related areas. Thus, the selection of potential informants was rather narrowly delimited from the beginning, which was deemed necessary to secure relevant responses. Requests to participate were made to seven representatives at seven different agencies via e-mail, who were asked to disseminate the request to relevant colleagues at relevant departments. The selection of informants at each agency was made exclusively in each department in order to secure anonymity among the informants as far as possible. The questionnaire was to be filled in between 18 October–15 November 2019, it received 11 responses in total from five different agencies: the Swedish Energy Agency, the Swedish Environment Protection Agency, Vinnova, the Swedish Agency for Growth Policy Analysis, and the Swedish Agency for Economic and Regional Growth (the Swedish Agency for Economic and Regional Growth (Tillväxtverket) supports regions and businesses in sustainable growth related to climate and other areas).

4. Results

The results from this exploratory study will be presented in two main categories: results that concern the current use of evaluations and which measures that are taken that may facilitate further use (Section 4.1), and the actual benefits and challenges raised by adopting and integrating a broader evaluation approach with a transformative focus at the Swedish agencies represented in this study (Section 4.2).

Given the exploratory nature of this study, and given that no inferences can be made regarding the reported answers, the results are presented as indications of the current evaluation practices among five Swedish agencies. While the results do provide valuable insights on evaluation use processes and practices, no claims are made to generalize beyond what the study at hand entails.

4.1. Design, Conduct and Use of Evaluations and Facilitating Measures for Use

4.1.1. Design and Conduct

Evaluation theory provides some key measures for facilitating and supporting the use of an evaluation (see Section 2). One key strategy is to ensure robust evaluation approaches and methods which build on relevant actors being involved who provide relevant knowledge for conducting the evaluation. In order to assess the Swedish evaluation practices in this regard, the questionnaire sent to the agencies asked informants which actors were commonly involved in the design of the evaluation; which actors conducted the evaluation; and which particular areas of knowledge were sought when commissioning actors to engage in the conduct of an evaluation.

The results concerning involvement in the design of an evaluation are illustrated in Figure 1 (the graph to the left) and show a clear emphasis on placing responsibility for the design of the evaluation on the staff at the agencies themselves: 10/11 reported that this was always the case. Expertise from external consultancies was reported to be used occasionally, with 6/11 reporting that this was sometimes or often the case. Actors from other state agencies or from academia were less frequently engaged in the designing phase: both categories received the lowest total score among the informants. Open-ended reflections from informants provided additional clarifications, with one mentioning that expertise from their own agency was often selected to represent different administrative units of the agency, including both specialist knowledge on the implementation and those with evaluation expertise. Another informant mentioned that organizations with an interest in the implementation also occasionally partook in the design of certain evaluations. Thus, it seems that while agencies who commission evaluations do take a large responsibility in their design, other actors may be invited to partake in the design process as well. The involvement of key stakeholders in an evaluation is outlined as being crucial to increase its credibility and utility to users [23,28,29], and while there seemingly is an emphasis among the informants on in-house participation in the design of the evaluations, this may, however, mirror the actual involvement of intended users, as will be discussed later. However,

consideration to actor involvement in the design of evaluations should be given in transformative evaluation, since users then may be found outside of the agency or immediate program boundary.

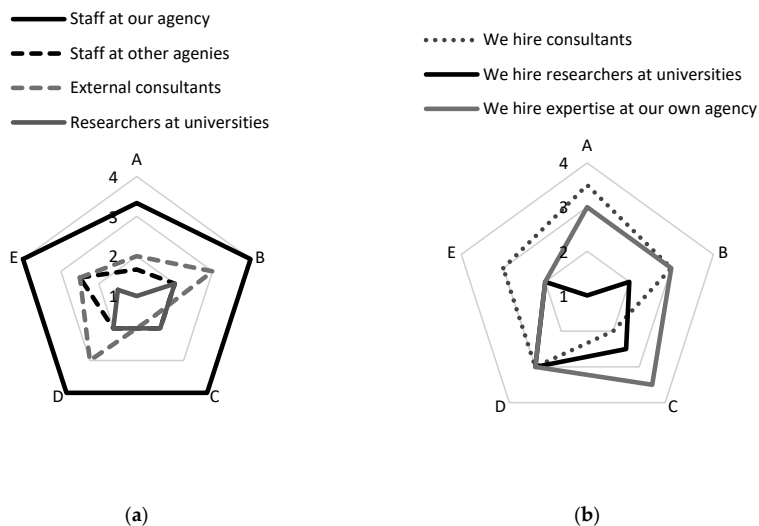


Figure 1. Responses from representatives at five different Swedish state agencies concerning actor involvement in the design and conduct of evaluations. (a) Answers from the questionnaire to the question "Who partakes in the design of the evaluations that you either commission or conduct at your agency?" (b) Answers from the questionnaire to the question "Which actors do you hire when you commission an evaluation?" Results show the mean value reported by informants from each agency (A–E), where 1 is "never" and 4 is "always".

Next, informants were asked about which actors were commissioned to perform evaluations (Figure 1, the graph to the right). In this regard, the informants reported the main tendencies to be both hiring consultants and conducting the evaluation in-house at the agency, while commissioning researchers from academia was less frequently done. As may be suspected, different evaluations call for different kinds of expertise, and one informant clarified that certain types of evaluations were mainly conducted at the agency (e.g., enquiries and official reports), whereas evaluations of implementations were commissioned to external consultants. The expertise most called for when commissioning an evaluation from external actors was reported to be evaluation expertise, which 10 informants (the 11th refrained from answering) assigned as always or often being the case, followed by expertise in economics, which was reported to be the case often or sometimes. The least sought-after expertise, as reported by informants, was modeling expertise. Two informants added that specialist knowledge for the topic at hand was also sought after, including an understanding of how innovation systems and policy work. On the topic of which kind of expertise is commissioned, there are differences to be noted between different agencies. For example, informants working at the same state agency consistently reported that certain kinds of expertise were seldom commissioned, while the same kinds of expertise were reported as sought-after by informants at other agencies. This can be explained by the fact that agencies have different core areas and consequently host different key competencies, or possibly that they (traditionally) approach the evaluation of implementations with certain predominant methods and kinds of expertise.

Putting these results in the context of evaluation theory, it is apparent that the contextual dimension is of importance when it comes to selecting the methods, key people and kinds of expertise to be involved. The implications for increasing use may thus be inherent either to the agency's general

practices for commissioning and conducting evaluations—e.g., related to certain traditions or key competencies that decree how an evaluation is designed and conducted—or to political or bureaucratic conditions, such as political interests or procurement/master agreements.

The next key aspect for increasing use, as outlined by evaluation scholars, is the communication of evaluation results. According to the evaluation literature, mechanisms for a successful dissemination and uptake of evaluation results include presenting results in formats and channels that are tailored for the specific target users and leveraging public channels that can increase the accessibility and awareness among broader audiences [24,27]. The questionnaire posed open-ended questions regarding who informants considered to be the most important users of the evaluations commissioned or conducted at the agency, and to whom evaluation results were communicated. The informants indicated a range of users, commonly including the agencies themselves, but also state departments such as the Swedish Government Offices and the Ministry of Enterprise, Energy and Communications, as well as actors at regional levels. Regarding the channels used for the dissemination of evaluation results, the informants again indicated a variety of approaches. Among the different approaches to dissemination suggested in the questionnaire (Appendix A), none received a clear answer of “never” or “always”. Channels that received a score that suggested a more frequent practice included the publishing of evaluation reports on the agency’s webpage, an active communication of results to those involved in the evaluation (e.g., interviewees) and the presentation of the evaluation at meetings where those affected by the evaluation were invited to take part physically or digitally. Approaches to communication that informants reported as being less frequently applied mainly included internal dissemination at the agency only or dissemination to local or regional agencies. Nevertheless, informants also elaborated on the communication of evaluation results, where some clarified that the means of dissemination and the selection of receivers depended on the purpose of the evaluation, again linking back to the context. Additional measures of communication that were raised by informants included making the evaluation results accessible on online web-portals related to the implementation, and using films published online. Relating back to evaluation theory, the reported efforts from the informants showcased attempts to promote evaluation results using different outlets, providing avenues for an increased access and use of the evaluation results.

4.1.2. Models of Use

Turning to the use of evaluations, the questionnaire sought to gain insights regarding the current perceptions of the intended use of evaluations. Informants were asked to rate statements that referred to the 11 different models of use outlined in Section 2.2, on a scale between 1–5, on how well they were perceived to apply to the informant’s agency, where 1 was “never” and 5 was “always” (see appendix A for a full list of statements as presented in the questionnaire). The results from the questionnaire are shown in Figure 2, where each model of use is presented as the aggregate ratings from all informants of the three statements underpinning the model (see Appendix B for a list of statements representing each model of use). In cases in which the statement was a reverse positive, the rating has been inverted. The number of responses per model of use could be up to 30 (10 informants answered these questions), with the response rate ranging from 27–30 per model of use. By using a Likert scale with a neutral midpoint, a division could be discerned between models of use that were rated as predominantly on the “never” side of the spectrum (1–3), as well as those predominantly found on the “always” side (3–5).

First, the results suggest that statements referring to all models of use were rated as high by some informants, and thus that the informants did not altogether denounce any model of use. Nevertheless, the results also suggest that some overall indications of certain models were more frequently recognized than others. The models of use that informants rated as least conforming with their perception of the evaluation practices at their agencies were mainly ritual and mobilizing use, followed by tactical use and overuse (Figure 2). Some of these models are deemed by scholars to be inappropriate or destructive [18,29]; and considering them from the perspective of supporting a

transition, they are not deemed to be very helpful in building a constructive and transparent knowledge base because they promote certain (covert) agendas (e.g., mobilizing or tactical use), which would require significant transparency if synthesized in a holistic perspective. This because they lack an active concern about evaluation results (e.g., ritual use) or because the findings are not properly put into context or perspective (e.g., overuse). Thus, while some informants gave these statements high ratings, the majority of ratings were found on the “never” side of the spectrum, and as such, they were deemed of less importance for current practices in terms of the use of evaluations to support a transition towards a more sustainable energy system and society.

The next category that could be discerned from the results included the models spanning the entire spectrum, balancing between high and low levels of agreement. These models included constitutive and legitimizing use. In order to further analyze the wide reported span of these models, the underpinning statements were assessed individually. For legitimizing use, the statement “give trust and support for decisions that concern a program” received a high total score (41) and thus largely represented the right side of the spectrum, whereas the other statements—“contribute to showing that we are doing things correctly” and “confirm what we know about a research program or policy instrument”—conversely received lower scores (27 and 22, respectively). For constitutive use, the statements were more evenly rated, with the lowest being “realize effects as early as possible, before the evaluation in itself is done” (23) and the highest being “spur improvements in a program by communicating that an evaluation is to be done to those whom the evaluation concerns” (36). Contrasting these models of use from the perspective of supporting a transition, they can be argued to be of more or less importance. Constitutive use—spurring action and willingness to perform well by informing evaluands that an evaluation is coming—can probably only generate minor effects on a transformative scale, at least compared to other, more targeted efforts for supporting a transition. Legitimizing use, on the other hand, can be powerful from a transformative perspective, both for supporting and for disfavoring such efforts. Vedung [29] argues that parties using legitimizing use to support their own (political) agenda cannot be faulted, and depending on the agenda for adopting legitimizing use, it can be both detrimental and constructive for a given cause. In terms of supporting a transition, such discussions, however, call for further insights concerning justice, equity and power relations, for example.

The last category to be detailed includes those models of use that received ratings that suggested that they were perceived to be commonly adopted in the informants’ departments. In this category, statements related to instrumental use were rated as the most agreed with, along with statements for enlightenment. Nevertheless, there were informants that tended not to agree with the statements as well; for example, the statement saying that an evaluation is intended to “lead to immediate changes in the program that is being evaluated”, which five informants rated with a 2 or 3.

The last model of use to elaborate further is unintended use. Illustrated through statements that placed emphasis on widening evaluation application areas beyond the evaluand in itself, either for using the knowledge in discussions not immediately related to the evaluand or for highlighting new areas for further investigation, the informants rated unintended use more towards the right side of the spectrum. The statement “spur further discussions about other programs” predominantly received ratings of 4 and 5. This model of use is of particular interest when seeking to support the use of evaluations for a transition, since it refers to making evaluations wider in their approach in order to also cater for effects outside of the program or to capitalize on learning that can be valuable to apply in other circumstances. Such use can support a transition in that evaluations are seen as carriers of knowledge and learning that are intended to be used beyond the sphere of the evaluand itself. That informants indicate unintended use as balancing between neutral and positive suggests an opening for applying transformative evaluation in Swedish agencies. The next section focuses on the perceptions of the benefits and challenges of conducting transformative evaluation.

Lastly, the questionnaire provided opportunities for informants to elaborate on the most important benefits and limitations of their current evaluation practices. These responses showed that some informants emphasized an independent evaluation practice, where the selection of methods, theories

and any delimitations were to a large degree at the discretion of the agencies, and that this independence also included the freedom to arrive at any conclusion, however inconvenient it may be for the evaluand. Others emphasized that the evaluation practice was recurring, mainly conducted externally and that evaluations were planned already at the design phase of an implementation. Conversely, the limitations of current evaluation practices included access to data, a lack of resources and notably a lack of time, requiring the balancing of ambition and the desire to conduct an evaluation in a certain way and the given deadlines for delivering the final product. Others emphasized limitations in the commissions of evaluations, which called for the application of indicators which may not be relevant, or a lack of systematization in evaluation processes; e.g., concerning routines for addressing and following-up measures as suggested by an evaluation. Another limitation mentioned by one informant was that there is currently an emphasis on quantitative evaluations, which limits the ability to perform evaluation to create understanding and learning.

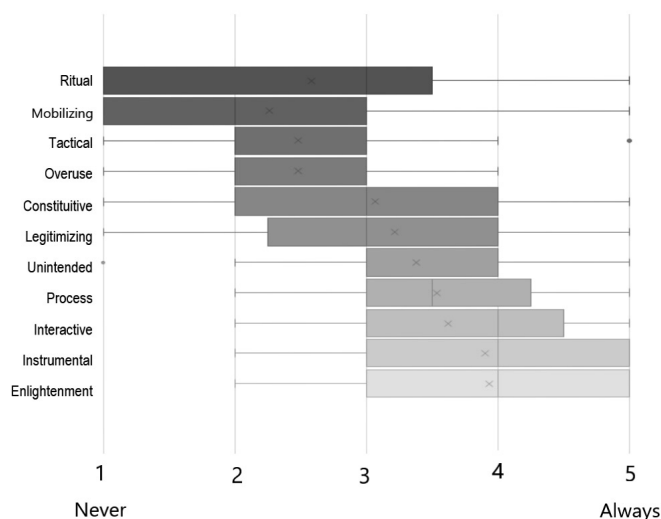


Figure 2. The aggregate ratings of the models of use of evaluations. Each model was represented by three statements which the informants were asked to rate between 1–5, where 1 meant “never” and 5 meant “always” (for the full list of statements, see Appendix B). Ten informants out of 11 answered these questions. Thus, each model of use could receive a maximum of 30 ratings (10 informants x three statements to rate). The results are presented as aggregate responses for all three statements per model of use. Answers are shown sample-wide, including all agencies that were represented in the questionnaire. The number of responses per model of use ranges from 27–30.

4.2. Benefits and Challenges of Transformative Evaluation

As described in Section 1.1, an approach for transformative evaluation has been proposed. In order to acquire insights about how such an approach may be adopted and integrated by Swedish state agencies, an initial scoping of benefits and challenges was conducted using focus groups in which four state agencies took part: the Swedish Energy Agency, the Swedish Environment Protection Agency, Vinnova and Growth Analysis. The input shared at these focus group sessions was grouped and categorized according to content.

The comprehensive findings from this initial scoping are presented in Table 2. Generally, the results show that representatives from the agencies articulated more challenges than benefits in terms of adopting transformative evaluation, but also that many of the challenges and benefits were related, as benefits tended to relate to the envisioned outcome of transformative evaluation, whereas the

challenges were related to the processes of adopting and executing such an approach. Thus, the agencies that took part in the focus group sessions seemingly saw a value in the proposed evaluation approach from the perspective of what it may deliver in terms of useful evaluation processes and results, but also expressed concerns about practicalities related to realizing these benefits.

Table 2. Benefits and challenges of Transformative evaluation, as articulated during focus group sessions with four Swedish agencies: the Swedish Energy Agency, the Swedish Environment Protection Agency, Vinnova and Growth Analysis. The first four rows in the table present categories that are seen as both beneficial and challenging; the following rows present categories bound to either benefits or challenges.

Benefits of Transformative Evaluation	Challenges of Transformative Evaluation
<i>A new way of approaching evaluation</i>	
Balances the requirements of smaller vs. more encompassing evaluations. Reduces stakes when evaluations become part of a bigger picture and gives room for experimentation.	Adoption requires change, which requires courage and acceptance from stakeholders. It does not harmonize with the current system for funding and the design of research and policy. Requires addressing current path dependencies.
<i>Competency requirements</i>	
The need for different evaluation competencies will be distributed more evenly. Less pressure on one stakeholder to host wide competencies.	A transformative approach requires the development of competencies by both commissioners and evaluators.
<i>Evaluation planning and logistics</i>	
The approach supports a pragmatic and deliberate planning of evaluation needs and calls for a connection between small-scale and large-scale evaluations.	The combining of evaluations of various scales into a more holistic picture means a continuous balancing of the need for details, the available data and the wider context. It is logistically challenging to involve many actors in an evaluation.
<i>Knowledge transfer and documentation</i>	
The evaluation approach supports a systematic documentation and logging of findings and insights, as well as a continuous knowledge transfer for learning between various evaluations and stakeholders.	The documentation and knowledge transfer required to execute the evaluation approach is challenging to realize.
<i>Guiding</i>	
A broader evaluation approach supports evaluation practices by being an inspiration, by maintaining ideas of what evaluation should (not) be and by assisting both evaluators and commissioners in providing a joint language for articulating expectations and designs.	<i>Ownership, agency and interests</i> A sustainability transition has stakeholders with vested interests, which calls for transparency and a continuous motivation for change. Issues to be resolved include the ownership of the evaluation approach and responsibilities and agency in managing such an approach. Responsibilities between levels such as units, agencies and state departments need to be determined.
<i>Potential</i>	
The approach highlights the transformational contributions of the evaluand and provides a possibility to identify drivers for change and how to support a sustainability transition.	<i>Current procurement and commissioning routines</i> There is a limited budget for evaluations, and commissions for evaluations are commonly limited to narrow requirements or are limited by what is allowed under a direct award. Master agreements may limit who can be commissioned to perform evaluations. There is a limited number of actors performing evaluations, which affects possibilities for variation.
<i>Goals and purposes for evaluation</i>	
	A sustainability transition may have potentially conflicting goals on an aggregate level. Evaluations are static, but the purposes and goals (of a transition) may change. Previously set goals may become irrelevant. There may be a mismatch between what is stated in (old) policy documents and what should be evaluated according to a broader transformative approach. Currently, the purpose of evaluating is commonly to determine whether to legitimize a new program period.
<i>Current use</i>	
	The use of evaluations today is limited by various factors.

Building on this initial scoping, the questionnaire sent to the agencies asked informants to rate statements concerning the challenges and benefits related to adopting transformative evaluation. Since the number of informants was rather low, these results should be seen as indicative and

should be regarded as complementing the focus group discussions regarding benefits and challenges. Nevertheless, the results from the questionnaire suggest that informants at all agencies represented believed that the current evaluation practices can be strengthened by collaborating with external actors, and that current practices can be revised and altered to cater for a transformative focus (see Appendix C for ratings of statements).

Open-ended questions concerning other benefits or challenges that were not captured in the abovementioned statements yielded some additional valuable comments. Regarding benefits, informants stated that the broadening of evaluation approaches was beneficial, and that such a broadening potentially could be done incrementally to ensure that different actors could adapt. Moreover, it was also emphasized that there should be a clearer and more deliberate focus on the use of evaluation, making it an integrated and interactive tool among different stakeholders in planning an intervention to strengthen its implementation and flexibility in a changing context.

As for the challenges, responses to a large degree corresponded to the challenges outlined in evaluation theory related to use of evaluations. Organizational resistance, as described by Weiss [27], includes challenges related to transforming current practices and routines, the money and time needed to do so and the acquiring of new skills, as well as the protection of the organization. Chelimsky [22] highlights issues inherent to the mandates and abilities to act within or across administrative units or departments, for example. While these concerns focus on the use of (challenging) evaluation results, they are arguably translatable to the challenges outlined in this study for transforming the evaluation practices, as well as the use of evaluation results for supporting a sustainability transition. Results from the questionnaire show that coordinating an evaluation between multiple stakeholders and agencies is challenging, particularly when there are conflicting goals between different departments and when findings require actions within another actor's jurisdiction. One informant highlighted that there is a lack of political priorities and coherence between different implementations of both research and policy, which further complicates the adoption of transformative evaluation. Moreover, it was emphasized that transformative evaluation needs to be clearly anchored in the evaluation processes in order to prevent it from becoming reliant on certain key people and thus hinging on their remaining in the process. Other issues highlighted in evaluation theory include access to relevant and reliable data for conducting an evaluation [27,40], which was also voiced by informants, notably for performing evaluations with a transformative focus. It was specifically mentioned that it is rare that implementations are designed to be readily evaluated, which makes data collection difficult.

On a slightly different note, one informant mentioned that transformative evaluation should not be adopted "for the sake of it"; instead, the need for knowledge should guide the evaluation inquiry, rather than an enforced intention of placing focus on a certain place. This is clearly a valuable point, which the author argues should be accommodated in the proposed evaluation approach in that it builds on a variety of different evaluations of different scopes and foci. Thus, the inquiry should guide the design of the evaluation, and the transformative focus can then be adopted to support a deeper learning.

Lastly, the questionnaire sought to shed light on whether a clear transformative focus in evaluations was perceived as being able to support a more active use of evaluations. The answers were varied, with four answering "yes", five answering "do not know" and one answering "no". Clearly, the issue of increasing the use of evaluations is not as easily addressed as simply re-framing the utility of evaluations; however, when including the open-ended answers in which informants were asked about their thoughts on adopting transformative evaluation, five said that it would be interesting and positive. One informant elaborated, saying that it would indeed support a more active use and increased learning from evaluations. Others elaborated on the challenges, including competence development, and how such an approach may divert from the inquiries that need to be made in evaluations.

5. Discussion

Building on insights gained from focus group sessions and the structured interview questionnaire, a key indication is that informants at Swedish state agencies express both a concern that evaluations are not currently used to their full potential and—importantly—that there is an interest in discussing and increasing their use. Another key indication is that participants at the focus group sessions and in the questionnaire expressed a curiosity and general openness to adopting an evaluation approach that seeks to provide insights on transformative efforts, with some claiming that it can strengthen evaluation practices and use. While the number of informants was rather low, particularly for the structured interview questionnaire, the combination of insights from the focus group sessions and the questionnaire still allows the study to provide an indication about issues that often otherwise remain theoretical.

5.1. Potential of Different Models of Use

The analysis of which models of use that were perceived as most agreed with shows models of use that seem quite supportive of a transformative evaluation approach, as they largely represent a constructive use in which knowledge is attained and used. These models include enlightening use, instrumental use, process use, interactive use and even unintended use, which in itself represents the conveyance of knowledge or insights from an evaluation beyond the evaluand: e.g., providing input for other evaluations for a transformative perspective. In theory, these models of use are arguably more supportive of a transformative evaluation approach than their counterparts of overuse or ritual use, which are less prone to promoting knowledge production and use for a transformative cause. Regardless of the model of use, the most important aspect is, however, that the evaluations are used not only to improve a program—this kind of use is clearly also important—but that they also are used to synthesize results for a more holistic picture.

A word is warranted on instrumental use in particular, as it was one of the models of use that received one of the overall highest ratings among the 11 models in this study. While this model is noted by scholars as not very prominent in reality [24], due to its rather direct approach to use which contrasts with the complexity of a real policy process [29,48], it has been shown to be applied in the European Commission, for example, in decision-making based on legislative evaluations, even under political opposition [31], or as indirectly applied as evaluation findings inform policy development ex-ante [49]. Thus, while the results of this study at first glance may seem idealistic and to showcase a mere vision of what the use of evaluation could be, they should be considered against the fact that instrumental use is not altogether disregarded in practice.

On this note, it also needs to be acknowledged that it is difficult to fully assess the degree to which informants rated statements in relation to their perception of the actual practice in their department or from the perspective of what the intended evaluation use is. Some individual statements underpinning the models of use may be argued to be difficult to disagree with from the perspective of what evaluation may be perceived to be in an ideal case. Thus, there may be a discrepancy in the reported uses, in that some may relate to the ideal case of what evaluation is envisioned to accomplish, and the use in reality. Another factor contributing to this discrepancy can also be rooted in the informant's insights regarding the entire evaluation processes in the department in question. For example, while some may conduct evaluations, they may not be responsible for making use of the results and may thus be uninformed of what comes after an evaluation is finalized.

Finally, while the 11 models of use applied in this study are treated as separate, it should also be noted that the lines between them may become vague, as the statements underpinning them become subjected to interpretation by the informants. Furthermore, the differences between the models of use are at times fine; for example, between legitimizing and mobilizing use, which both refer to use intended to serve a stated viewpoint but with slightly different intended reactions of generating acceptance or activation among actors. Thus, the models of use should perhaps not be solely viewed

as independent entities, but rather as parts of a spectrum and at times linked and interlaced with each other.

5.2. Strengthening the Use of Evaluations

Turning to the means for strengthening the further use of evaluations in order to support a transition, the results from the questionnaire suggest that the modes of communicating the results, and the intended main users, depend on the purpose of conducting the evaluations. While this may seem fairly obvious, it does raise further questions regarding whether the purpose, the main users, and the channels used for communication are indeed prepared to support a transition. A previous study of current Swedish evaluation practices implied that there is rarely a stated objective within an evaluation to provide knowledge to support a transition [38], and the focus of evaluations is often geared towards a program level. Therefore, while the purpose and reported uses are still seemingly not yet fully aligned for conducting evaluations with a transformative focus, this study indicates a rather robust and varied approach to disseminating results. Communication is emphasized in evaluation theory as paramount for the successful use of evaluations, and provided that the purpose and intended use are focused towards supporting transitions, there is a large potential in capitalizing on current communication practices in terms of targeting various actors and using tailored communication formats (reports, seminars, films, etc.) and channels (interpersonal, physical and digital).

Lastly, the Swedish state agencies that took part in this study—both in the focus group sessions and in the questionnaire—expressed both excitement and concern regarding adopting transformative evaluation. On the one hand, it was acknowledged that evaluations today are not always used to their full potential, and taking a transformative approach is perceived as bringing many benefits, both for enhanced use and for supporting a more holistic approach. Based on the insights gained in this study, it seems that Swedish agencies as a minimum perceive evaluations to be used in constructive ways, supporting the adoption of transformative evaluation. The incorporation of actors outside of an agency is not uncommon and could likely become more deliberate and possibly even extended if the agencies were to adopt a transformative evaluation approach that builds on cooperation across departmental boundaries. Subsequently, when the focus of the evaluation moves towards transformative contributions, the intended users will also move from a mainly program-oriented level to actors in other parts of the system.

On the other hand, concerns included the costs in terms of the time and resources required to adopt such an approach and the difficulties that may come with crossing institutional boundaries and mandates between various state departments. Another concern is that any evaluation approach taken needs to be guided by what kind of knowledge is sought, and to impose a certain approach would risk a skewed evaluation. These are all valuable concerns, which call for future research, notably in terms of testing how transformative evaluation can be applied in a real evaluation setting in order to ensure that it does not contribute to a misleading outcome.

5.3. The Role of Framing Evaluations for a Transition

Another issue to consider for further research is the role of framing in evaluations. Evaluation theory states that users need to be involved in evaluations to enhance their use; however, to support a transition, users must also see the value in evaluations supporting such transitions. Thus, it can be argued that framing theory may provide key insights to ensuring a more extensive use of evaluations, in that evaluations are presented and accepted as carriers of information that can support a transition. On the one hand, framing theory may support the promotion of a framing that places evaluation findings and their synthesis in a transition context, emphasizing the importance of this knowledge-production method for aligning and strengthening transformative efforts. On the other hand, it can also easily be argued that evaluations themselves are subjected to, and are conducted under, the influence of frames, which leads the design and the focus towards certain endpoints—another reason for further uncovering the role of frames in evaluations.

There may be a potential in framing evaluations as carriers of information that can support a transition, not for the sake of misleading or imposing decisions but for shifting focus away from the program level and towards harnessing the potential of evaluations. Such a framing needs to be communicative, meaning that the commissioning, designing, conducting and communication of an evaluation should uphold this frame in order to promote cognitive frames used by receivers to address and act on evaluations [50–52]. It is important to note, however, that frames are crafted by both including and excluding information and viewpoints, which may affect how receivers respond to them [52]. Thus, frames that seek to support evaluations' transformative potential need to be carefully crafted, to be transparent and to be made accessible to users. Transformative evaluation advocates an ongoing synthesis and alignment of multiple evaluations, which balances findings guided by different perspectives (and frames) and thus represents a more versatile and holistic depiction of the effects that research initiatives and policy instruments provide in terms of realizing a transition.

6. Conclusions

This study shows that the current practices among the represented Swedish state agencies when using evaluation findings are varied, but importantly that the reported modes of using evaluations are generally constructive and arguably supportive of learning and use on a transformative scale. The reported processes for designing and conducting evaluations show collaboration between actors both in public and private sectors, which can support a wider approach to evaluation. The informants recognize a value in moving towards transformative evaluation approaches, related to collaboration and increased evaluation use, for example. Challenges that are outlined in relation to adopting a transformative evaluation approach include overcoming institutional boundaries to coordinate and collaborate on evaluations and balancing the need for different evaluation scales and foci. To capitalize on the potential of increasing the use and usefulness of evaluations, the time seems ripe to deliberately address how to further elaborate the communication of evaluations, how to support cross-agency coordination and the articulation of intended goals, and how to develop an extended collaboration with relevant actors and stakeholders of a transition towards a more sustainable energy system and society.

Funding: This work was funded by the International Institute for Industrial Environmental Economics (IIIEE) at Lund University. This work builds on previous work conducted in the Transition Governance project (2016-2019) supported by the Swedish Energy Agency [grant number 39938-1]. The APC was funded by Lund University.

Acknowledgments: The author is grateful for the valuable feedback and support provided by Lena Neij in terms of the conduct of this study and for the input provided by Maria Johansson on the design of the structured interview questionnaire. The author is also thankful to Evert Vedung for reading the manuscript and for providing input and inspiring discussion.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

A structured interview questionnaire was sent to seven Swedish agencies that work with the evaluation of research initiatives and policy instruments within the fields of energy and/or climate related issues. The original questionnaire was sent out in Swedish. An English translation is provided below.

Part 1: Background

1. At which state agency are you working?

- The Swedish Energy Agency
- The Swedish Environment Protection Agency
- Growth Analysis
- Vinnova

- Formas
- National Board of Housing, Building and Planning
- The Swedish Agency for Economic and Regional Growth
- Other (open answer)

2. What are your main tasks, concerning evaluation?

- I work strategically with issues related to evaluation
- I commission evaluations
- I conduct evaluations
- Other (open answer)

3. What is your experience in working with evaluation-related tasks?

Rate on a scale from 0–10.

Part 2: Evaluation Today

4. Who takes part in the design of the evaluations that you either commission or conduct at your agency? (Select one of the following: never, sometimes, often, always)

- Staff at our agency
- Staff at other agencies
- External consultants
- Researchers at universities
- Other (open answer)

5. Which actors do you hire when you commission an evaluation? (Select one of the following: never, sometimes, often, always)

- We hire consultants
- We hire researchers at universities
- We hire expertise at our own agency
- We hire others (open answer)

6. Which specific knowledge are you seeking if you commission an evaluation, or look for expertise, from an external actor? (Select one of the following: never, sometimes, often, always)

- Evaluation expertise
- Economics
- Organizational theory
- Behavioral expertise
- Modeling expertise
- We have the expertise needed within our own agency
- Other (open answer)

7. Who do you believe is/are the most important user/s of the evaluations you commission or conduct at your agency?

Open answer

8. To whom are the evaluation results communicated?

Open answer

9. How are the evaluations results communicated? (Select one of the following: never, sometimes, often, always)

- The report is only distributed internally at our agency
- The report is published on our agency's webpage
- The report is sent to other state agencies
- The report is sent to municipal or regional agencies
- The report is sent to those taking part in the evaluations; e.g., through interviews
- We invite parties to an open seminar for discussing the evaluation
- We broadcast the presentation of seminars online, where the evaluation is discussed
- We arrange physical or digital meetings with those affected by the evaluation
- Other (open answer)

10. The following questions refer to the purpose and the intended use of the evaluations which you commission or conduct at your agency. The questions follow the same format, where you are asked to rate on a scale between 1–5 how well the following statements apply to your department, where 1 is never and 5 is always.

Evaluations of research initiatives or policy instruments that we either commission or conduct at my department are intended to . . .

Lead to immediate changes in the program that is being evaluated

Provide general knowledge about how a certain type of program works

Give trust and support for decisions that concern a program

Show that "something is being done"

Be used to create support among others for the evaluated program

Be used as a knowledge base to be used for issues that are outside of the evaluated program

Evaluations of research initiatives or policy instruments that we either commission or conduct at my department are intended to . . .

Not emphasis on describing why the results show what they show

Lead to important insights during the evaluation process

Prevent hasty decisions, by allowing the evaluation process to take time

Buy additional time in a decision-making process

Be performed because it is expected that an evaluation is performed

Contribute to knowledge development that creates immediate measures for improvement in the evaluated program

Evaluations of research initiatives or policy instruments that we either commission or conduct at my department are intended to . . .

Be broad and robust enough to be the only basis for decisions regarding the evaluated program

Be conducted to show that the program at hand has been followed-up on

Influence actors to consider what needs to be done to meet the expectations of the planned evaluation

Increase the understanding of how a program should be implemented, rather than lead to concrete measures in the program

Spur improvements in a program by communicating that an evaluation is to be done

Indicate areas outside of the evaluation boundaries that need further investigation

Contribute to showing that we are doing things correctly

Evaluations of research initiatives or policy instruments that we either commission or conduct at my department are intended to . . .

Provide an overview of which aspects (e.g., administration, behaviour, economy, market) are affected by a program

Be used in combination with other material in decision-making

Be conducted as per usual, so that actors affected by the evaluation will know what to expect

Confirm what we know about a research program or policy instrument

Show how the evaluated program should be changed

Be used for decisions about the program that are entirely based on what the evaluation shows

Spur further discussions about other programs

Evaluations of research initiatives or policy instruments that we either commission or conduct at my department are intended to . . .

Focus on pre-determined criteria and indicators, regardless of whether the program or the situation has changed

Primarily contribute with learning and knowledge during the evaluations process, through interactions between different actors

Be used to convince opponents

Realize effects as early as possible, before the evaluation itself is done

Only to be one part of the knowledge basis in a decision process about a program

Spur change and improvement already during the evaluation process, through dialogues with different actors

Be used to encourage actors to support a program or a viewpoint

11. What do you consider to be the largest benefits of the current evaluation practices at your agency?

Open answer

12. What do you consider to be the largest limitations of the current evaluation practices at your agency?

Open answer

Part 3: Evaluation with a transformative focus

13. Are you familiar with the concept of "transition" (transformative change)?

Yes/No

14. Do you believe that a deliberate focus on capturing and supporting transformative change for a more sustainable society may spur a more active use of evaluations?

Yes/No/Do not know

15. What are your thoughts on allowing evaluations that are commissioned or conducted at our agency to take a transformative focus?

Open answer

Info-page about the proposed broader approach to evaluation with a transformative focus

16. Have you had the possibility to read or acquaint yourself with ‘Vägledning för utvärdering av transformativ omställning’ sedan tidigare?

Yes/No

17. Do you have any viewpoints you wish to share regarding the ‘Vägledning för utvärdering av transformativ omställning’?

Open answer

Part 4: Possibilities and delimitations

18. Rate on a scale between 1–5 the extent to which you agree with the statement, where 1 is do not agree and 5 is absolutely agree. (Do not know also available)

Our evaluation-related work can be strengthened through cooperation with other agencies

It is difficult to work with a broad transformative evaluation approach because there may be disagreements between different actors on what a transition should entail

We can change our specifications of requirements when we commission or conduct an evaluation to include a transformative focus

A broad transformative evaluation approach is limited by the agency’s budget for evaluation

Our evaluation-related work can be strengthened through cooperation with actors from academia and external consultants

To coordinate different governmental agencies and other actors under the umbrella of a broad evaluation approach with a transformative focus is difficult

A broader evaluation approach with a transformative focus would contribute to a better planning of evaluations than the case today

19. Rate on a scale between 1–5 the extent to which you agree with the statement, where 1 is do not agree and 5 is absolutely agree. (Do not know also available)

Objectives and purposes of evaluations may be in conflict with each other when they are brought together—this limits the utility of a broad transformative evaluation approach

The evaluation expertise available at our agency as well as externally, is enough for a broader evaluation approach with a transformative focus

It is difficult to maintain the systematic documentation and knowledge transfer between different actors that is required for a broad transformative evaluation approach

The development of competency needed among those commissioning and those conducting evaluations for a broader evaluation approach with a transformative focus can easily be acquired

A broader evaluation approach with a transformative focus is hindered by current routines for procurement and framework agreements for evaluations

We dare to try and we are supported to try new ways of doing things in my department

We are already using our evaluations in a way that supports transformative change in the energy system and in society

It is difficult to apportion ownership, responsibilities and mandate to make decisions about steps of transformative processes between different actors

20. Do you see any other possibilities connected to applying a broader evaluation approach with a transformative focus that is not mentioned in the above statements?

21. Do you see any other delimitations connected to applying a broader evaluation approach with a transformative focus that is not mentioned in the above statements?

Appendix B

Eleven models of use were identified in evaluation theory. Based on the characteristics of each model, three statements were designed to capture the essence of what the model of use entails. This appendix shows the statements used to illustrate each model of use. These statements were scrambled and presented as stand-alone statements in the self-administered questionnaire sent out to Swedish state agencies, as presented in Appendix A.

Table A1. Eleven models of use were identified in evaluation theory. Based on the characteristics of each model, three statements were designed to capture the essence of what the model of use entails. Original text in Swedish in brackets.

Model of Use	Characteristics
1. Instrumental	Show how the evaluated program should be changed (<i>Visa hur på det utvärderade programmet bör förändras</i>)
	Lead to immediate changes in the program that is being evaluated (<i>Leda till direkta korrigeringar i det utvärderade programmet</i>)
	Contribute to knowledge development that creates immediate measures for improvement in the evaluated program (<i>Bidra till kunskapsuppbyggnad som skapar omedelbara förbättringsåtgärder i det utvärderade programmet</i>)
2. Enlightenment/conceptual	Increase the understanding of how a program should be implemented, rather than lead to concrete measures in the program (<i>Öka förståelsen för hur ett program bör genomföras snarare än att leda till konkreta handlingar</i>)
	Provide general knowledge about how a certain type of program works (<i>Ge övergripande kunskap om hur en viss sorts program fungerar</i>)
	Provide an overview of which aspects (e.g., administration, behaviour, economy, market) are affected by a program (<i>Ge en överblick över vilka olika aspekter (t.ex. administration, beteende, ekonomi, marknad) som påverkas av ett program</i>)
3. Legitimizing/Reinforcing use	Contribute to showing that we are doing things correctly (<i>Bidra till att visa att vi gör saker rätt</i>)
	Give trust and support for decisions that concern a program (<i>Ge förtroende och stöd för beslut som rör ett program</i>)
	Confirm what we know about a research program or policy instrument (<i>Bekräfta det vi vet om en forskningsinsats eller styrmedel</i>)
4. Interactive	Be broad and robust enough to be the only basis for decisions regarding the evaluated program (<i>reverse positive</i>) (<i>Vara bred och robust nog att utgöra det enda beslutsunderlaget för beslut om programmet.</i>)
	Only be one part of the knowledge basis in a decision process about a program (<i>Endast vara en del av kunskapsunderlaget i en beslutsprocess om ett program</i>)
	Be used in combination with other material in decision-making (<i>Användas i kombination med annat material vid beslutsfattande</i>)
5. Ritual/Symbolic use/ Mechanical use	Be conducted to show that the program at hand has been followed-up on (<i>Utföras för att visa att programmet i fråga har följts upp</i>)
	Be performed because it is expected that an evaluation is performed (<i>Utföras för att det förväntas att en utvärdering utförs</i>)
	Be conducted as per usual, so that actors affected by the evaluation will know what to expect (<i>Utföras som vanligt, så att aktörer som berörs av utvärderingen vet vad som väntas</i>)

Table A1. Cont.

Model of Use	Characteristics
6. Mobilizing use/Persuasive use	Be used to encourage actors to support a programme or a viewpoint (Användas för att uppmuntra aktörer att stödja ett program eller en ståndpunkt)
	Be used to convince opponents (Användas för att övervinna meningsmotståndare)
	Be used to create support among others for the evaluated program (Användas för att skapa stöd hos andra för programmet)
7. Overuse	Focus on pre-determined criteria and indicators, regardless of whether the program or the situation has changed (Fokusera på förutbestämda kriterier och indikatorer, oavsett om programmet eller situationen har förändrats)
	Be used for decisions about the program that are entirely based on what the evaluation shows (Användas för programbeslut som helt baseras på vad utvärderingen visar)
	Not put emphasis on describing why the results show what they show (Inte lägga stor vikt vid beskrivning av varför resultaten visar det som de visar)
8. Process use	Spur to change and improvement already during the evaluation process, through dialogues with different actors (Sporra till förändring och förbättring under utvärderingsprocessen genom dialoger med olika aktörer)
	Primarily contribute with learning and knowledge during the evaluations process, through interactions between different actors (Främst bidra med lärande och kunskap under utvärderingsprocessens gång genom interaktionen mellan olika aktörer)
	Lead to important insights during the evaluation process (Leda till viktiga insikter i utvärderingsprocessen)
9. Constitutive/Anticipatory use	Realize effects as early as possible, before the evaluation itself is done (Få effekter så tidigt som möjligt, innan själva utvärderingen är klar.)
	Influence actors to consider what needs to be done to meet the expectations of the planned evaluation (Påverka aktörer att fundera över vad som behöver göras för att leva upp till förväntningarna av den planerade utvärderingen)
	Spur improvements in a program by communicating that an evaluation is to be done to those whom the evaluation concerns (Sporra till förbättring genom att den kommande utvärderingen kommuniceras till berörda aktörer)
10. Tactical use	Buy additional time in a decision-making process (Köpa ytterligare tid i en beslutsprocess)
	Show that "something is being done" (Visa på att 'något görs')
	Prevent hasty decisions, by allowing the evaluation process to take time (Förhindra förhastade beslut, genom att utvärderingsprocessen tillåts ta tid)
11. Unintended use	Spur further discussions about other programmes (Sporra vidare diskussioner om andra program)
	Be used as a knowledge base to be used for issues that are outside of the evaluated program (Användas som kunskapsunderlag för frågor som står utanför det utvärderade programmet)
	Indicate areas outside of the evaluation boundaries that need further investigation (Påvisa områden utanför utvärderingens gränser som kräver ytterligare utredning)

Appendix C

Two figures that display how informants at five Swedish state agencies rated positively and negatively framed statements regarding the benefits and the challenges related to adopting a transformative evaluation approach are presented. The statements were guided by an initial scoping of perceptions at focus group sessions with four state agencies.

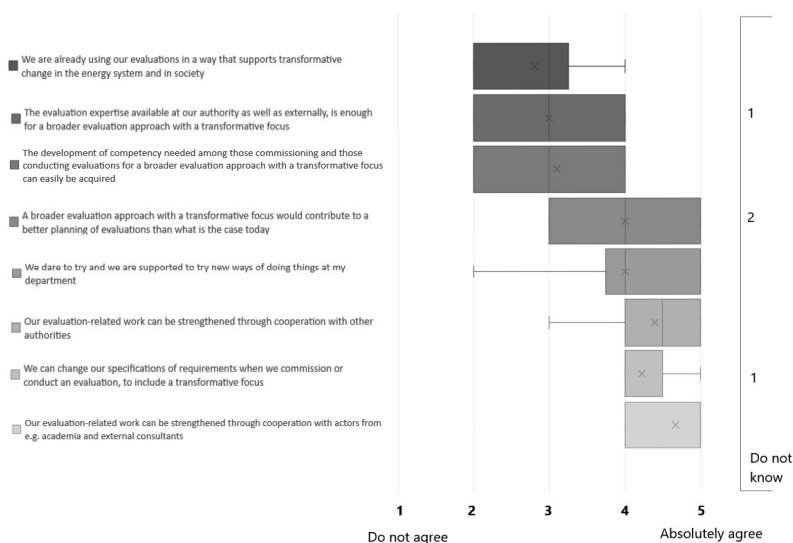


Figure A1. Positively framed statements related to benefits and challenges to adopting a transformative evaluation approach, which informants were asked to rate in a self-administered questionnaire between 1–5, where 1 meant “do not agree” and 5 meant “absolutely agree”. Ten informants out of 11 answered these questions. Answers are shown sample-wide, including all agencies that were represented in the questionnaire. The number of responses, including “do not know” for each statement is ten, except for the last statement with nine answers.

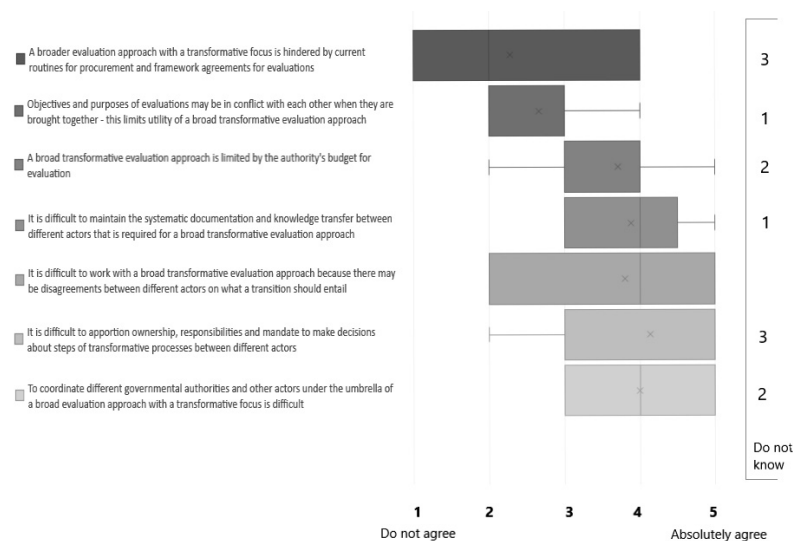


Figure A2. Negatively framed statements related to benefits and challenges to adopting a transformative evaluation approach, which informants were asked to rate in a self-administered questionnaire between 1–5, where 1 meant “do not agree” and 5 meant “absolutely agree”. Ten informants out of 11 answered these questions. Answers are shown sample-wide, including all agencies that were represented in the questionnaire. The number of responses, including “do not know” per statement is 10, except for statement number three with nine answers.

References

1. European Environment Agency. *Trends and Projections in Europe 2018: Tracking Progress Towards Europe's Climate and Energy Targets*; Publications Office of the European Union: Luxembourg, 2018; ISBN 978-92-480-007-7.
2. International Energy Agency; Organization for Economic Co-operation and Development. *World Energy Outlook 2018*; OECD: Paris, France; IEA: Paris, France, 2018.
3. Masson-Delmotte, V.; Zhai, P.; Pörtner, H.-O.; Roberts, D.; Skea, J.; Shukla, P.R.; Pirani, A.; Moufouma-Okia, W.; Péan, C.; Pidcock, R.; et al. (Eds.) IPCC Global Warming of 1.5 °C. In *An IPCC Special Report on the Impacts of Global Warming of 1.5°C Above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*; 2018; in press.
4. Geels, F.W. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Res. Policy* **2002**, *31*, 1257–1274. [CrossRef]
5. Farla, J.; Markard, J.; Raven, R.; Coenen, L. Sustainability transitions in the making: A closer look at actors, strategies and resources. *Technol. Forecast. Soc. Chang.* **2012**, *79*, 991–998. [CrossRef]
6. Markard, J.; Raven, R.; Truffer, B. Sustainability transitions: An emerging field of research and its prospects. *Res. Policy* **2012**, *41*, 955–967. [CrossRef]
7. Weinstein, M.P.; Turner, R.E.; Ibáñez, C. The global sustainability transition: It is more than changing light bulbs. *Sustain. Sci. Pract. Policy* **2013**, *9*, 4–15. [CrossRef]
8. Somanathan, E.; Sterner, T.; Sugiyama, T.; Chimanikire, D.; Dubash, N.K.; Essandoh-Yeddu, J.; Fifita, S.; Goulder, L.; Jaffe, A.; Labandeira, X.; et al. National and sub-national policies and institutions. In *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., et al., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2014.
9. European Commission. *Better Regulation Guidelines*; [European Commission Staff Working Document]; European Commission, 2017. Available online: <https://ec.europa.eu/info/sites/info/files/better-regulation-guidelines.pdf> (accessed on 7 October 2020).
10. Edler, J.; Berger, M.; Dinges, M.; Gok, A. The practice of evaluation in innovation policy in Europe. *Res. Eval.* **2012**, *21*, 167–182. [CrossRef]
11. Edler, J.; Cunningham, P.; Gök, A.; Shapira, P. (Eds.) *Handbook of Innovation Policy Impact*; EU-SPRI Forum on Science, Technology and Innovation Policy; Edward Elgar: Cheltenham, UK; Northampton, MA, USA, 2016; ISBN 978-1-78471-184-9.
12. Huitema, D.; Jordan, A.; Massey, E.; Rayner, T.; van Asselt, H.; Haug, C.; Hildingsson, R.; Monni, S.; Stripple, J. The evaluation of climate policy: Theory and emerging practice in Europe. *Policy Sci.* **2011**, *44*, 179–198. [CrossRef]
13. Mela, H.; Hildén, M. *Evaluation of Climate Policies and Measures in EU Member States: Examples and Experiences from Four Sectors; The Finnish Environment*, 19th ed.; Finnish Environment Institute: Helsinki, Finland, 2012.
14. Schoenefeld, J.J.; Jordan, A.J. Environmental policy evaluation in the EU: Between learning, accountability, and political opportunities? *Environ. Politics* **2019**, *28*, 365–384. [CrossRef]
15. Højlund, S. Evaluation use in the organizational context—changing focus to improve theory. *Evaluation* **2014**, *20*, 26–43. [CrossRef]
16. King, J.A.; Alkin, M.C. The centrality of use: Theories of evaluation use and influence and thoughts on the first 50 years of use research. *Am. J. Eval.* **2019**, *40*, 431–458. [CrossRef]
17. Milzow, K.; Reinhardt, A.; Söderberg, S.; Zinöcker, K. Understanding the use and usability of research evaluation studies1,2. *Res. Eval.* **2019**, *28*, 94–107. [CrossRef]
18. Patton, M.Q. *Utilization-Focused Evaluation*, 4th ed.; Sage Publications: Thousand Oaks, CA, USA, 2008; ISBN 978-1-4129-5861-5.
19. Shulha, L.M.; Cousins, J.B. Evaluation use: Theory, research, and practice since 1986. *Eval. Pract.* **1997**, *18*, 195–208. [CrossRef]
20. Weiss, C.H. The many meanings of research utilization. *Public Adm. Rev.* **1979**, *39*, 426. [CrossRef]
21. Weiss, C.H. Have we learned anything new about the use of evaluation. *Am. J. Eval.* **1998**, 21–34. [CrossRef]

22. Chelimsky, E. A strategy for improving the use of evaluation findings in policy. In *Evaluation Use and Decision-Making in Society: A Tribute to Marvin C. Alkin*; Christie, C.A., Vo, A.T., Alkin, M.C., Eds.; Information Age Pub. Inc: Charlotte, NC, USA, 2015; pp. 73–90.
23. Newcomer, K.E.; Hatry, H.P.; Wholey, J.S. Planning and designing useful evaluations. In *Handbook of Practical Program Evaluation*; Wholey, J.S., Hatry, H.P., Newcomer, K.E., Eds.; Jossey-Bass, Wiley: San Francisco, CA, USA, 2010.
24. Shadish, W.R.; Cook, T.D.; Leviton, L.C. *Foundations of Program Evaluation: Theories of Practice*; Reprinted; Sage Publications: Newbury Park, CA, USA, 1991; ISBN 978-0-8039-3551-8.
25. Alkin, M.C.; King, J.A. Definitions of evaluation use and misuse, evaluation influence and factors affecting use. *Am. J. Eval.* **2017**, *38*, 434–450. [CrossRef]
26. Cousins, J.B. Commentary: Minimizing evaluation misuse as principled practice. *Am. J. Eval.* **2004**, *7*.
27. Weiss, C.H. *Evaluation: Methods for Studying Programs and Policies*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 1998; ISBN 978-0-13-309725-2.
28. Rossi, P.H.; Lipsey, M.W.; Freeman, H.E. *Evaluation: A Systematic Approach*, 7th ed.; Sage: Thousand Oaks, CA, USA, 2006; ISBN 978-0-7619-0894-4.
29. Vedung, E. *Public Policy and Program Evaluation*; Transaction Publishers: New Brunswick, NJ, USA, 1997; ISBN 978-0-7658-0687-1.
30. Vedung, E. *Six Uses of Evaluation; In Nachhaltige Evaluation?: Auftragsforschung Zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann*; Hennefeld, V., Meyer, W., Silvestrini, S., Stockmann, R., Eds.; Waxmann Verlag: Münster, Germany; New York, NY, USA, 2015; pp. 187–210.
31. van Voorst, S.; Zwaan, P. The (non-)use of ex post legislative evaluations by the European Commission. *J. Eur. Public Policy* **2019**, *26*, 366–385. [CrossRef]
32. Fleischer, D.N.; Christie, C.A. Evaluation use: Results from a survey of U.S. American evaluation association members. *Am. J. Eval.* **2009**, *30*, 158–175. [CrossRef]
33. Ledermann, S. Exploring the necessary conditions for evaluation use in program change. *Am. J. Eval.* **2012**, *33*, 159–178. [CrossRef]
34. Cunningham, P.; Edler, J.; Flanagan, K.; Larédo, P. The innovation policy mix. In *Handbook of Innovation Policy Impact*; Edler, J., Cunningham, P., Gök, A., Shapira, P., Eds.; Edward Elgar: Cheltenham, UK; Northampton, MA, USA, 2016; pp. 505–542.
35. Edler, J.; Fagerberg, J. Innovation policy: What, why, and how. *Oxf. Rev. Econ. Policy* **2017**, *33*, 2–23. [CrossRef]
36. Martin, B.R. R&D policy instruments—a critical review of what we do and don’t know. *Ind. Innov.* **2016**, *23*, 157–176. [CrossRef]
37. Neij, L.; Sandin, S.; Benner, M.; Johansson, M.; Mickwitz, P. Bolstering a transition for a more sustainable energy system: A transformative approach to evaluations of energy efficiency in buildings. *Energy Res. Soc. Sci.* **2020**, in press.
38. Sandin, S.; Neij, L.; Mickwitz, P. Transition governance for energy efficiency—insights from a systematic review of Swedish policy evaluation practices. *Energy Sustain. Soc.* **2019**, *9*. [CrossRef]
39. Cook, T.D. Lessons learned in evaluation over the past 25 years. In *Evaluation for the 21st Century: A Handbook*; Chelimsky, E., Shadish, W.R., Eds.; Sage Publications: Thousand Oaks, CA, USA, 1997; ISBN 978-1-4833-4889-6.
40. Zrakat, C.A.; Clark, W. Environmental changes and their measurement: What data should we collect and what collaborative systems do we need for linking knowledge to action? In *Evaluation for the 21st Century: A Handbook*; Chelimsky, E., Shadish, W.R., Eds.; Sage Publications: Thousand Oaks, CA, USA, 1997; pp. 329–336.
41. Alkin, M.C.; Taut, S.M. Unbundling evaluation use. *Stud. Educ. Eval.* **2002**, *29*, 1–12. [CrossRef]
42. Wholey, J.S. Use of evaluation in government—the politics of evaluation. In *Handbook of Practical Program Evaluation*; Wholey, J.S., Hatry, H.P., Newcomer, K.E., Eds.; Jossey-Bass: San Francisco, CA, USA, 2010; pp. 651–667.
43. Leviton, L.C. Evaluation use: Advances, challenges and applications. *Am. J. Eval.* **2003**, *24*, 525–535. [CrossRef]
44. Bryman, A. *Social Research Methods*, 4th ed.; Oxford University Press: Oxford, UK; New York, NY, USA, 2012; ISBN 978-0-19-958805-3.

45. Wilkinson, S. Focus group research. In *Qualitative Research Methods-Theory, Method and Practice*; Silverman, D., Ed.; Sage: Thousand Oaks, CA, USA, 2004; pp. 177–199.
46. Denzin, N.K. *The Research Act: A Theoretical Introduction to Sociological Methods*; AldineTransaction: New Brunswick, NJ, USA, 2009; ISBN 978-0-202-36248-9.
47. Evans, J.R.; Mathur, A. The value of online surveys: A look back and a look ahead. *Internet Res.* **2018**, *28*, 854–887. [CrossRef]
48. Weiss, C.H. The interface between evaluation and public policy. *Evaluation* **1999**, *5*, 468–486. [CrossRef]
49. Højlund, S. Evaluation use in evaluation systems—the case of the European Commission. *Evaluation* **2014**, *20*, 428–446. [CrossRef]
50. Carnahan, D.; Hao, Q.; Yan, X. Framing methodology: A critical review. *Oxf. Res. Encycl. Politics* **2019**. [CrossRef]
51. Chong, D.; Druckman, J.N. Framing theory. *Annu. Rev. Polit. Sci.* **2007**, *10*, 103–126. [CrossRef]
52. Entman, R.M. Framing: Toward clarification of a fractured paradigm. *J. Commun.* **1993**, *43*, 51–58. [CrossRef]



© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

IIIEE Dissertations

Leonidas Milios

Policy Framework for Material Resource Efficiency: Pathway Towards a Circular Economy
IIIEE Dissertations 2020:3

Julia Nußholz

Circular Business Model Design: Business Opportunities from Retaining Value of Products and Materials
IIIEE Dissertations 2020:2

Katherine Whalen

Circular Business Models That Extend Product Life: Going Beyond Recycling to Create New Circular Business Opportunities
IIIEE Dissertations 2020:1

Jessika Luth Richter

Towards a Circular Economy with Environmental Product Policy. Considering dynamics in closing and slowing material loops for lighting products
2019

Jonas Sonnenschein

Green growth and rapid decarbonisation? Assessing policy objectives, instrument choice and behavioural mechanisms
2019

Lars Strupeit

Scaling up solar photovoltaic use. A system-oriented assessment of experiences
2017

Alvar Palm

Residential solar photovoltaics deployment: barriers and drivers in space
2017

Thomas Parker

The Meaning of Environmental Management. An Interpretive Study of Managing Emergent or Evolutionary Environmental and Energy Strategy

2016

Jessika Luth Richter

Extended Producer Responsibility for Closing Material Loops. Lessons from energy-efficient lighting products.

2016

Olga Chkanikova

Sustainable Supply Chain Management in Food Retailing. Insights into corporate practice of managing supplier relationships

2016

Nora Smedby

Local environmental governance. Assessing proactive initiatives in building energy efficiency

2016

Matthias Lehner

Sustainability in Store. How Retailers Bring Sustainability to their Customers

2015

Bernadett Kiss

Building Energy Efficiency. Policy, learning and technology change.

IIIEE Dissertations 2013:1

Niina Kautto

Towards More Coherent and Sustainable Biomass Policy: Examining European biomass-to-energy planning

IIIEE Dissertations 2011:2

Panate Manomaivibool

Advancing the Frontier of Extended Producer Responsibility. The management of waste electrical and electrical equipment in non-OECD context

IIIEE Dissertations 2011:1

Martin Kurdve

Chemical Management Services from a Product Service System Perspective. Experiences of fluid management services from Volvo Group metalworking plants

IIIEE Dissertations 2010:1

Panate Manomaivibool

Making Sense of Extended Producer Responsibility. Towards a framework for policy transfer
IIIEE Dissertations 2009:7

Tareq Emtairah

Lost in Transition: Sustainability Strategies and Social Contexts
IIIEE Dissertations 2009:6

Åke Thidell

Influences, Effects and Changes from Interventions by Eco-labelling Schemes. What a Swan can do?
IIIEE Dissertations 2009:5

Helen Nilsson

Finding a Balance. Placing Farmers' Markets in the context of sustainability in modern society
IIIEE Dissertations 2009:4

Dagmara Nwrocka

Extending the Environmental Focus to Supply Chain. ISO 14001 as an interorganizational tool?
IIIEE Dissertations 2009:3

Beatrice Kogg

Responsibility in Supply Chain. Interorganisational management of environmental and social aspects in the supply chain. Case studies from the textile sector
IIIEE Dissertations 2009:2

Charlotte Leire

Increasing the Environmental and Social Sustainability in Corporate Purchasing. Practices and tools
IIIEE Dissertations 2009:1

Chris van Rossem

Individual Producer Responsibility in the WEEE Directive – From Theory to Practice?
IIIEE Dissertations 2008:3

Camelia Tepelus

Destination Unknown? The Emergence of Corporate Social Responsibility for Sustainable Development of Tourism
IIIEE Dissertations 2008:2

Luis Mundaca
Markets for Energy Efficiency – Exploring the new horizons of tradable certificate schemes
IIIEE Dissertations 2008:1

Adriana Budeanu
Facilitating Transitions to Sustainable Tourism
IIIEE Dissertations 2007:4

Carl Dalhammar
An Emerging Product Approach in Environmental Law – Incorporating the life cycle perspective
IIIEE Dissertations 2007:3

Kes McCormick
Advancing Bioenergy in Europe: Exploring bioenergy systems and socio-political issues
IIIEE Dissertations 2007:2

Kaisu Sammalisto
Environmental Management Systems – a Way towards Sustainable Development in Universities
IIIEE Dissertations 2007:1

Murat Mirata
Industrial Symbiosis: A tool for more sustainable regions?
IIIEE Dissertations 2005:1

Andrius Plepys
Environmental Implications of Product Servicing. The Case of Outsourced Computing Utilities
IIIEE Dissertations 2004:3

Naoko Tojo
Extended Producer Responsibility as a Driver for Design Change – Utopia or Reality?
IIIEE Dissertations 2004:2

Oksana Mont
Product-service systems: Panacea or myth?
IIIEE Dissertations 2004:1

Zinaida Fadeeva
Exploring cross-sectoral collaboration for sustainable development: A case of tourism
IIIEE Dissertations 2003:1

Philip Peck

Interest in Material Cycle Closure? Exploring evolution of industry's responses to highgrade recycling from an industrial ecology perspective

IIIEE Dissertations 2003:2

Peter Arnfalk

Virtual Mobility and Pollution Prevention: The emerging role of ICT based communication in organisations and its impact on travel

IIIEE Dissertations 2002:1

Mårten Karlsson

Green concurrent engineering: A model for DFE management programs

IIIEE Dissertations 2001:2

Kaisu Sammalisto

Developing TQEM in SMEs: Management Systems Approach

IIIEE Dissertations 2001:1

Håkan Rodhe

Preventive Environmental Strategies in Eastern European Industry

IIIEE Dissertations 2000:7

Nicholas Jacobsson

Emerging Product Strategies: Selling Services of Remanufactured Products

IIIEE Dissertations 2000:6

Karin Jönsson

Communicating the Environmental Characteristics of Products

IIIEE Dissertations 2000:5

Pia Heidenmark

Going Organic?

IIIEE Dissertations 2000:4

Peter Kisch

Preventative Environmental Strategies in the Service Sector

IIIEE Dissertations 2000:3

Thomas Lindhqvist

Extended Producer Responsibility in Cleaner Production

IIIEE Dissertations 2000:2

Desta Mebratu
Strategy Framework for Sustainable Industrial Development in sub-Saharan Africa
IIIEE Dissertations 2000:1

Peter Arnfalk
Information technology in pollution prevention: Teleconferencing and telework used as tools
in the reduction of work related travel
IIIEE Dissertations 1999:1

Thomas Parker
Total Cost Indicators: Operational Performance Indicators for managing environmental
efficiency
IIIEE Dissertations 1998:2

Kent Lundgren
Förnyelsebara energibärares nuvarande och framtida konkurrenskraft - föreställningar om
konkurrenskraft
IIIEE Dissertations 1998:1

Lars Hansson
The Internalization of External Effects in Swedish Transport Policy: A Comparison Between
Road and Rail Traffic
IIIEE Dissertations 1997:2

Mårten Karlsson
Green Concurrent Engineering: Assuring Environmental Performance in Product
Development
IIIEE Dissertations 1997:1

Erik Rydén
Car Scrap: Throw it Away or Make it Pay?
IIIEE Dissertations 1995:2
Also available in Swedish: Bilskrot: möjlighet eller miljöhot?
IIIEE Dissertations 1995:1

