Review of evaluations of policy instruments for energy efficiency in buildings in Sweden

Sofie Sandin
International Institute for Industrial Environmental Economics (IIIEE) at Lund University
Lund University
P.O. Box 196
SE-221 00 Lund
Sweden
sofie.sandin@iiiee.lu.se

Lena Neij
International Institute for Industrial Environmental Economics (IIIEE) at Lund University
Lund University P.O. Box 196
SE-221 00 Lund
Sweden
lena.neij@iiiee.lu.se

Per Mickwitz
Finnish Environment Institute SYKE
P.O.Box 140
FI-00251 Helsinki
Finland
per.mickwitz@ymparisto.fi

Keywords
evaluation, energy efficiency policy, buildings, case studies

Abstract
The built environment accounts for approximately 32 % of the global final energy use but also holds great potential for energy efficiency still to be harnessed (Lucon et al., 2014). To develop innovative policies for energy efficiency that are capable of realizing transformative changes, we argue that there is an urgent need for new systemic evaluation frameworks that provide learning and knowledge on how policy initiatives provide energy efficiency in the built environment. Moreover, there is a need to understand to what extent evaluation practices today apply evaluation frameworks with a systemic and transformative approach.

This paper presents a critical assessment of 30 evaluations commissioned by Swedish authorities 2005–2015 on policy instruments for energy efficiency in buildings. The assessment is based on a systemic theoretical framework developed and based on general insights from evaluation theory and transition theory. The main categories of assessment are: the data and methods applied in the evaluation, the value judgment in evaluation and the use of evaluations. The results show some important aspects of a systemic approach but also reveal room for improvements. The reviewed evaluations used multiple methods and multi-criteria analysis, providing a thorough base for evaluation. However, the design, methods and data used could be much stronger if taking a broader system approach and including aspects of side effects, attribution, rebound-effects and triangulation. The design could also more explicitly assess aspects of transitions related to visioning, experimentation and learning. The value judgment in the evaluations used one to three criteria, of which impact and effectiveness were the ones most frequently used, leaving room for further reflections and analysis. Moreover, use of the evaluations was not explicitly mentioned.

Introduction
The built environment alone accounts for approximately 32 % of the global final energy use, and holds a potential energy efficiency estimated to 50–75 % and 50–90 % in existing and new buildings respectively, including changes in design practices, technology and behavior (Lucon et al., 2014). To harness this potential, policy interventions are necessary to overcome market failures, provide new knowledge and accelerate changes in socio-technical systems towards an energy efficient built environment. In order to be effective such policy interventions need to be evaluated to create learning of actions taken.

Earlier research on evaluation practices of environmental policies show that evaluations have focused on effectiveness and goal achievement (Huitema et al., 2011; Mundaca et al. 2016). Furthermore, traditional policy interventions and evaluation approaches have focused on incremental changes (Furubo, 2013) and it is unclear to what extent these existing approaches are capable of capturing and supporting transformative processes towards an energy efficient society.

The objective of this study is to advance knowledge on current evaluation practices for policy instruments targeting energy efficiency in the built environment. The research is to shed light on the extent to which evaluations are able to support learning and transitions towards energy efficiency. For this, a critical review of energy policy evaluations has been done using

a systemic framework for evaluation. The review covers evaluations of policy instruments for energy efficiency in buildings, commissioned by Swedish authorities 2005–2015.

Theoretical framework

The theoretical framework of this paper builds on evaluation theory and the categorization developed by Alkin (2013). Alkin describes the field of evaluation using a metaphor of a tree with three roots and three branches. The roots underpin the rationale of evaluations and represent epistemology, systematic social inquiry and social accountability. These roots emphasize social and behavioral aspects of evaluations and the evaluation's role in decision-making. The branches of the tree represent (I) the data and methods applied in the evaluation, (II) the value judgment in evaluation and (III) the use of evaluations.

The paper further stresses the design of a framework capable of supporting a transition process and the realization of radical changes. Based on transition theory (Geels & Schot, 2007; Kemp, 1994; Rotmans et al., 2001), we emphasize a system-wide, as well as a multi-actor based approach in the evaluations. We also bring forward key concepts in transition theory such as visioning, experimenting and learning. In terms of visioning we emphasize a long-term perspective and the combination of ex-post and ex-ante evaluations. In terms of experimentation we stress the acknowledgement of radicalness in the design of projects, programs and policy, as well as in intended outcome. This also encompasses the evaluation of path dependences and lock-ins within the system, as well as efforts to destabilize them. We also emphasize learning and reflexivity. Learning, the act of acquiring new knowledge, skills and values, is an essential driver of transitions; reflexivity, stresses the challenge of setting goals and the need to reflect on the choice of evaluation approach in relation to the object and the objective(s) of the evaluation.

The theoretical framework of this paper, based on evaluation theory and transition theory as described above, is designed to structure a systematic and critical review of evaluations of policy instruments, and an assessment of these evaluations in terms of how they capture radical changes and transformative efforts in the socio-technical system. The framework is composed of three main categories (I) data and methods applied in evaluations (II) value judgments in evaluations, and (III) use of evaluations, holding 16 sub-categories designed for recognizing theoretical elements that are relevant for capturing transition in evaluations (appendix A). The framework applied is a development of the framework presented in Mickwitz et al (2016) which also provides a more detailed theoretical foundation.

The review is conducted based on evaluations collected using primarily state web pages and interpersonal contacts with Swedish authorities (e.g. the Swedish Energy Agency and the National Board of Housing, Building and Planning) as well as with external consultants conducting the evaluations.

Results

In all, the review covers 30 evaluations of policy instruments for energy efficiency in buildings that have been conducted or commissioned by Swedish authorities in 2005–2015 (appendix B). These evaluations cover the following policy initiatives:

- Legislative instruments: revision of building codes and energy requirements (two evaluations).
- Financial instruments: subsidies and tax reductions for energy efficiency measures and subsidies for performing energy audits (six evaluations).
- Informative instruments: demonstration projects, municipal energy advisory programs, energy performance certificates for buildings and online information portals (17 evaluations).
- Other instruments: technology procurement programs and cooperative network programs for energy efficiency in buildings (Lågan, Bebo, BeLOK, BeLivs and HyLOK) (five evaluations).

The evaluations commonly find that the policy instruments have been working well, but have the potential to improve the performance further. Policy instruments that are generally receiving the most favorable results in the evaluations are the municipal energy advisory programs and the cooperative network programs, which are praised for being important measures for increasing awareness of energy efficiency.

To improve the performance of policy instruments we review the evaluation practice being used, and how this practice is providing learning and knowledge on how policy initiatives provide energy efficiency in the built environment. Below we present the results of the review using the theoretical framework presented above.

DATA AND METHODS APPLIED IN EVALUATIONS

A key challenge in traditional evaluation practice is how to provide credible results and vital learning, and an important aspect of this challenge is the choice of methods applied in the evaluation. The methods provide information about the effects of the policy instrument, or the impact of the policy instrument if taking into account other factors influencing the outcomes and the nature of the particular context (e.g. attribution, side effects, rebound effects). By using mixed-methods, several data sets or multiple analysts, triangulation can be used to validate the results.

In the reviewed sample of evaluations, a variety of methods were used to evaluate the results of the policy instruments (Figure 1). The far most frequently collected data and used methods were document analysis and interviews, but other methods such as models and surveys were also frequently applied. Most of the evaluations performed were based on two (16/30) or three methods (10/30), one was based on one single method and three were based on four different methods. Surveys were used to provide data of both qualitative nature (e.g. opinions, importance of policy instrument) and quantitative nature (e.g. energy use). The results from both interviews and survey analyses were however largely guided toward a qualitative manner, focusing on e.g. the opinions of stakeholders regarding their perceptions of the services provided or the usefulness of the policy instrument. The quantitative methods that were used in the sample are thus models and statistical analysis, giving a strong favoring of qualitative methods (87 %).

The methods applied were used to evaluate various evaluation criteria, either as single criterion or multi-criteria evaluations. The selection of criteria was based on the purpose of the

evaluation, i.e. which questions the evaluation was intended to answer, and thus shaped the focus. Among the reviewed evaluations, the focus was commonly directed towards intended effects of the instruments (22/30), e.g. reduction of energy use, amount of measures taken for energy efficiency, technology development etc.

To assess processes of change, evaluation theory advocates the use of intervention theory (Vedung, 2009). The use of properly constructed assessments based on intervention theory can act as guiding lines throughout the evaluation process to ensure that predictable effects will not be overlooked (Vedung, 2009). Several evaluations in the review have referred to intended means of implementation and outcomes of a policy instrument (17/30). The evaluations did not explicitly use intervention theory, but used such an approach to illustrate the functions and mechanisms of a policy instrument. The scope of these constructions was varying from a brief explanation of the intended flow of the process to more elaborate schematic flow charts covering e.g. actors involved.

Side effects can be accounted for in evaluations to provide a wider and systemic understanding of the assessed process of change. Such side effects were mentioned in seven evaluations, in the form of e.g. marketing and competition advantages, altered behavior in terms of energy use, increased knowledge and creation of joint platforms between authorities and businesses. Only one evaluation incorporated a more thorough consideration of side effects by identifying environmental impacts connected to the policy instrument, and assessing their costs (National Board of Housing, Building and Planning, 2014). Rebound effects, i.e. the diminished net-effects of energy efficiency measures due to increased comfort use, were considered in one evaluation (National Board of Housing, Building and Planning, 2009 b).

The rather extensive use of interviews and surveys in the reviewed evaluations required involvement of many actors, and the actors selected to take part in the inquiries are likely to affect the outcome. In order to voice various opinions a multi-actor perspective is favorable**.** In the reviewed sample, 27 evaluations incorporated stakeholders and approximately half of these (14/27) did take a multi-actor approach, involving two or more groups targeted or otherwise involved in the implementation of the policy instrument. These were commonly authorities, beneficiaries and representatives from businesses and organizations, whereas actors outside of this boundary were left out.

Only twelve of the reviewed evaluations used a counterfactual construction for the assessment of attribution. The prevailing type was derived from interviews or surveys (11/12) where evaluators sought to determine the impact of the instrument on the decisions and actions of actors, followed by baselines and calculations (3/12) of scenarios without the instrument. Two of the reviewed evaluations used multiple methods for the construction of counterfactuals, which were synthesized from two respectively three methods, the latter combining surveys, calculations and a reference group (National Board of Housing, Building and Planning, 2009 a; Swedish Energy Agency, 2013). The choice of method for constructing the counterfactuals was however rarely discussed in the reviewed evaluations. Multi-actor involvement occurred in the interviews or surveys, but reflections about the selection of actors to be involved or the amount of respondents needed for constructing a robust counterfactual was not apparent.

Evaluation theory further stresses the need for triangulation, i.e. the importance of mixed-methods, complementing datasets and analysts to verify the results (Patton, 2002). The results of the review do indicate a use of multiple methods, however, triangulation in the form of using multiple methods, data sources or analysts as a tool for systemic control of the consistency and validation of results was scarce. One evaluation used document studies covering energy performance certificates to validate findings from interviews with building owners and tenants (National Board of Housing, Building and Planning, 2009 c), but the combination of different data sources or methods was otherwise commonly not used for testing of consistency of findings, but rather as complements to each other. In the cases of interviews in groups and stakeholder involvement in seminars or workshops to discuss preliminary results found by evaluators, it was seemingly a means to get feedback on the results, rather than an intention of illuminating alternative opinions or inconsistencies.

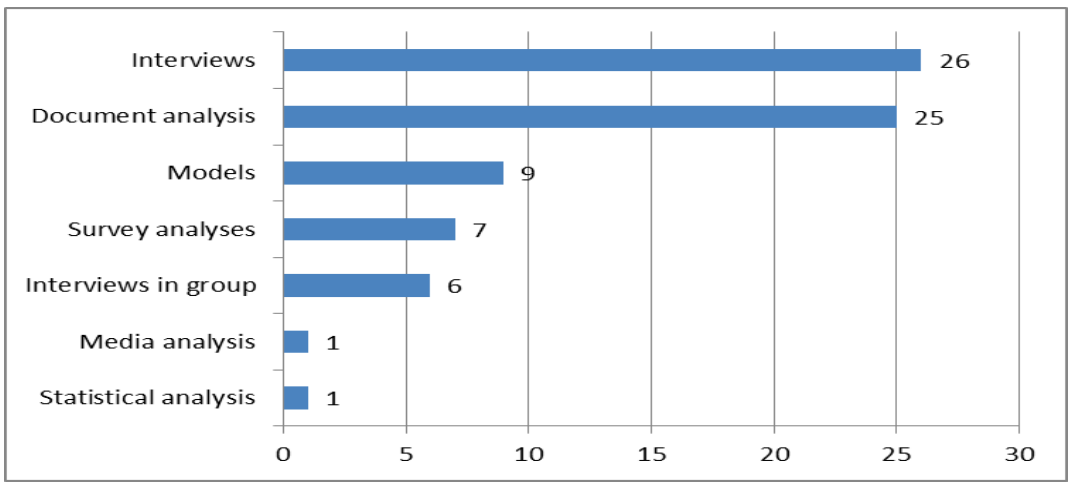

Figure 1. Type of methods used.

In this review we have not only applied traditional evaluation theory but also included aspects of transition theory to assess to what extent the methods support learning and transitions towards energy efficiency. In order to support a transition we see a need to combine methods that can provide a systemic perspective, a dimension of scale and insights to visioning (e.g. time perspective, ex-post vs ex-ante evaluations), experimenting (radical policy and potential outcome, lock-in effects) and learning (knowledge and skills).

In the reviewed sample the system approach was investigated, distinguishing system components as actors, institutions and technological factors in the socio-technical system. The results show a strong favoring of evaluating one single component (25/30), solely focusing on actors. Institutional aspects were considered in five evaluations, of which two were ex-ante evaluations concerning future tightening of regulations and building codes (Helmersdotter et al., 2015; National Board of Housing, Building and Planning, 2014). This implies that institutional structures tend to be overlooked if the purpose of the evaluation in itself is not geared towards them. Seven evaluations conducted by the same external consultant (Sweco, 2014 a-g), put the evaluated policy instruments in larger schematic pictures relating them to institutional structures, however without evaluating potential effects between them. Technological factors were acknowledged in two evaluations, one regarding technology procurement (Sweco, 2014 e) and the other regarding tightening of building codes (National Board of Housing, Building and Planning, 2014). As described above, side effects providing a wider system approach were rarely discussed or evaluated.

The transition literature also highlights visioning and long-term perspectives as important in supporting transition processes (Loorbach, 2010). In the reviewed documents we do not find any long-term evaluation strategies, however, we see a frequent evaluation practice especially in the case of policy instruments concerning information and financial programs (energy advisory, cooperative network programs and subsidies for energy audits). Visioning could also be captured and supported in evaluations through a combination of ex-post and ex-ante evaluations. The review reveals that the combination of such methods is applied, but has room for extension. One evaluation (Swedish Energy Agency, 2013) was constructed as both ex-post and ex-ante, creating an initiative for both learning and prediction of future outcomes simultaneously.

The transition literature further emphasizes experimentation (Kemp et al., 1998) and in the review we have looked into the assessments of radical outcome, potential lock-in effects and path dependences. In all, we find five evaluations taking into account criteria of experimentation. Three of them, concerning cooperative network programs (BeBo, BELOK and BeLivs) (Sweco, 2014 c), technology procurement (Sweco, 2014 e) and a demonstration program for passive houses (Grontmij, 2009) state that the evaluated instruments provided platforms for experimentations. Two evaluations, both concerning various investments in measures for energy efficiency (National Board of Housing, Building and Planning, 2009 a; Samakovlis & Vredin Johansson, 2007), stated that the instruments were unsuccessful in providing support for experimentation concerning unconventional technology and technology development. This furthermore stands in connection to consideration of lock-in effects and upholding of path dependences. In order to focus an evaluation towards radical changes, forces upholding the status quo need to be acknowledged and incorporated, along with the efforts geared towards disrupting them (Kivimaa & Kern 2016).

VALUE JUDGMENTS IN EVALUATIONS

The nature of evaluation is normative and assessing the value requires 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 various stakeholders (which will be discussed in the next sub-section), social justice and reflexivity. All these aspects are crucial to take into account in order to evaluate and support a transformative process. In the reviewed evaluations it is however not apparent to which extent stakeholders such as beneficiaries, businesses, organizations and other authorities have been invited to partake in decisions regarding evaluation design and which criteria or methods to be used in the evaluation process.

The number of criteria used for assessment within the reviewed evaluations ranged from one to three with a fairly even distribution; the average number is 1.8 criteria per evaluation. The most frequently used criteria in the reviewed sample were effectiveness and impact (Figure 2). These criteria evolve to a large extent around outcomes that can be measured in e.g. saved amount of kilowatt hours or in monetary terms, whereas criteria aimed at the mechanisms behind a successful implementation, such as acceptability, relevance and coordination with other policies were subordinated. There are many well-known shortcomings related to effectiveness as the sole or foremost criterion for value judgment. The clearest limitations are that it disregards side effects and unanticipated effects; it does not consider costs; and the relevance of the goals is not examined (Mickwitz 2002).

For capturing radical changes and transformative efforts, criteria should be sensitive not only to effectiveness and goal attainment, but also to the drivers for change and their implications. In the reviewed sample, criteria seemed to be predetermined and guided by purpose or requirements set by the commissioner. Discussion concerning the legitimacy of value claims was altogether not apparent, leaving issues concerning questioning of evaluation approach, value constructions or value dissonance in the evaluation process unmentioned.

It is also important to include criteria that are able to promote reflexivity in terms of challenging established goals and needs. Such reflexive elements were however not a predominant feature in the reviewed sample of evaluations. In all, we found nine evaluations with a reflexive approach concerning the instrument goals or the thought behind the instruments.

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 (Vedung, 2009). Involvement of stakeholders will help gear the evaluation towards issues of importance that can lead to essential learning; stakeholders are to be involved in the evaluation process to focus it, to make it timely, to participate in decisions on methods and data collection, to be involved in interpreting findings and to influence value judgements (Alkin, 2013). In

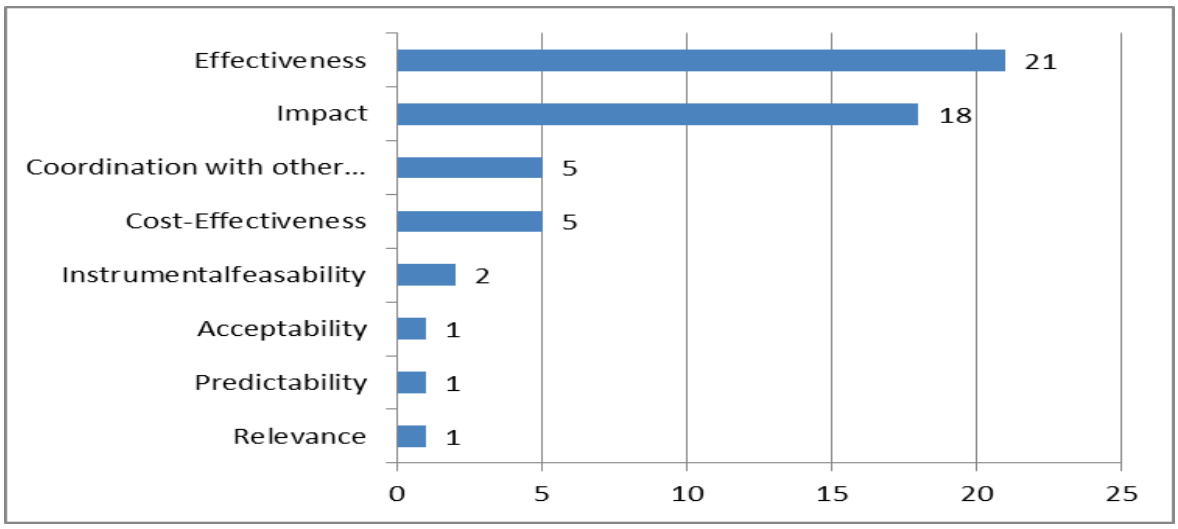

Figure 2. Value criteria used for assessment.

this paper we argue that for a realization of transitions, a reflection on use of evaluations is essential.

According to the reviewed sample, the predominant way for Swedish authorities to conduct evaluations of policy instruments is to commission external consultants. Among the reviewed evaluations the majority were conducted by external consultants (23/30). One evaluation was a collaboration between both authority and consultant (Ekander et al., 2015).

In terms of involvement and use, it is unclear to which extent the authorities have been involved in the design of the evaluation and how the results of the evaluation have been used. For the commissioned evaluations an account of the mission was sometimes stated, however limited to describing the expectations of the evaluation in general terms. The purposes of some evaluations explicitly included the intention of bringing knowledge about the implementation back to authorities and policy makers, but the receiving and making use of this knowledge were however not clarified within the evaluation reports.

The means of involvement in the reviewed evaluations was dominated by interviews; the number of respondents per study ranging from one to 700. Discussion about the scope and size of the group of respondents was not common in the reviewed evaluations, and the involvement was seemingly limited to data collection rather than for facilitation of further use. Evaluations conducted by one consultant did include seminars where a group of selected stakeholders were invited to discuss preliminary results. This may have had a facilitating effect on future use of the results, but was not the primary aim. In general, facilitation of further use or the intended time frame for use were not clearly stated in the reviewed evaluations. However, the majority of the reviewed evlauations were freely available on state webpages which may increase use.

Discussion & conclusion

The array of evaluations of policy instruments implemented in Sweden for increasing energy efficiency in buildings is fairly wide and covers many different types of policy instruments. The critical review of these evaluations shows several strengths in relation to evaluation theories but also weaknesses; more specifically the review illuminate certain areas for improvement regarding practices and possibilities of a deliberate focus on transition efforts and radical changes.

Starting with the framing of Swedish practices for commissioned evaluations, there is a favoring towards multiple methods and multi-criteria analysis, providing a thorough base for evaluation. Chosen criteria are, however, often limited to effectiveness and impacts, whereas criteria for investigating mechanisms aiding or counteracting the outcomes of a policy instrument tend to be underrepresented. Regarding the focus and the system perspectives of designing the evaluations, the review results indicate a rather narrow line of focus which tends to leave potential system aspects, side effects and rebound effects uninvestigated. Such delimitation runs the risk of overlooking effects caused elsewhere than in the target area. The use of counterfactuals for assessing the attribution of the policy instrument is also limited and the robustness of the counterfactual constructions vary from the methods and amount of data that have been used for their synthesis. Triangulation for systemic validation of findings is scarce, even though e.g. multiple methods and data sources commonly are used, they are used to complement each other rather than as tools for comparison.

In all, there is room for improvements in relation to the design, methods and data used to assess the policy instruments. A systemic evaluation strategy, designed at the initial policy design phase, would provide the availability of good and robust data and the potential to provide a robust analysis of changes in the socio-technical system, by including multiple actors, side effects, rebound effects and triangulation. Moreover, the design processes determining purposes and role of evaluations were not clearly outlined in the reviewed evaluations, and should ideally be the product of a dialogue between commissioner and stakeholders representing various groups that share an interest in the policy instrument. As for stakeholder involvement in general, the results show a high share of evaluations that do include actors, however, limiting the involvement to data collection where selected groups are asked to partake in surveys, interviews or workshops. The amount of respondents within an evaluation furthermore varies widely, ranging from one to 700, which consequently affects the robustness of the outcomes,

but is seldom discussed. The involvement of the commissioner and stakeholders in the design process clarifying the role of the evaluation and the expected results will make the client's choice of relevant methods more accurate, and will also provide more precise input to the evaluation of change in the socio-technical system as well as to the evaluation of specific criteria such as impact and effectiveness.

The assessment of the reviewed evaluations' ability to capture radical changes indicates a need for further development of the evaluation practice. The evaluations show little emphasis on evaluating the potential of transitions in terms of visioning, experimentation and learning. Such criteria were not absent in the reviewed evaluations, but not predominant either. Attention to visioning is paid in one evaluation combining ex-post and ex-ante evaluation methods. Attention to experiments is paid in five evaluations, of which two are noting a negative trend in which the evaluated policy instruments have failed to facilitate technological development.

As for learning and the construction of a knowledge base, the focusing on ex-post evaluation in the reviewed sample does indicate that such a platform exists. Learning in terms of concrete suggestions for improvement of the policy instrument did occur, as well as proposals for future evaluation. However, to render the knowledge available and encourage further use, some measures for ensuring the conveying of knowledge between commissioners, evaluators and other users are needed. To clearly state strategies for use within the evaluation may aid in creating a dynamic knowledge exchange and facilitate the reintroducing of findings when designing both future evaluation practices and policy instruments.

Finally, for evaluations to be able to capture radical changes and transformative processes, and notably to be able to fuel such processes, they need to be reflexive. Established goals and norms are upholding the current system, and in order to take big leaps such constructions need to be questioned. This is not uncommon among the reviewed evaluations, but the extent to which reflexivity is actively used needs to increase and evolve into a systematic continuous feature.

For policymakers

For a successful evaluation strategy that is able to create a knowledge base for learning and able to capture radical changes and large-scale system transformations, we identify the following aspects as important:

- Policy makers should engage in discussions with evaluators regarding evaluation focus, design and methodology. A dialogue and clear intentions for further use of the evaluation results may aid in providing relevant outcomes.
- The focus of the evaluation should take a system perspective capable of including both multiple levels and actors, as well as side effects and rebound effects.
- The evaluation design and methodology need to robustly be able to capture relevant perspectives and mechanisms influencing the implementation, the counterfactual should be assessed more thoroughly and triangulation should be used to provide a wide coverage and to assert findings.
- Efforts for experimentation and radical changes should be regarded, as well as a long-term perspective in which upscaling of innovation or accumulation of incremental changes is considered.

Although this review mirrors Swedish evaluation practices, the results are highly relevant also in a broader European context.

References

Alkin, M.C. (Ed.), 2013, Evaluation roots: a wider perspective of theorists' views and influences, 2nd ed. ed. SAGE Publications, Los Angeles.

Ekander, F., Olin, M. & Ågren, R-M., 2015, Final report: the effects of the energy and climate advisory in 2013 to private individuals, companies and organizations (Slutrapport: effekter av energi- och klimatrådgivningen 2013 till privatpersoner, företag och organisationer). Swedish Energy Agency.

Furubo, J.-E., Rist, R.C., Speer, S., 2013, Evaluation and turbulent times: reflections on a discipline in disarray. pp. 1–32. Transaction Publishers, New Brunswick.

Geels, F.W., Schot, J., 2007, Typology of sociotechnical transition pathways. Research Policy 36, 399–417. doi:10.1016/j.respol.2007.01.003

Grontmij, 2009, Evaluation of the Swedish Energy Agency's program for passive houses and low-energy houses 2005–2008 (Utvärdering av Energimyndighetens program för passivhus och lågenergihus 2005–2008). Swedish Energy Agency.

Helmersdotter Eriksson, A., Sandén, P. & Jacobsson, R., 2015, Mapping and analysis of actors' views on an appropriate tightening of the energy requirements of current building codes (Kartläggning och analys avseende aktörers syn på en lämplig skärpning av energikraven i nu varande byggregler). Sweco Strategy.

Huitema, D., Jordan, A., Massey, E., Rayner, T., van Asselt, H., Haug, C., Hildingsson, R., Monni, S., Stripple, J., 2011, The evaluation of climate policy: theory and emerging practice in Europe. Policy Sciences 44, 179–198. doi:10.1007/s11077-011-9125-7.

Jansson, T., Grudin, M. & Henningsson, K., 2016, Final evaluation of the program for buildings with very low energy use (Lågan) (Slututvärdering av Programmet för byggnader med mycket låg energianvändning (Lågan)). Faugert & Co Utvärdering AB.

Kemp, R., 1994, Technology and the transition to environmental sustainability. Futures 26, 1023–1046. doi: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, 175–198. doi:10.1080/09537329808524310.

Kivimaa, P. Kern, F., 2016, Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions Research Policy 45 (1), 205–217.

Loorbach, D., 2010, Transition Management for Sustainable Development: A Prescriptive, Complexity-Based Gover-

nance Framework. Governance 23, 161–183. doi:10.1111/j.1468-0491.2009.01471.x

Lucon, O., Ürge-Vorsatz, D., Zain Ahmed, A., Akbari, H., Bertoldi, P., Cabeza, L.F., Eyre, N., Gadgil, A., Harvey, L.D.D., Jiang, Y., Liphoto, E., Mirasgedis, S., Murakami, S., Parikh, J., Pyke, C., Vilariño, M.V., 2014, 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.

Mickwitz, P., 2002, Effectiveness Evaluation of Environmental Policy: the role of intervention theories, Administrative Studies 21 (4), 77–87.

Mickwitz, P., Neij, L., Benner, M., Johansson, M. & Sandin, S., 2016, A theory based assessment of evaluations intended to inform transitions towards sustainability, working paper, IIIEE, Lund University. (presented at the 2016 European Environmental Evaluators Network (September 2016, Copenhagen) and the European Evaluation Society (EES) international conference (September 2106, Maastricht).

Mundaca L., Neij, L., Markandya A., Hennicke P., and Yan J., 2106, Towards a Green Energy Economy? Assessing policy choices, strategies and transitional pathways, Applied energy, 179, 1283–1292.

National Board of Housing, Building and Planning, 2009 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 (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). National Board of Housing, Building and Planning.

National Board of Housing, Building and Planning, 2009 b, Evaluation of the financial support for installation of energy efficient windows or biofuel devices (Utvärdering av stödet för installation av energieffektiva fönster eller biobränsleanordningar). National Board of Housing, Building and Planning.

National Board of Housing, Building and Planning, 2009 c, 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 (Utvärdering av systemet med energideklarationer. Uppdrag nr 12 Uppföljning av energideklarationer enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A). National Board of Housing, Building and Planning.

National Board of Housing, Building and Planning, 2014, More stringent energy conservation requirements - the accounting of the government's instructions of reviewing and tightening energy regulations in building codes, report 2014:19 (Skärpta energihushållningskrav – redovisning av regeringens uppdrag att se över och skärpa energireglerna i Boverkets byggregler, rapport 2014:19). National Board of Housing, Building and Planning.

Patton, M.Q., 2002, Qualitative research and evaluation methods, 3 ed. ed. Sage Publications, Thousand Oaks, California.

Rotmans, J., Kemp, R., Van Asselt, M., 2001, More evolution than revolution: transition management in public policy. Foresight, Vol 3. Iss 1, 15–31.

Samakovlis, E. & Vredin Johansson M., 2007, Evaluation of the cost-effectiveness of climate investment programs (En utvärdering av kostnadseffektiviteten i klimatinvesteringsprogrammen). National Institute of Economic Research.

Sweco, 2014 a, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.6 – Energikartläggningsstöd). Sweco Energuide AB.

Sweco, 2014 b, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.8 – Energi- och Klimatrådgivning). Sweco Energuide AB.

Sweco, 2014 c, 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 – Network programs for driving development (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.18 – Beställargrupper formerade för att driva utvecklingen). Sweco Energuide AB.

Sweco, 2014 d, 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 (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). Sweco Energuide AB.

Sweco, 2014 e, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering –

en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.20 – Projektstöd för marknadsintroduktion och teknikupphandling). Sweco Energuide AB.

Sweco, 2014 f, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.21 – Energiaktiv.se). Sweco Energuide AB.

Sweco, 2014 g, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.22 – Byggnaders energiprestanda – energideklarationer). Sweco Energuide AB.

Swedish Energy Agency, 2013, Financial support for energy audits – A socio-economic evaluation (Energikartläggningscheckar – En samhällsekonomisk utvärdering). Swedish Energy Agency.

Vedung, E., 2009, Public policy and program evaluation (4th ed.). Transaction Publishers, New Brunswick.

Acknowledgements

We are very grateful to the Swedish Energy Agency (project "Transition governance") for their financial support.

Appendix A – Framework for systemic 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 un-intended) 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 analyzed? – 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; lock-in effects)?; Is the evaluation considering learning (knowledge and skills).

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 judgment the conclusions are based on?

C. USE OF EVALUATIONS

C.1. Has 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 B – Full list of evaluations in review sample

Ekander, F., Olin, M. & Ågren, R-M., 2015, Final report: the effects of the energy and climate advisory in 2013 to private individuals, companies and organizations (Slutrapport: effekter av energi- och klimatrådgivningen 2013 till privatpersoner, företag och organisationer). Swedish Energy Agency.

Grontmij, 2007, Evaluation of the support for implementation of energy audits within OFFROT (Utvärdering av stöd för genomförande av energikartläggningar inom OFFROT). Swedish Energy Agency.

Grontmij, 2009, Evaluation of the Swedish Energy Agency's program for passive houses and low-energy houses 2005–2008 (Utvärdering av Energimyndighetens program för passivhus och lågenergihus 2005–2008). Swedish Energy Agency.

Helmersdotter Eriksson, A., Sandén, P. & Jacobsson, R., 2015, Mapping and analysis of actors' views on an appropriate tightening of the energy requirements of current building codes (Kartläggning och analys avseende aktörers syn på en lämplig skärpning av energikraven i nu varande byggregler). Sweco Strategy.

Holm, D. & Martinsson, F., 2012 a, Evaluation of external projects – Construction and Industry – For the Swedish Energy Agency. 5 – Stoppsladd – Stage 2 (Utvärdering av externa projekt – Bygg och Industri – För Energimyndigheten. 5 – Stoppsladd – Etapp 2). IVL, Swedish Environmental Research Institute.

Holm, D. & Martinsson, F., 2012 b, Evaluation of external projects – Construction and Industry – For the Swedish Energy Agency. 8 – BeAware (Utvärdering av externa projekt – Bygg och Industri – För Energimyndigheten. 8 – BeAware). IVL, Swedish Environmental Research Institute.

Holm, D. & Martinsson, F., 2012 c, Evaluation of external projects – Construction and Industry – For the Swedish Energy Agency. 10 – Energy efficiency and conservation of our cultural heritage (Utvärdering av externa projekt – Bygg och Industri – För Energimyndigheten. 10 – En-

ergieffektivisering och bevarande i vårt kulturarv). IVL, Swedish Environmental Research Institute.

Holm, D. & Martinsson, F., 2012 c, Evaluation of external projects – Construction and Industry – For the Swedish Energy Agency. 12 – Radical energy efficiency and renovation of the Million Programme (Utvärdering av externa projekt – Bygg och Industri – För Energimyndigheten. 12 – Radikal energieffektivisering och upprustning av miljonprogrammet). IVL, Swedish Environmental Research Institute.

Holm, D. & Martinsson, F., 2012 c, Evaluation of external projects – Construction and Industry – For the Swedish Energy Agency. 13 – Herrgårdsängen (Utvärdering av externa projekt – Bygg och Industri – För Energimyndigheten. 13 – Herrgårdsängen). IVL, Swedish Environmental Research Institute.

Jansson, T., Enberg, J., Stålfors, S., Terrell, M. & Karlström, S., 2014, Evaluation of HyLok, state tenants for energy efficiency in premises (Utvärdering av HyLok, statliga Hyresgäster för energieffektivisering av Lokaler). Faugert & Co Utvärdering AB.

Jansson, T., Grudin, M. & Henningsson, K., 2016, Final evaluation of the program for buildings with very low energy use (Lågan) (Slututvärdering av Programmet för byggnader med mycket låg energianvändning (Lågan)). Faugert & Co Utvärdering AB.

Jansson, T. & Terrell, M., 2012, Mid-term evaluation of HyLok, state tenants for energy efficiency in premises (Halvtidsutvärdering av HyLok, statliga Hyresgäster för energieffektivisering av Lokaler). Faugert & Co Utvärdering AB.

Jansson, T. & Terrell, M., 2013, Mid-term evaluation of the program for buildings with very low energy use – LÅGAN (Halvtidsutvärdering av program för byggnader med mycket låg energianvändning – LÅGAN). Faugert & Co Utvärdering AB.

National Board of Housing, Building and Planning, 2009 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 (Utformningen reducerade effekterna – Boverkets utvärdering av OFFrotstödet. Redovisning av stöd till investeringar i energieffektiversering och konvertering till förnybara energikällor i lokaler som används för offentlig verksamhet 2005:205, enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A). National Board of Housing, Building and Planning.

National Board of Housing, Building and Planning, 2009 b, Evaluation of the financial support for installation of energy efficient windows or biofuel devices (Utvärdering av stödet för installation av energieffektiva fönster eller biobränsleanordningar). National Board of Housing, Building and Planning.

National Board of Housing, Building and Planning, 2009 c, 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 (Utvärdering av systemet med energideklarationer. Uppdrag nr 12 Uppföljning av energideklarationer enligt regleringsbrev för budgetåret 2009 avseende Boverket. M2008/4791/A). National Board of Housing, Building and Planning.

National Board of Housing, Building and Planning, 2014, More stringent energy conservation requirements - the accounting of the government's instructions of reviewing and tightening energy regulations in building codes, report 2014:19 (Skärpta energihushållningskrav – redovisning av regeringens uppdrag att se över och skärpa energireglerna i Boverkets byggregler, rapport 2014:19). National Board of Housing, Building and Planning.

Samakovlis, E. & Vredin Johansson M., 2007, Evaluation of the cost-effectiveness of climate investment programs (En utvärdering av kostnadseffektiviteten i klimatinvesteringsprogrammen). National Institute of Economic Research.

Stern, P. Enberg, J., Fridholm, T. & Ärenman, 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 och Co Utvärdering AB.

Sweco, 2014 a, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.6 – Energikartläggningsstöd). Sweco Energuide AB.

Sweco, 2014 b, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.8 – Energi- och Klimatrådgivning). Sweco Energuide AB.

Sweco, 2014 c, 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 – Network programs for driving development (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.18 – Beställargrupper formerade för att driva utvecklingen). Sweco Energuide AB.

Sweco, 2014 d, 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 (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). Sweco Energuide AB.

Sweco, 2014 e, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.20 – Projektstöd för marknadsintroduktion och teknikupphandling). Sweco Energuide AB.

Sweco, 2014 f, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.21 – Energiaktiv.se). Sweco Energuide AB.

Sweco, 2014 g, 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 (Styrmedels förutsättningar att styra mot ökad energieffektivisering – en utvärdering av 24 styrmedel – En rapport till Näringsdepartementet. 2.22 – Byggnaders energiprestanda – energideklarationer). Sweco Energuide AB.

Swedish Energy Agency, 2011, Long-term follow-up of the energy and climate advisory in 2008 and 2009, ER 2011:20 (Långtidsuppföljning av energi- och klimatrådgivningen 2008 och 2009, ER 2011:20). Swedish Energy Agency.

Swedish Energy Agency, 2013 a, Effects of the energy and climate advisory in 2011 (Effekter av energi- och klimatrådgivningen 2011). Swedish Energy Agency.

Swedish Energy Agency, 2013 b, Effects of the energy and climate advisory in 2012 to private individuals, companies and organizations (Effekter av energi- och klimatrådgivningen 2012 till privatpersoner, företag och organisationer). Swedish Energy Agency.

Swedish Energy Agency, 2013 c, Financial support for energy audits – A socio-economic evaluation (Energikartläggningscheckar – En samhällsekonomisk utvärdering). Swedish Energy Agency.