ENERGY EVALUATION EUROPE
2020 Conference
Here East, Olympic Park, London
Monday 29 June to Wednesday 1 July 2020
Formerly IEPPEC
Accelerating the energy transition for all: Evaluation's role in effective policy making

The dilemma of evaluating impacts of informational measures

Katja Schumacher, Öko-Institute, Germany, k.schumacher@oeko.de;
Martin Steinestel and Reinhard Loch, Consumer Association of North Rhine-Westphalia, Germany;
Dominik Jessing, ifeu, Germany

ABSTRACT

Informational measures are an integral part of an energy efficiency strategy targeting broad groups to individual checks on site. While the importance of informational measures is undisputed, their actual impact is often unknown. Evaluators especially face the challenge to assess their individual impacts.

For smaller scale measures a theory-based bottom-up evaluation approach might be used. It places emphasis on the impact chain of a measure or activity. For each part of the chain, evaluators aim to gather information. But the information on whether action follows and is caused by the measures is very difficult to obtain.

The lack of contact information prevents a survey based evaluation. Alternatively, insights from evaluations in similar fields might be used. If even this is not possible, the use of default values (empirical approach or expert judgement) might provide a rough estimate.

In this paper, we show the range of impact estimates that result from applying different methodologies. We develop a multi-stage model to evaluate informational measures. The multi-stage model breaks the evaluation task down in individual steps. It is very transparent, well-structured and each step can be replicated and recalculated. We exemplify this by looking at two specific informational measures (public informational lectures and DIY-Youtube clips on energy efficiency measures). The parameterisation remains the greatest challenge. Our paper assesses the pros and cons and concludes with recommendations for policy makers and further research.

Introduction

Informational measures are an integral part of a strategy to reduce energy consumption in households[1]. They might range from broad campaigns, over internet-based information to energy checks on site at consumers' homes or at offices or production sites. Informational measures aim to raise awareness, sensitize users and stimulate changes in behavioural, user routines or purchase/investment decisions towards low energy/climate friendly action. While their importance for energy conservation or climate strategies is undisputed, their actual impact is often unknown. Informational measures often complement other measures and are part of comprehensive policy packages. Nevertheless, evaluators often face the challenge to assess the individual impact of informational measures. Funders or regulators provide funding for these measures, are required to prove cost-effectiveness to auditing offices and ask evaluators to perform ex-post evaluations, most often to also compare them to ex-ante estimations.

To provide quantitative impacts of informational measures, evaluation studies tend to follow a top-down approach in comparing energy consumption indicators before and after the informational measure or in relation to a control group that was not affected by the measure. They rely on available statistical data on a

[1] In this paper, we understand measures in a sense of activities resulting from policy instruments, e.g. from climate programmes or energy efficiency codes. We focus on households.

disaggregation level that matches the object/subjects targeted by the measures (i.e. number of electric vehicles in a region, heating energy consumption) and correct for influencing factors outside their measure. While this might work for large scale informational programs, it is not suitable for most smaller scale informational measures (such as for example for information via DIY video clips that explain how to insulate hot water pipes).

For these measures, the European Commission recommends in their standard (DIN EN 16212, 2012) to follow a bottom-up evaluation approach. The approach is theory based (Leeuw, 2003) and places emphasis on the impact chain of a measure or activity. It starts bottom-up from a) the number of people, firms, business etc. within the target group that are reached by activities within the informational measure, b) considers how many of these people, firms, business take action in line with the information provided, c) whether the action is caused by the informational measure, d) what kind of action they actually take (modal split, energy efficiency investment in appliances or technology, change in nutrition etc.), and e) how long the change in action might last (lifetime of an appliance or just until the rainy season starts). For each of these steps, evaluators aim to gather information. While the number of participants might be counted and thus known, the assessment of whether action follows and is caused by the measures is more difficult to know from the target group.

Surveys provide a good way to close the information gap (Andor and Fels, 2018; Delmas et al., 2013; Hafner et al., 2019; Rand Europe, 2012). But often contact information is not available to conduct a survey, in particular if contacts with the target group are rather short. In other cases, if contact information is available the response rate is low, in particular if the time span from providing the informational measure to conducting the survey is long. A longer time span for a follow up is important, however, to give the target group time to induce changes or take investment decisions. If surveys are not possible, the European Commission standard recommends using insights from other evaluation studies in similar fields to arrive at impact estimations. If even this is not possible, e.g. because no comparable study has been done, the use of default values might provide a last resort. Such values can be deducted from previous surveys, the literature or expert judgement and provide only rough estimates (Dahlbom et al., 2009).

In this paper, we use two examples for informational measures (public lectures by energy consultants on specific energy topics and Do-it-yourself Youtube video clips on insulating hot water pipes) to show the range of impact estimates from different methodologies. The informational measures are part of an energy counselling project (ENERGIE2020) of the German Consumer Association of North Rhine-Westphalia (Verbraucherzentrale NRW or VZ NRW in the following). The paper starts out with introducing the methodological approaches, followed by a literature review on evaluation of informational measures. We then match our approach to the multi-stage evaluation model from the ENERGIE2020 project and apply it to the two example measures, assess the pros and cons and conclude with recommendations for policy makers and further research.

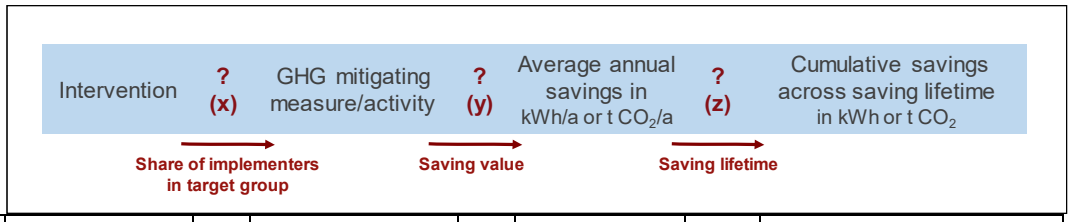
Approaches and challenges to evaluation of informational measures

The investigation of informational or "soft" measures, such as short consultations and lectures and DIY-videos, can best be based on the approach of theory-based evaluation (Leeuw, 2003). For this purpose, impact models are developed which describe the mechanisms by logically linking interventions (project activities) with the impacts and thereby allowing statements on the evaluability at the various levels of action. The core element of each impact model are measure-specific impact chains that establish a logical connection between measures and effects. These impact chains provide the basis for quantifying model variables and thus to derive a quantification of possible impacts. While the accuracy of the impact varies depends on data availability and quality and can be represented as a range, the logical connection of the effect chain must be unambiguous.

The approach corresponds to the recommendations of DIN EN 16212 Energy efficiency and energy saving calculation - top-down and bottom-up methods (DIN EN 16212, 2012). The European standard provides a general approach for ex-ante and ex-post evaluations of energy or greenhouse gas emissions savings. The general approach is applicable to energy savings of buildings, vehicles, appliances, industrial processes, etc. However,

the implementation of the EU standard depends on the availability of suitable data. Reference values or default values that can be used alternatively are not given.

The following illustrates a typical impact chain for the assessment of energy or GHG savings.



Example 1	Public lectures by energy consultants on specific energy topics		Homeowner investment or behavioural change		Energy efficient lighting, refurbishment of heating etc.		lifetime of technology or behavioural change
Example 2	Do-it-yourself Youtube video clips on insulating hot water pipes		Homeowner insulate hot water pipes		reduction of energy loss		lifetime of heating pipes

Figure 1. Typical impact chain for informational measures to save GHG in households

To assess energy or GHG savings from measures/interventions, three main information gaps (marked with the letters x, y and z in Figure 1) arise, which can be addressed with the help of adequate data, standardised assumptions and expert estimates. The implementation probability of the intervention (x) describes the relation between project activities and the effects on the target group (here mainly energy savings or GHG-reducing actions). The savings value of the energy- or GHG-reducing action (y) shows the savings effects (e.g. kg CO2eq.). Finally, in order to correctly calculate the savings effects, the "savings lifetime " (z) must be recorded.

For gaps y and z, data from other evaluations are already partially available. The most challenging task is to estimate the share of implementers in the target group, gap x. With this gap, it should be noted that in real application, several interventions can often have a simultaneous effect on the target group. Furthermore, "soft" measures usually do not lead directly to energy- or GHG-reducing actions, but to changes in the target group, which can have an influence on their actions, for example by providing knowledge, increasing motivation or open-mindedness, by improving the framework conditions of the target group. All these aspects can play an important role in motivating actors to take more actions in the short, medium or long term. These interrelationships must be included or separated when estimating the share of implementers (see Figure 2) in order to determine the impact of the programme to be evaluated. However, this requires elaborate evaluation methods, such as control group designs, in which groups comparable with the target group are considered, which may have benefitted from other interventions, but were not affected by the intervention to be evaluated.

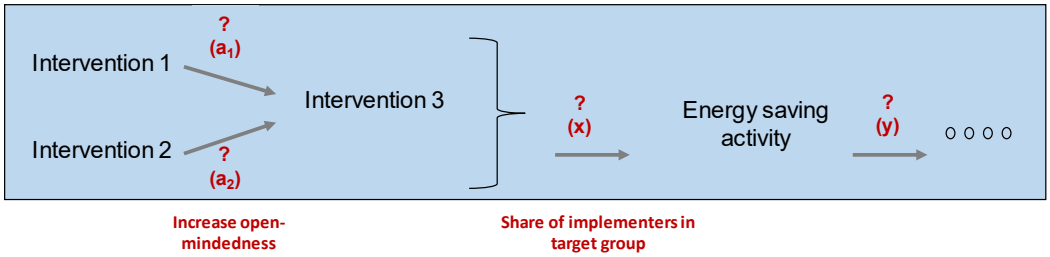


Figure 2. GHG impact chain for multi-causalities

The preferred way to fill the information gaps in the impact chain is to actually measures energy consumption before and after the intervention. In many cases, however, the effort involved is disproportionate. Furthermore, the target group is not known in detail, so that measurements cannot be carried out or measured values cannot be inquired, and a comparable control group cannot be found. Control group approaches also involve a great deal of effort (and high costs), as consumers with comparable characteristics who have not participated in the informational intervention must be identified. Contact details are usually not known and information can only be obtained through extensive consumer surveys. Data protection provides an additional challenge. The cost of applying this approach to informational measures is most often disparate in relation to the total cost of the project.

Alternatively, surveys of the participants (before and after and sometime afterwards) offer good information for the evaluation of the share of implementers and the attributability to the measure. Surveys offer the opportunity to find out whether the information provided was helpful, whether it led to actions or investment decisions (or whether actions would have taken place even without the advice) and whether further measures were taken during implementation. However, even though participants are usually counted when providing an informational measure (e.g. public lecture), it is not always possible to collect contact data, so that surveys may not be possible (or not possible without considerable additional effort). Furthermore, participants often find themselves unable to attribute their actions to a cause (i.e. they are unable to provide a weight or number for the question “How much did the measure contribute to your action?”). Finally, especially in the case of "soft measures", where the target group is sometimes only reached with low intensity (e.g. short consultation at a trade fair compared to at-home-energy consultation), data collection in the target group is always a challenge. Here, limited recall of the soft measures or lack of willingness of the target group and protection of personal data are strong obstacles.

Literature values, expert estimates and empirical values from other evaluations, if available, must therefore be supplemented. In some cases, however, quantification is not scientifically feasible and therefore cannot be carried out.

Literature Review

In a first step, we investigated the existing literature on the evaluation/assessment/effectiveness of informational interventions. We included publications in academic journals, grey literature and project reports in our search. The literature can be subdivided into a) meta-studies, which provide a summary overview of studies on a topic, b) empirical studies on specific issues, and c) guidelines/instructions on evaluation and effectiveness assessment.

It can be summarised that the literature on the attributability of interventions to energy efficiency/savings effects is rather scarce and most often not backed up with empirical values. Furthermore, no study known to us deals specifically with public lectures, short consultations or DIY Youtube video clips. "Energy tips" or "energy audits" come closest to these topics.

Most studies evaluate the effect of informational interventions by measuring or calculating the reduction of energy consumption in the target group or by conducting interviews with the target group. In a next step, they

examine whether these reductions can be linked to the intervention (attributability). For this purpose, they use control group approaches, counterfactual scenarios or statistical regression analyses. Experimental designs are also used; in these cases, the set-up is designed differently for different groups and it is analysed to what extent one group saves more energy in the experiment than the other group.

Even though these approaches are called bottom-up approaches, they usually start the impact analysis "from the back", i.e. starting from the achieved savings and consecutively trying to determine what contribution the intervention had to the savings. This approach requires access to the target group in order to measure, calculate or inquire about energy savings afterwards. This means that the target group is fully known, and the contact details of the target group are available, so that information can still be obtained from the target group after a certain period of time. Also, the aim of the intervention is usually homogeneous and completely known.

This approach is not feasible for interventions such as public lectures, short consultations or DIY video clips, as there is only a brief (and sometimes indirect) contact with the target group. Furthermore, the contact may pursue several goals with its primary aim to convey information and promote the open-mindedness and empowerment of consumers and therefore indirectly contribute to improving energy efficiency. A value for the reduction in energy consumption cannot be inquired from the target group or measured or calculated. The approach pursued so far is therefore to explore to what extent consumers become active on the basis of the lectures, short consultations or video clips, what further information they needed and what type of activity they carry out after receiving the informational intervention. It would be desirable to find out (inquire) what impulse/contribution the lecture or short consultation gave to these activities.

The following three evaluations from the literature provide estimates and values for some of the factors of the impact model and are thus discussed in more detail.

Highlighting three studies

Methodological manual for the evaluation of the National Climate Initiative

The evaluation methodology of the NCI (impact chain analysis with a bottom-up approach - division into gaps "x" effectiveness, "y" savings value, "z" lifetime of savings) is presented in an evaluation manual (Schumacher et al., 2019). Guideline values for the effectiveness of informative interventions are given, which can provide a point of reference if measurement, questioning or surveys are not possible. The guideline values refer to the percentage of implementers in the target group, i.e. they describe the percentage of people who actually implement a desired climate-friendly action after taking part in the informational intervention. In the NCI evaluation, the guideline values differ by type and intensity of intervention. A broad campaign intervention is associated with a lower probability of implementation than a bilateral and interactive contact with specific advice. The NCI methodology does not provide guiding values for the attributability of intervention to action or impact. It should be noted that the guiding values in Table 1 provide probabilities for addressing changes in user routines or decisions on smaller investments. For larger scale investment decisions, the numbers do not apply. Projects within the NCI usually apply only one type of intervention per homeowner.

Table 1. Guidelines for likelihood of action (share of implementers in the target group)

Type of intervention	Intensity of intervention	Guideline Values for effectiveness
Broad campaigns	Simple contact Intensive contact	Max. 2% 2%-5%
Specific advice	Intensive contact (stationary) Very intensive contact (on-site)	5%-10% 10%-15%
Decision making help	Specific information for individual questions in decision making phase	8%-12%

Source: NCI evaluation methodology (Schumacher et al., 2019)

Changing energy behaviour - Guidelines for behavioural change programs

A study by Dahlbom et al. (2009) draws experience from 41 case studies of behavioural change programmes in 10 EU countries and provides estimations on achieved energy savings. For example, it shows that specific information programmes lead to energy savings of 3%. When tailored to specific target groups, specific information measures even lead to reductions of 16%; when combined with financial incentives they lead to even more (see Table 2). Dahlbom et al. (2009) compare energy savings (in %) before and after the intervention. However, it is not clear from the study whether the values have been adjusted for interaction or attributability, i.e. it cannot be said with certainty whether other influencing factors have also contributed to the reductions. If attribution is not taken into account, the values should rather be regarded as an upper limit of the estimate. In light of the interventions that are in focus of our paper (public consultations, DIY videos), only "information dissemination" and "specific information" from their list are comparable. "Tailor-made information" has a certain comparability with on-site consultations, since information and recommendations are specifically tailored to the respective household.

Table 2. Average energy savings of different interventions

Intervention	Likely Saving*
Contracts and reward	6%
Financial incentives	3%
Fin. Incentives + information	5%
Information dissemination	1-2%
Specific information	3%
Financial support	9%
Tailor-made info	16%
Tailor-made + fin. Support	16+%
Ecoteams (high impact amongst small population)	15%

Source: Dahlbom et al., 2009, p. 22. *Whether the savings rates reflect lifetime savings or savings in the first year of intervention is not explained in the study.

Achieving energy efficiency through behaviour change: what does it take?

Similar to the study by Dahlbom et al. (2009), the study by the European Environment Agency (EEA, 2013) brings together results from the literature on energy savings from different types of intervention. They emphasize that the figures should be treated with caution, as the savings depend strongly on the design of the project and the evaluation, e.g. the size of the group considered, the duration of the project, the demographic composition, the possibility of including a control group. Nevertheless, they give an indication of the savings that can be achieved with the measures under consideration. In most cases, annual energy bills before the measure are compared with energy bills after the measure. It remains unclear whether only one year is considered or whether the savings are calculated over a period of years. The interventions are mostly more intensive than the ones in focus in our paper. Nevertheless, they also show a broad range of savings.

Table 3 Potential energy savings from interventions aimed at behavioral change

Intervention	Range of energy savings
Feedback	5–15 %
Direct feedback (including smart meters)	5–15 %
Indirect feedback (e.g. enhanced billing)	2–10 %
Feedback and target setting	5–15 %
Energy audits	5–20 %
Community-based initiatives	5–20 %
Combination interventions (of more than one)	5–20 %

Source: EEA (2013). *Whether the savings rates reflect lifetime savings or savings in the first year of intervention is not explained in the study.

Main insights from the literature

In summary, the following insights can be taken from the literature research:

- Most studies refer to a specific intervention and determine savings by measuring before and after energy consumption.
- Only a few studies provide quantitative data on energy savings of different types of informational interventions. Most studies refer to investment interventions or give qualitative assessments.
- Some studies list concrete values for energy savings. However, the different factors of the impact model are not separated (implementation factor >< attribution factor). It can be assumed that, depending on the evaluation methodology and design, the attribution factor is not taken into account and that time related factors such as project duration and savings lifetimes are handled inconsistently.
- Savings are higher the more specific, intensive and tailored an intervention is. Broad campaigns have a significantly lower impact per addressee. The effects range from 2% (broad campaign, single contact) to 15% (specific advice, decision-making knowledge, intervention tailored to target group).
- For some interventions (e.g. in the area of education, training programs, but also networking), it is not possible to find default values for savings in the literature. Therefore, a scientifically sound and justifiable approach to determining energy or emissions savings for these measures seems doubtful.

The ENERGIE2020 project and it multi-stage evaluation model

The ENERGIE2020 project[2] is performed by the Consumer Association of North Rhine-Westphalia (Verbraucherzentrale NRW - following: VZ NRW), a nonprofit organization. The project informs private households on energy efficiency improvement, implementation of renewables and smart technologies. Private consumers are targeted by lectures, trade fair contacts, website and social media tools and individual consultations for retrofitting houses in an office or on-site at home.

An evaluation objective is to assess the impact of lectures, short advisory contacts on trade fairs and social media tools. Many advisory services (personal consultations and on-site consultations) have already been evaluated in previous projects (Jessing et al., 2016). As shown in the literature review, for informational or "soft" interventions such as lectures, trade fair contacts, or for campaigns and public relations in general, it is a challenge to estimate an order of magnitude of achievable emission reductions.

The multi-stage evaluation model of the ENERGIE2020 project

For this reason, VZ NRW has designed a multi-stage evaluation model that defines several steps and factors along the impact chain. The model tries to estimate the size of net effects as impacts.

The first step of the impact chain, the **output** (or number or proportion of those reached in the target group) is assessed using different methods - depending on the approach and the available savings figures.

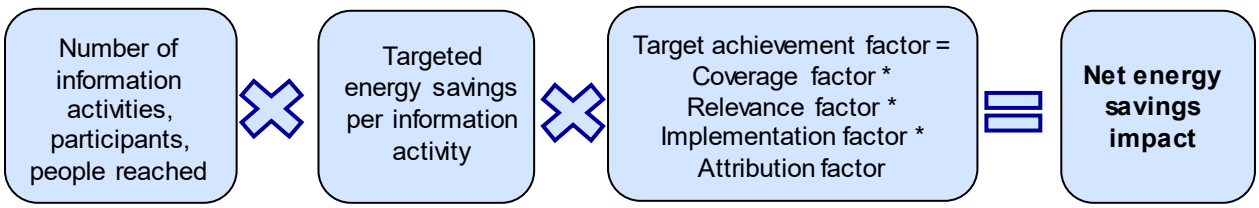


Figure 3. Energy savings impact chain of multi-stage evaluation model by VZ NRW

To assess the next step, the **outcome** of an informational intervention, the model first considers the target savings. The target savings from an informational intervention can be estimated, calculated or specified. Possible indicators are a comparison of before and after energy consumption, typical savings values of a measure (empirical values), percentage savings of energy consumption, average values etc. The value is assumed to be an annual and average real final energy saving of an addressee (gross saving = target effect TE). The model is static and calculates only first-year-savings and first-year-impacts. To ensure transparency and comparability of the measures and their effects, the model does not take account of lifetimes and project periods (compare Voswinkel, 2018).

The impact model distinguishes between the gross impact, which can be determined by the group size and the target effect (implementation), and the net impact, which takes only the share of energy savings actually induced by the measure into account (attribution). The net effect therefore represents the actual project impact. To determine this net impact (net effect NE), a target achievement factor (TAV) is introduced, so that the net effect equals NE = TE * TAF

The target achievement factor consists of several elements that are derived and multiplied along the causal chain: TAF = CF * RF * IF *AF

[2] The ENERGIE2020 project is co-funded by the EU EDRF (European Regional Development Fund) and the state of North Rhine-Westphalia.

- Coverage factor CF - describes the proportion of the (total) target group (e.g. households, owners of single-family houses) reached by the measure. Example: A newspaper advertisement on the subject of heating replacement only reaches those who read the newspaper at all.
- Relevance factor RF - records for which share the intervention is relevant at all and is understood and accepted. Example: A newspaper advertisement on the subject of a heating replacement is only relevant for those who own heating systems and whose heating systems are due a renewal or a change of energy source in the foreseeable future (e.g. roughly every twentieth to tenth heating system owner).
- Implementation factor IF - describes what proportion of the intended target effect is implemented. More specifically, it describes how many people in the target group implement the energy-saving or GHG-reducing action for which they were advised. Example: How many participants of a lecture on the topic of heating replacement subsequently replace their heating?
- Attribution factor AF - takes into account in how far the replacement can be attributed to the intervention. Specifically, it determines whether - and if so, to what extent - the implementers carry out their action based on the intervention. Example: To what extent did the public lecture on heating replacement contribute to replacing participant's heating? This means, here we correct for other influencing factors.

Intervention-Example 1: Public lecture on energy renovation of buildings

A public lecture/talk has the intervention type of a consultation - here a group consultation. In contrast to an individual consultation on site at the building, a lecture usually covers a smaller range of possible measures and has a lower influence on planned or additional measures. The target group (private owners of detached and semi-detached houses) and the number of people reached are known. For a lecture for private homeowners on energy-related building modernization, we do not need to worry about the coverage and relevance factor as they can be set at 100% as a first approximation assuming that only people attend for whom the topic is relevant. To get an approximation for the implementation and attribution factor, we conducted a three-stage survey with attendees. Participants were asked to fill in a survey before the lecture (1st stage) and right after the lecture (2nd stage). Furthermore, they were asked again after 3 to 5 months. The period was chosen to ensure that participants could still recall the talk but also had enough time to initiate or more intensively contemplate changes. The survey covered questions on how participants knew about the lecture, why they participated, how they liked the lecture, how they describe their knowledge on the topic before and after the lecture, whether they need additional information and which kind of information (after the lecture), whether and how they were planning on following up with activities (question before the lecture and immediately after the lecture) as well as whether and how they have indeed initiated follow-up implementation activities (3-5 month after the lecture) and whether the lecture had an impact on these activities (attribution factor). The survey included closed and open questions. About 50% of the participants who agreed to provide contact details at the time of the talk responded to the survey (n=155). (Jessing and Schumacher, 2019)

The survey revealed that a substantially higher number of implementation activities (up to 18%) were planned/initiated after the lecture than were planned before the lecture. The implementation factor IF can thus be assumed to be in the order of 10%-18%. As to the attribution effect, on average about 5% of the participants who initiated follow-up activities said that the lecture gave the decisive information and impulse, 40% said it was very helpful and another 40% said it was of supplementary help. Weighting these categories with 100%, 50% and 25%, we derive an attribution factor AF between 25% and 36%. This results in a target achievement factor TAF of between 3% and 6.5% for lectures on building modernization within the project Energy 2020. This means that about every twentieth participant in a lecture carries out an (additional) modernisation on the basis of attending

the lecture. From this, the target value and an assumed energy-emission mix can be used to determine both gross and net reduction effects for final energy and GHG.

Intervention-Example 2: Do-it-yourself-Video on heating pipe insulation

A DIY-youtube video is an informational and explanatory video that provides specific instructions/manuals on insulation and aims at implementation. It is considered an informational intervention, specifically a consultation with one-way interaction as people cannot ask questions or engage in discussions compared to lectures or on-site consultation. Nevertheless, implementation rates can be expected to be relatively high because the information is very specific and leads through each implementation step. The impact of the video depends on whether the viewer was already highly motivated and was "only" looking for specific instructions (high likelihood of implementation) or whether the viewer is new to the topic and was pointed to it via a campaign or a reference from VZ NRW (lower likelihood of implementation). An indication for the coverage factor CF can be the ratio of the views of relevant length (reflecting a stable level of audience retention during the time of the main content of the clip; realized by analysis of the audience retention-diagram; not just the first few seconds) to the total number of views.

Figure 4 shows an example for a DIY video clip of insulating heating pipes by the VZ NRW[3]. It shows a significant drop of viewers after about 27 seconds. We derive that approx. 70% (of a total of 16,600 viewers in the year 2019) watched the entire video (with the exception of the last few closing seconds).

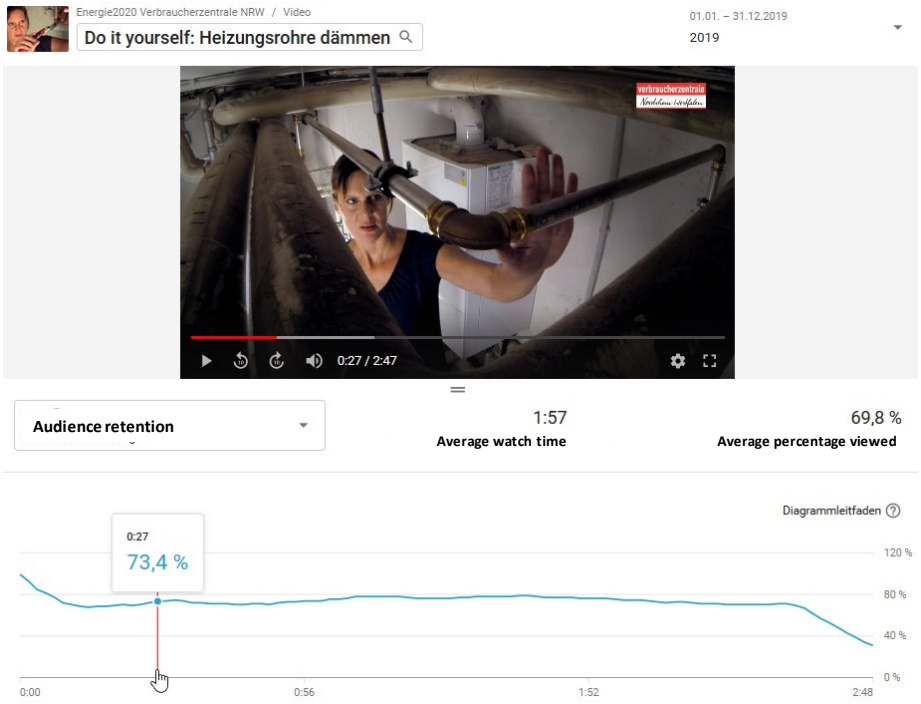


Figure 4 Ratio of viewers over the length of the 2:48 minute video (in 2019)

A measurement of the other factors is not available and must be estimated. The number of views of the video might include other interested parties who are not planning a specific implementation or persons who watch the video repeatedly. The relevance factor (RF) is thus exogenously set to a conservative factor of 50%-60%. The implementation factor IF can be assumed to be in the range of 70%-80% due to the implementation proximity of the topic. People who are not interested in the specific topic are considered unlikely to watch a 3-

[3] https://www.youtube.com/watch?v=mJhMb3SUh0s

minute video on heating pipe insulation or other very specific activities. The attribution factor is assumed to be in the range of 20% to 30%, reflecting the impact of the video on actually carrying out the activity. Multiplying these factors, we receive a target achievement factor (TAF) of 5% to 10% for the DIY videos on heating pipe insulation by VZ NRW. This means that about every twentieth to tenth viewer of a VZ NRW video therefore carries out (additional) heating pipe insulation.

In the transdisciplinary discourse, this assessment approach needs to be repeatedly re-examined and adapted to new findings. Some caveats remain: For example, it is not possible to find out how many viewers watch the video more than one time or how many viewers take action because of the video. Ideally, the video viewers would answer a few questions while watching the video, e.g. a) whether the video is relevant for them (after about 30 seconds – relevance factor), b) whether they plan to implement action (and insulate their heating pipes – implementation factor) and c) whether the video gives a decisive impulse and is helpful for implementation (attribution factor). It needs to be seen whether viewers would be willing to answer these questions or rather stop the video and whether the answers allow for additional insights.

Reflections on the multi-stage evaluation model

In both examples, the energy savings and emission reduction impacts can be derived from the target achievement factor, the respective target value of the savings and an assumed energy-emissions mix, both gross and net. In the VZ NRW impact model, only first-year savings are estimated. Lifetime savings can be derived but depend on assumptions on a static or dynamic baseline and the number of years the effect lasts. To separate assumptions on lifetime calculations from assumptions needed to derive the target achievement factor, it is recommended to focus on first year savings first.

It needs to be noted that the approach shown above can only provide rough estimates. However, the multi-stage impact model by VZ NRW follows a very transparent, well-structured evaluation approach. Each step can be replicated and recalculated. The approach takes into account the different intervention logics of measures, allows different procedures to be applied and is therefore very suitable for estimating a range of values typical for the intervention. The parameterisation represents the greatest challenge. Uncertainties in the parameters multiply across the model and can distort the results. The literature research, however, shows that parameters lie within the range of values derived in other evaluation with different methods. Our model (individual factors within our model) can gradually be improved and uncertainties be reduced once further evaluations and experience become available and knowledge is shared.

Conclusions and recommendations

The multi-stage impact model introduced in this paper has proven very useful for evaluations of informational interventions. It is very transparent and comprehensible and reflects the qualitative cause-impact chain. The model can be flexibly applied and can thus be tailored precisely to the intervention to be considered. Compared to evaluations in the literature, the impact chain is much more detailed and separates different factors that influence the impacts of the interventions. This ensures that not only the implementation in the target group is assessed, but also that the specific influence of the project or measures is separated from other influencing factors (attributability). Separating different factors is the strength of the approach, but also the challenge for parameterisation. Uncertainties regarding the parameterisation differ depending on the way data is collected as well as on the quality of data. They can range from "no uncertainty" (measured data) to "high uncertainty" (rough estimate/assumptions). Furthermore, the uncertainties of the parameterisation of individual factors multiply over the chain of effects and influence the savings effects.

The literature survey reveals that quantifying savings of informational measures turns out to be a challenge. The literature overview shows a range of possible savings that can form a framework for the activities in the ENERGIE2020 project. Often though, literature values are based on measurements or surveys (before-and-after consideration) and directly consider the energy savings without covering the intermediate stages along the

impact chain. Thus, the exact impact mechanism (whether it was due to high coverage, high relevance or high implementation rates), can often not be understood. The literature research indicates that the attribution factor (also called impact performance factor) is mostly disregarded, which is why the results in the literature must always be considered in the light of their assumptions and the underlying uncertainties.

We conclude that while informational interventions play an important role in stimulating investment decisions or behavioural change their impact is very challenging to evaluate. Further research and empirical evidence will help to refine the evaluation methodology. Our multi-stage model can provide a good basis.

Recommendations for evaluators, funders and policy makers include:

- Stimulate exchange of expertise with evaluators and the scientific community
- Ensure regular review of the parameterisation based on publications in the literature, own surveys or evaluation activities of third parties.
- Encourage/conduct surveys and follow-up surveys to enhance empirical evidence. Approaches and insights from the area of marketing and communication should be taken into account when designing and evaluating survey, in particular if online tools or videos are included.
- Give more emphasis to indicators at the outcome rather than impact level (especially if savings on the impact level cannot be assessed or only under very high uncertainty). Such outcome indicators have been applied for examples in consultations in the education sector and campaigns. Possible outcome indicators that deserve emphasis include: knowledge enhancement, awareness raising, increase of open-mindedness, self-efficacy. In addition, for all interventions general aspects should also be highlighted which are not or barely quantifiable, such as the visibility of project implementers and advisory agency, development of confidence and trust relationships that allow for future action (e.g. consumer advice agencies set agendas and act in the interest of consumers), compatibility, relative advantage etc..

These recommendations will help take this research further and provide a better basis for evaluating informational measures. In addition, publications at conferences and in journals will help to disseminate and further evolve the methodology and its empirical foundation. It should be noted though that the challenges in collecting data as well as the uncertainty about the quality of data will continue to be a challenge and limit the insights that can be gained from evaluation. A trade-off will therefore remain between effort being put into evaluation of informational measures and additional insights gained. Policy makers might thus be inclined to "earmark" funding for informational measures and not to evaluate their impact.

Bibliography

Andor, M. A. and Fels, K. M., 2018, 'Behavioral Economics and Energy Conservation – A Systematic Review of Non-price Interventions and Their Causal Effects', *Ecological Economics, (*148) 178–220.

Dahlbom, B., Greer, H., Egmond, C. and Jonkers, R., 2009, *Changing Energy Behaviour - Guidelines for Behavioural Change Programmes,* Madrid (https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/behave_guidelines_for_behavioural_change_programmes_en.pdf) accessed 6 May 2019.

Delmas, M. A., Fischlein. Miriam and Asensio. Omar I., 2013, 'Information strategies and energy conservation behavior_ A meta-analysis of experimental studies from 1975 to 2012', *Energy Policy, (*61) 729–739.

DIN EN 16212, 2012, *Energieeffizienz- und -einsparberechnung – Top-Down- und Bottom-Up-Methoden.*

EEA, 2013, *Achieving energy efficiency through behaviour change: what does it take?,* No 5 (https://www.eea.europa.eu/publications/achieving-energy-efficiency-through-behaviour) accessed 6 May 2019.

Hafner, R., Elmes, D., Read, D. and White, M. P., 2019, 'Exploring the role of normative, financial and environmental information in promoting uptake of energy efficient technologies', *Journal of Environmental Psychology, (*63) 26–35.

Jessing, D., Mellwig, P., Werle, M., Duscha, M., Krieg, O. and Erren, J., 2016, *Evaluation der „Energieberatungbei Ihnen zu Hause" im Projekt „Private Haushalte in Nordrhein-Westfalen für die Energiewende gewinnen (ENeRWin)"* (https://www.energie2020.nrw/sites/default/files/migration_files/media246228A.pdf) accessed 6 February 2020.

Jessing, D. and Schumacher, K., 2019, *Endbericht Evaluation der Veranstaltungsteilnehmenden und Einschätzung des Wirkmodells im Projekt „ENERGIE 2020",* Verbraucherzentrale Nordrhein-Westfalen (https://www.verbraucherzentrale.nrw/sites/default/files/2020-02/Eval_Energie2020.pdf) accessed 6 February 2020.

Leeuw, F. L., 2003, 'Reconstructing Program Theories: Methods Available and Problems to be Solved', *American Journal of Evaluation, (*24/1) 5–20.

Rand Europe, 2012, *What Works in Changing Energy - What Works in Changing Energy-Using Behaviours in the Home?*

Schumacher, K., Tews, K., di Nucci, R. and Zell-Ziegler, C., 2019, *Feinkonzept zur Evaluierung der Nationalen Klimaschutzinitiative - Methodenhandbuch,* Öko-Institut; FFU Berlin, unpublished/unveröffentlicht.

Voswinkel, F., 2018, *Accounting of energy savings in policy evaluation - How to get at least 8 different (correct!) results from the same data,* Wien.