Using systematic reviews to inform environmental policy-making

Matilda Miljand
Stockholm University, Sweden

Katarina Eckerberg
Umeå University, Sweden

Abstract
There is a demand for scientific knowledge to make informed decisions in environmental policy. This study examines expectations of knowledge use, and how knowledge stemming from systematic reviews (SR) is being used through an analytical framework that distinguishes between instrumental, conceptual and legitimising evaluation use, as well as between process and product use. Empirically, we investigate knowledge generated from six SRs conducted through the Mistra Council for Evidence-based Environmental Management from the perspectives of those carrying out the SR and their targeted stakeholders. Our study reveals ways in which SRs are used and some characteristics that improve and some that hamper their usefulness. While the systematic method and the comprehensiveness of the SRs contribute positively to the usefulness, we found that the SRs produced were simultaneously too focused (lacking multiple perspectives), and too general (providing evidence on the effects of an intervention only at the general level) thereby restricting their usefulness. The time and resources it takes to produce an SR can also affect its usefulness compared to a traditional review.

Keywords
decision making, evaluation, evidence-based environmental policy, knowledge use, stakeholder involvement, systematic review

Introduction
Syntheses methods have long been applied when reviewing scientific literature. In the environmental science field, the traditional systematic review (SR) method was introduced in the...
early 2000s, with expectations that this could improve environmental policy-making (Pullin et al., 2004; Sutherland et al., 2004). SRs are literature reviews carried out using explicit accountable methods. They follow structured guidelines for how the review should be conducted, including how to select relevant literature and how to analyse the empirical data. This enables a more comprehensive description of the state-of-the-art in a field, new empirical conclusions to be drawn and theory to be developed (Miljand, 2020a). Traditional SR is guided by the idea of repeatability, objectivity and comprehensiveness (Miljand, 2020b), with the ‘aim to limit systematic error (bias) [in the literature review], mainly by attempting to identify, appraise and synthesise all relevant studies’ (Petticrew and Roberts, 2006: 9).

The argument for introducing the idea of evidence-based policy (EBP) into the field of environment management was that such decisions had often leaned on individual managers’ experiences and traditions, rather than on scientific evidence (Pullin et al., 2004; Pullin and Knight, 2003; Sutherland et al., 2004). EBP can be summarised as the conscientious, explicit and judicious use of current best evidence in decision-making (Sackett et al., 1996). According to the guidelines for carrying out SR in environmental management a

‘key objective of [SR] is to inform decision-makers of the implications of the best available evidence relating to a question of concern, and enable them to place this evidence in context, in order to make a decision on the best course of action’ (Collaboration for Environmental Evidence, 2013: 61).

So far, the prospects and challenges of using SR to inform environmental policy (Bilotta et al., 2014; Collins et al., 2019) and how to engage stakeholder in the SR process have been studied (Haddaway et al., 2017). Based on the experiences of researchers involved in making syntheses, Wyborn et al. (2018) find that research synthesis can contribute to the understanding of a problem, the establishment of networks and to changes in policy and practice. While making no definitive claims about the effects, they encourage these insights to be ‘taken as provisional hypotheses that could be explored or tested in future research’ (Wyborn et al., 2018: 81). They emphasise that future research should focus on the recipients of the syntheses. In their literature review, Oliver et al. (2014) note that most studies on EBP have focused on identifying factors that facilitate or create barriers for the use of scientific knowledge. They call for new ways of exploring EBP that also investigate how decision-makers use different forms of knowledge. This would require more complex theories than are currently used in EBP research to better understand the policy processes that the evidence is intended to inform. Oliver et al. (2014) thus point to an important knowledge gap in EBP studies, namely how evidence is used, rather than solely the extent to which it is used. With this article, we wish to contribute to this gap by investigating different ways of use that an SR can give rise to and analyse what factors may lead to a certain way of use.

Empirically, we examine the use of SRs conducted by the Mistra Council for Evidence-based Environmental Management (EviEM). This case selection is based on the fact that EviEM focused solely on SRs according to the well-established international guidelines (CEE guidelines) for carrying out SRs in the environmental field. In addition, EviEM were amply resourced for carrying out SRs and disseminating the results from the reviews, and can be viewed as a ‘best case’ for our purpose. This programme was funded by the Swedish Foundation for Strategic Environmental Research (Mistra) with the aim of providing comprehensive scientific evidence for decision-making (Mistra, 2011). It ran from 2012 until 2018 and published a total of 10 SRs and 4 systematic maps (SMs), mainly within natural sciences. An SM
is ‘identical to traditional SR with the exception that it does not include any synthesis’ (Miljand, 2020b: 48). The interest in SR methods has grown since their introduction into the environmental field in the early 2000s. For example, the journal Environmental Evidence dedicated to the publication of SRs in environmental science has published some 60 full SRs (4/11 2020). This development further motivates our study.

Aim and research questions

The overall aim is to explore how SRs, conducted to inform environmental policy, are being used by stakeholders involved in developing environmental policy. To nuance the understanding of how SRs are used to inform environmental policy, we borrow the conceptualisation of evaluation use proposed by Vedung (2021) and analyse SRs as a form of evaluation knowledge.

We ask (1) How do those who carry out SRs involve stakeholders in the review process and how do they expect the SRs to be used? (2) To what extent and how are the SRs used by stakeholders? And (3) What affects the perceived usefulness of the SRs?

The study is based on a combination of interviews, document analysis and personal experience from within the EviEM programme.

Analytical framework: different ways of knowledge use

While there are many different conceptualisations of knowledge use from the 1970s and onwards, three main ways of use can be distinguished: (1) instrumental use, when ‘evaluation findings inform decision making and lead to change in the object of evaluation’, (2) conceptual use, when ‘results lead to a better understanding or a change in the conception of the object of evaluation’ and (3) symbolic use, when evaluation results ‘are used to justify or legitimize a preexisting position, without actually changing it’ (Ledermann, 2012: 160). These conceptualisations build to a great extent on the seminal work by Weiss (1979, 1998, 2013). She presents seven different models based on different rationales and understanding of policy (Weiss, 1979).

Patton (1998) and Forss et al. (2002) highlight that use not only takes place when there is a published report. The evaluation process itself can give rise to “process use,” defined as “changes in thinking and behaving that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process” (Patton, 1998: 225). Patton (1997) further emphasises that ultimate use by the intended users (stakeholders) is determined by the extent to which evaluation is utilisation-focused from beginning to end. Use would then depend on whether these users have committed, from the onset, to invest in the process and on the process being designed to facilitate such involvement (Contandriopoulos and Brousselle, 2012). Drawing on insights from Weiss (1982, 1998, 2013), Patton (1998), Albæk (1988, 1996), Boswell (2009) and Dahler-Larsen (2011), Vedung (2015, 2021) defines six ways of use: instrumental, conceptual, legitimising (similar to Ledermann, 2012), plus tactical, ritual and constitutive ways of use (see Table 1). Like Patton (1998) and Forss et al. (2002), Vedung emphasises that use can occur during the evaluation process as such. However, a development with the framework by Vedung (2015) is that process use is seen as an additional dimension for each of the ways of use.

Table 1 helps to understand how evaluations may be used in general, but not all are fully applicable to our case study. The three latter ways of use (tactical, ritual and constitutive)
Table 1. Twelve ways to use evaluation knowledge, based on a typology by Vedung (2015) six uses of evaluation and Vedung (2021) the twelve-part typology of evaluation uses.

| Process          | Product                                                                 |
|------------------|-------------------------------------------------------------------------|
| Instrumental     | Instrumental process use includes cognitive, attitudinal and behaviour changes in individuals, or organisations, resulting either directly or indirectly from engagement in the evaluation process and learning to think evaluatively. Instrumental product use occurs when evaluative findings are employed as direct input into decision making. |
| Conceptual       | Conceptual process use is the use of evaluation process activities to create in-sights into the intervention of those concerned, resulting in enlightenment but not direct and immediate decision-making and action. Conceptual product use implies that evaluation influences on how users think. E.g. evaluations assist in labelling, categorising and framing intervention difficulties, causes of intervention difficulties and effects of such difficulties. |
| Legitimising     | Legitimising process use is when the evaluation processes are used to strengthen or weaken legitimacy. Legitimising product use means that the evaluation is used for legitimising purposes, it is employed to endorse previously established positions of the user. |
| Tactical         | Tactical process use is when the evaluation is used to argue that decision-making must be postponed until sometime in the future, when the evaluation will present its findings and recommendations. The evaluation is therefore an excuse for inactivity. In tactical product use it is the end-findings that are used as an argument that decision-making cannot be performed now because the evaluation findings have been presented at an inappropriate point in time or are unclear, or weakly underpinned, or deficient in some other way. |
| Ritual           | Ritual process use is when the evaluation is referred to as an activity that is announced and performed according to the organisation’s standard operating procedure. The evaluation is presented as one of a series of actions that is always done in a particular situation and performed in the same way each time. This e.g. to signal that one’s organisation is rational, modern, cost-conscious, ready to listen and learn about effects and side effects, or inclined to take well-grounded decisions. Ritual product use is when it is the end-findings that are talked about as being treated according to predetermined standard operating procedures and according to established routines and practices. This can e.g. be done by emphasising that once the evaluation has presented its findings they will, as usual, be studied carefully and handled in the same way as is always done by the organisation. |
| Constitutive     | Constitutive process use refers to the power of evaluation to institute, establish, or enact a practice in the field. Anticipation is the driving mechanism behind constitutive use. By anticipating an evaluation, those exposed to the evaluation change their behaviour. Constitutive product use occurs when findings from one evaluation creates an anticipation among others that when they are audited they will be assessed in similar manner and thus causing change in their behaviour. |

largely presuppose that evaluation activities are permanent and ongoing (Andersen, 2020), with institutionalised evaluation activities according to expected organisational needs and routines (Leeuw and Furubo, 2008). In our case, it was a temporary and independent research programme (EviEM) that decided whether to initiate each SR and on what topic. Even if EviEM received input from the targeted recipients of the respective SRs, tactical (to delay a
decision) or ritual use (in accordance with the recipient organisation’s routine) is deemed quite
unlikely, since EviEM did not have permanent ties with any of the invited public agencies. The
aim and design of each SR was decided upon by EviEM executive committee (with only loose
ties with the recipients for each SR). This situation is elaborated further in section 5.1-5.2.
Tactical, ritual and constitutive use was therefore not asked about in the interviews. However,
we describe how EviEM works with selecting topics and involving stakeholders and discuss
the implication of this for the potential for ritual, tactical and constitutive use in the results
section and concluding analysis. Below, we describe in more detail instrumental, conceptual
and legitimising use, the most relevant ways of use for this study. We specify further in section
4 how the analytical framework was applied.

**Instrumental use**

Instrumental use means that the results are applied in a direct way to influence what the
decision-maker decides to do next, such as to terminate, extend, change the content or design
of a policy or programme (Weiss, 1998). Policy is predominantly seen as a process of con-
scious and rational problem solving. Once a problem has been identified, research can provide
decision-makers with knowledge to help identify an appropriate solution to the problem or
choose between alternative solutions (Amara et al., 2004). In this understanding of scientific
knowledge, knowledge consists of facts which are neutral, objective and clearly distinguish-
able from values (Bijker et al., 2009) and can be conveyed to decision-makers at appropriate
stages of the policy process (Owens, 2015). The connection between knowledge and policy is
seen as linear and uni-directional, where knowledge informs a presumably rational and ordered
process of public policy formation.

**Conceptual use**

Conceptual use occurs when scientific knowledge affects users’ understanding of a pro-
gramme, policy or societal problem without instrumentally leading to the results being applied
in a specific decision. Weiss (1980) argues that the idea that knowledge directly affects a par-
ticular policy decision is problematic because it does not reflect how decisions are made, and
thus how knowledge is actually used. Weiss challenges the view that public officials are
engaged in clearly defined ‘decision making’; instead she depicts a decision-making process
characterised by fragmented authority. This can imply division of responsibility within an
organisation/agency, where no individual feels that s/he has a major influence over the deci-
sion, or that the power in a multi-level governance system is dispersed between different lev-
els. Hence, the process is the result of gradual, often diffuse steps through which decisions
take shape. In such a fragmented policy process, scientific knowledge may only partly give an
‘answer’ that political actors can apply. Instead, use of research becomes a much more diffuse
activity, more to orient oneself about problems (Weiss, 1977, 1980). Government officials
may use research to help understand or define a policy problem, using it to get new ideas and
perspectives, and to set the agenda for future policy measures (Weiss, 1977).

**Legitimising use**

When evaluations are used for legitimising purposes, they endorse previously established
positions. The findings are used to support one’s position, rather than for illumination.
Evaluations might be ‘used to exonerate (whitewash) or denunciate (blame) policies, practices, and institutions. The de facto task of evaluation is to supply argumentative ammunition for political and administrative discursive battles, where alliances are already formed and frontlines already exist’ (Vedung, 2015: 199). Legitimising use is often considered as problematic in the utilisation discourse, however, Vedung (2015) identifies a scholar – Majone – holding a more positive view on legitimising use. Majone (1989) argues that in politics, it is not enough to make decisions, even if they are correct – justificatory arguments are needed too. Once a decision is made, it must be justified, legitimised and explained. Moreover, decisions are often made in response to external pressures or are based on personal convictions, and in

‘such cases arguments are needed after the decision is made to provide a conceptual basis for it, to show that it fits into the framework of existing policy, to increase assent, to discover new implications, and to anticipate or answer criticism’ (Majone, 1989: 31)

As policies exist for some time, their political support must be constantly renewed through new arguments to justify a policy in a constantly changing environment (Majone, 1989).

**Method and material**

EviEM’s first six SRs that were completed at the time of our study were examined, including four traditional SRs and two systematic maps (SM), see Table 2 for their respective aim and coverage.

We use SR as a label for both SR and SM in our empirical analysis, and mention SM only when a separation is analytically warranted. The first six SRs were chosen on the premise that sufficient time is needed after a report is published before it is reasonable to investigate how it is received and used. Having said that, we are aware of the methodological difficulties in pinpointing what knowledge is used at what point in time, as well as the many analytical pitfalls when trying to follow the practical use of specific evidence into policy decision making with concern to relevance (timeliness, salience and actionability), credibility and accessibility (Contandriopoulos and Brousselle, 2012; Owens, 2015). For example, the users’ understanding of relevance and credibility is very much influenced by their pre-existing opinions and preferences (Contandriopoulos and Brousselle, 2012). For those reasons, we do not attempt to provide factual evidence for use, but rather for the perceptions of such use taking place by the involved actors in each SR. Furthermore, we are aware that learning within EviEM itself might show slightly different results in subsequent SRs, when stakeholder involvement might have become more developed. Still, we believe that the six SRs are sufficiently varied in terms of scope and target audiences within environmental policy-making to justify this selection. Each SR was carried out as a project led by a project manager from EviEM tasked with coordinating the process. In addition, a group of approximately five researchers (one chairing the team) carried out the SR. The project manager was responsible for making sure that the SR followed the SR guidelines, while the chair was responsible for the scientific content of the SR.

Using the instrumental, conceptual and legitimising knowledge ways of use in Table 1, we analysed both expectations of knowledge use that the producers of SRs (both project managers and researchers) had, as well as the stated use (by the recipient stakeholders) of the SRs.
| Publication year | Aim of the review                                                                                                                                                                                                 | General conclusions                                                                                                                                                                                                                                                                                                                                 | Type of review | No. of included articles in the review/articles identified |
|-----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------------------------------------------------|
| 2015            | The aim of this SR is to find out what science has to say about the impact of reindeer grazing on mountain vegetation. The question to be answered is: what are the impacts of reindeer/caribou (Rangifer tarandus L.) on arctic and alpine vegetation? | Overall, the studies report on large variations in the response of vegetation to reindeer. Average effects are usually small, but there are also both significantly positive and significantly negative effects within all vegetation categories considered.                                                                                                               | Systematic review | 41/6 638                                               |
| 2016            | The aim of this SR is to compile all the available evidence, synthesizes the results and assesses how effective wetlands are for nitrogen and phosphorus removal. The question to be answered is: how effective are created or restored freshwater wetlands for nitrogen removal and phosphorus retention? | On average, created and restored wetlands significantly reduce the transport of nitrogen and phosphorus and may thus be effective in efforts to counteract eutrophication. However, restored wetlands on former farmland were significantly less efficient than other wetlands at phosphorus removal.                               | Systematic review | 93/5 853                                               |
| 2015            | The aim of this SR is to investigate whether biomanipulation – reducing the populations of roach and bream – improve water quality in eutrophic lakes. The question to be answered is: what is the influence of a reduction of planktivorous and bentivorous fish (performed directly or indirectly through stocking of piscivores) on water quality in temperate eutrophic lakes? | Removal of planktivorous and bentivorous fish is indicated to be a useful means of improving water quality in eutrophic lakes. Biomanipulation tends to be particularly successful in relatively small lakes with short retention times and high phosphorus levels. More thorough fish removal increases the efficacy of biomanipulation. Nonetheless, successes and failures have occurred across a wide range of conditions. | Systematic review | 124b/1 2 908                                            |
| 2015            | The aim of this SM is to examine the evidence relating to how different farming methods affect the amounts of organic carbon in arable soils. The question to be answered is: what are the effects of agricultural management on SOC stocks? | The map identified four broad categories of arable farmland practices, each having been investigated by between 238 and 307 studies: soil amendments, crop rotations, use of fertiliser and tillage. The SR covered areas with a temperate or boreal climate and found studies from the majority of these areas, but there was a lack of studies from Russia, a sign of a knowledge gap in the literature. The mapping process also identified several common problems with the research methods used in many studies. | Systematic map | 735/24 547                                               |
| 2018            | The aim of this SR is to investigate how the phase-outs of PFASs – substances used in many commercial products – has affected concentrations in the environment. The question to be answered is: what is the effect of phasing out long-chain per- and polyfluoralkyl substances on the concentrations of perfluoroalkyl acids and their precursors in the environment? | In regions where regulations and phase-outs have been implemented, human concentrations of PFOS, PFDS, and PFOA are generally declining. Rapid declines for PFOS-precursors have also been consistently observed in human studies. In contrast, limited data indicate that human concentrations of PFOS and PFOA are increasing in China where the production of these substances has increased. For abiotic and biological environmental samples there are no clear patterns of declining trends. Results suggest that declining PFOS, PFOS-precursor and PFOA concentrations in humans likely resulted from removal of certain PFASs from commercial products or from food packaging. | Systematic review | 92/10170                                               |
| 2015            | The aim of this SM is to assess what evidence exist about whether protected forests should be left completely untouched, or if some kind of active management is required if their biodiversity is to be preserved. The question to be answered is: what is the impact of active management on biodiversity in boreal and temperate forests set aside for conservation or restoration? | The SM identifies a wealth of evidence on the impact of active management practices that could be utilised to conserve or restore biodiversity in forest set-asides. As such it should be of value to conservation managers, researchers and policymakers. Moreover, since the map also highlights important knowledge gaps – such as a scarcity of studies on hydrological restoration and on certain groups of invertebrates – it could inspire new primary research. Finally, it provides a foundation for more complete synthesis of specific subtopics. | Systematic map | 798/16 484                                               |

SR: systematic reviews; SM: systematic map.

aAfter removal of duplicates.

bAfter full text assessment, 233 articles remained, however, data were extracted from 124 as the other articles did not provide relevant information.
The first research question was posed to both SR project managers and researchers. The second and third questions addressed project managers, researchers and stakeholders in the studied SRs. The interviews covered a number of different themes including the interviewee’s role in the SR process, their general view of SRs and knowledge use, the design and conclusions of the specific report, EviEM’s efforts to involve and disseminate knowledge to various stakeholders and finally, if and how the knowledge from the specific SR had been used in a policy process. In total, 16 semi-structured interviews (between 49 and 108 minutes) were carried out in 2018-2019 (see Table 3). Interviews were recorded with the permission of the participants and transcribed. The interviewees were given the opportunity to validate the transcripts from the interviews, but all declined. We decided not to reveal the names of interviewees to allow them to speak more freely. All quotes presented are from interviews, with the role of the interviewee in parentheses.

It is possible to make generalisations from a case study (Flyvbjerg, 2006). However, in contrast to statistical generalisation, it is not about generalising to a population, but rather generalising to other similar circumstances or situations (Yin, 2013). This analytical generalisation means extracting abstract level of ideas from case study results, by producing a theory of how and why the studied events occurred. With our case study of EviEM, we aim to make analytical generalisations about how and why SRs are used within the environmental science and policy field. While the frequencies of use should not be generalised to a population, the conclusion that the SR reports are used in different ways and that aspects of the SRs make them more likely to be used, does have a broader relevance.

An operationalisation of the analytical framework is presented in Table 4. A distinction was made in the analysis between the strength of different accounts of use. When someone stated that they had used the results, but without giving a concrete example of how, it was considered a weak signal of use. We also included all statements of third party use – when the interviewee gave an account of how someone else had used the SR. If the interviewee could describe how s/he or he or her organisation had used the SR, this was considered a strong signal of use.

**Results**

The presentation of results follows the order of the three research questions. We begin by depicting how EviEM has worked with selecting topics, including stakeholders and the implication this has for use, followed by how the interviewees expected the SRs to be used, and whether and how stakeholders stated that SRs were used. Finally, we present the various perceptions of the usefulness of SR.

**EviEM’s approach to selecting topics and including stakeholders**

According to EviEM’s guidelines for selecting topics to review (Mistra EviEM, 2012), suggestions for topics may come from its own Executive Committee, stakeholders or scientists. The suggestions were then discussed and elaborated in consultation with relevant stakeholders. For each potential topic the EviEM secretariat carried out a pilot study as basis for deciding which topics to approve for a full review. The Executive committee required that ‘the question should concern an area in which the scientific results are incompletely surveyed (or perhaps disputed) but where some grounds for conclusions nonetheless exist’ (Mistra EviEM: website). In addition to this necessary criterion, seven additional criteria for selecting topics
were applied, for example, if the issue was controversial, relating to valuable natural assets or particularly costly measures. In addition to initially approving the topic, the Executive committee also approved the final report before it was published.

EviEM used a five-stage approach for interacting with stakeholders: (1) identification of stakeholders, (2) identification of policy- and practice-relevant topics, (3) framing and prioritisation of review questions, (4) establishment of the specific scope of a review and (5) a public review of a draft review protocol, that is, the plan for the SR as designed by the project manager and researchers (Land et al., 2017). This interaction aimed at making the focus of the reviews relevant for the stakeholder. When the SR was almost finalised, a draft version was sent to a smaller group of stakeholders for comments (mainly on presentation of the results). The final results were thereafter communicated to stakeholders through targeted oral presentations, videos and reports, and made publicly accessible on the EviEM webpage http://eviem.

Table 3. Overview of conducted interviews. Three different categories were interviewed, covering all SRs and SMs. Two project managers oversaw multiple SRs and one interviewed stakeholder was involved in two of the SR processes.

| Interviewee categories     | No. of interviewees |
|----------------------------|---------------------|
| Project managers           | 3                   |
| Chairs/Researchers         | 6                   |
| Stakeholders               | 7                   |
| Total                      | 16                  |

Table 4. Operationalisation of the analytical framework for knowledge use. Defining strong and weak signs of relevance in interviewee answers.

| Process         | Product                                                                 |
|-----------------|-------------------------------------------------------------------------|
| Instrumental    | Stating that the SR had been used in the organisation (weak sign).     |
|                 | Identifying concrete situations in which the findings from the report have been used as input to a decision (strong sign). |
| Conceptual      | Stating that the SR had influenced how s/he thinks about the issue studied (weak sign). |
|                 | Providing example(s) of how, e.g. assisted in labelling, categorising or framing intervention difficulties, causes of intervention difficulties, and effects of such difficulties (strong sign). |
| Legitimising    | Stating that the SR had been used to legitimise the use of an intervention or action (weak sign). |
|                 | Providing example(s) of when the SR had been employed to endorse previously established positions of the user or organisation (strong sign). |

SR: systematic reviews.
It is worth noting that there was little to no interaction with stakeholders while the review was ongoing, instead the interaction focused on the planning stage and after the review was finished.

**The implications of this for ritual, tactical and constitutive use**

Ritual use implies that an evaluation is used to show that an activity is announced and performed according to the organisation’s usual procedure, hence this use occurs only in certain organisational contexts. It could occur if an SR were carried out for internal use by an organisation itself, for example, to shed light on a new issue that the organisation had to deal with. Furthermore, SR could be employed for tactical use, that is, to deliberately delay a decision within an organisation. Since the SR review process can span several years, initiating an SR could be a way of avoiding making a decision by referring it to the outcome of the SR. However, EviEM was an independent organisation where no one could dictate or pay for a particular review to be carried out. Furthermore, EviEM did not work towards any specific stakeholder but rather tried to include a variety of stakeholders’ opinions, meaning that no single stakeholder could take control over the selection process. It is therefore unlikely that ritual use would occur under this organisational context as this kind of use is dependent on the ability to affect the initiation of the review. Ritual use also requires continuity and that SR becomes part of the ‘normal procedure’. Further, EviEM’s independent organisation makes tactical use more difficult, though less so than ritual use. An organisation could not decide to initiate an SR at EviEM as a reason to postpone a decision, however, it would be possible to justify not making a decision before an ongoing review is finished. Even if it is admittedly difficult to investigate, we did not find any indication of tactical use of the six investigated SRs.

SRs may have constitutive effects in the case of an ongoing evaluation practice. Constitutive use refers to the power of evaluation to institute, establish, or enact a practice in the field; thereby certain norms and procedures for what ought to be done are shaped by the logics and practices inherent to the evaluation activity (Dahler-Larsen, 2011). By anticipating an evaluation, those exposed to the evaluation change their behaviour beforehand. For such effects to occur the execution of SRs must be organised in such a way that they are carried out regularly to enable expectations. There are aspects of SR that makes constitutive use less likely compared to other evaluation methods. SRs are compilations of existing research and not evaluations of how a specific organisation work or a specific intervention that an organisation has taken. It is therefore less likely that the expectation of an SR would change how one works. Yet, it is possible that the expectation that an SR will come to the conclusion that a certain intervention is effective or ineffective can cause the attitude to this intervention to change in advance. The fact that EviEM was a time-limited research programme means that it cannot be expected to be recurring and thus give rise to constitutive effects.

**Expected use of finalised SRs**

Our interviews consistently show that SRs are indeed expected to be used to help improve environmental policy-making by decision-makers. The instrumental use of the results was emphasised as the primary expected use by the project managers. Examples were given for the specific reports:
We found evidence that biomanipulation could be a good method. [...] But here we have at least identified a method that works. [...] So I would say that it is concrete environmental protection so to speak. It could be used for that. (Project manager)

Another pointed to using the results to promote a certain intervention:

I think it will affect the priority [...] It has sometimes been emphasised that we do not really know if wetlands really work [to remove nutrients] [...] And yes, our investigation shows that there are such wetlands [where it does not work], but we have also shown that the vast majority of them work very well and we have also pointed out the reasons why some work poorly. (Project manager).

The same project manager also emphasised that the most important use is to legitimise further wetland interventions. Moreover, several researchers expected the findings to be instrumentally used.

There were two important policy decisions that I would pull out of this. There is pressure to reduce the use of these substances because it really affects people, at least people’s exposure. And then, to see if it actually has an effect out in the environment, you need good environmental monitoring. (Researcher)

In addition, both project managers and researchers expected non-use of the SR and the two SMs that did not arrive at any clear conclusions for environmental policy practice. Instead, they pointed to their usefulness to the research community in terms of justifying the need for more research and revealing methodological weaknesses in existing research. As one project manager put it:

the two main groups of stakeholders that this knowledge appeals to would be researchers who want to carry out primary research [...] and then those who can fund the gaps or the [identified knowledge] clusters. (Project manager). The researchers shared a doubt whether the SR would be used. This was expressed in one SR as: The stakeholders were probably disappointed, they probably did not get what they wanted. It felt like the issue was a bit outdated when the result came out. There was no huge rush after the report as far as I know. (Researcher)

They also expressed a worry that expectations from stakeholders could not always be met. One stated that the conclusions were unsurprising for the researchers, but that the stakeholders were probably disappointed as it was not possible to provide clear conclusions. Another mentioned that many of their SR questions could not be answered due to few and disparate studies.

Mainly, the outcome of the SRs was seen as a mixture of instrumental and legitimising use, but there was also an anticipation of non-use. This could be understood in relation to the conclusions of the SR, where there is a tendency, perhaps not entirely surprising, that instrumental and legitimising use is expected when the SR finds concrete effects of an intervention. When the results are more unclear, non-use is expected.

**Process use**

As mentioned above, most involvement of stakeholders took place at the beginning of all the SR processes, when the topic was chosen and the question formulated. The project managers
attested that the questions coming from practitioners tended to be too broad at the onset, requiring considerable discussion before they could be applied to SR. This made the SR formulation stage pivotal by way of discussing with the stakeholders the most appropriate research question to examine.

Several interviewees highlighted the need to involve stakeholders both from the start and during the process, to help keep the SR group’s focus on the ‘right’ questions from a policy perspective. Others felt that stakeholders should be involved primarily at the beginning when the SR question was formulated. There was agreement that stakeholder involvement in SR processes could be a challenge due to the technical nature of the evidence gathering and synthesising. We found that in practice this limit which stakeholders can effectively contribute to the process. One researcher stated that a scientific background is required to be able to fully read and comprehend a SR. An unforeseen finding in our study was that the stakeholders that were most active in the review process, and therefore were interviewed for this study, were civil servants from national agencies (therefore with resources to participate) who had a natural science research background. This helped them to understand the SRs. A broader group of stakeholders with other backgrounds might have different views on their involvement in the SR process. This should be considered if the results of our study are to be generalised.

We now turn to the question of process use. One instance of conceptual process use was identified, where one stakeholder felt that the involvement had provided knowledge about the SR method as such and to some extent contributed to expanding her or his knowledge network: ‘I got a better understanding of how SR works, the process. I had a good idea, but it was interesting to get some more insight into it. And I probably got a slightly wider contact network as well’. (Stakeholder). No instances of instrumental or legitimising use were found with concern to the process even if it can be regarded as a valuable knowledge exchange in itself. It can also facilitate product use due to better understanding of how the report – as a product – has been prepared.

While we did not explicitly investigate how the researchers themselves had used the SRs, this was brought up in the discussion on related questions. It became clear that the SRs were expected to be used for scientific purposes with the interviewees providing examples of such use. One stakeholder had used the SR to motivate further research, in presentations, and in teaching. The SMs were also found to have mainly scientific use, according to project managers, researchers and stakeholders. An aspect of scientific use is how involvement in an SR process leads to learning among researchers (c.f. Solomon et al., 2001). We found that the SR process had provided insights about the researchers’ own research field. Not so much that the empirical conclusions of the SR were new, but they gained insights on questions of research design, method choices, quality deficiencies and so on. These insights can provide a solid foundation when giving advice to decision-makers. Hence the learning that we have identified among researchers can, at a later stage, be used in decision-making.

**Product use of finalised SRs**

When interviewing the stakeholders, we identified one weak signal of instrumental use (see Table 5) of the SR reports (product use), where it had been used by a government official when in contact with those working with implementing an intervention. According to the interviews with researchers, another SR had also been directly used. S/he stated that consultant firms had been in contact and that this had led to action. ‘If I look at the process afterwards, the contacts
that I have had as chair with users is without doubt with consultant firms that are asking stuff’. (Researcher)

Conceptual use refers to when the SR has influenced how the stakeholder thinks about the studied issue. We found five indications of weak conceptual product use. These were mainly general statements about insights gained from the SR. One stakeholder expressed that the report had provided an international outlook that is often missing. Another said that the report provided evidence for a greater effect of the intervention than expected (even if the direction of the effect was expected). In a similar vein one stakeholder said that the SR report confirmed previously held understandings but could show that not only did the level of PFAS (poly- and perfluorinated alkylated substances) remain stable (as expected) but an increase could be detected. Two others stated that the reports have led to a (limited) reflection around management measures in nature reserves, and had provided some insights into water retention in the landscape.

We found four instances of legitimising use of the SR report, including one strong. Stakeholders gave accounts of how the SR had motivated why a project was important, and verified what was known, ‘it did not affect the process because it was already underway. It validated and substantiated what we did. And that’s an effect there too’. (Stakeholder). It thus strengthened the belief that there is a need to limit the emissions of PFAS,

it will reinforce the evidence we already have supporting the need for a limitation. And that limitation might come to pass regardless if this report had come or not’. (Stakeholder). In the wetland SR, the conclusions were used to advise landowners, “Yes, we use it in contact with landowners to some extent. Many people ask “how good are they really, how much nitrogen is removed?” And then you can point out that it was a very comprehensive study where you went through studies from Europe and North America and then you came to the conclusion that [...] much [nitrogen] is removed. (Stakeholder)

Several researchers affirmed that an SM may provide insight into what has been studied and not. Hence, the SM helped to focus the need for further studies, valuable from a research perspective but less so for practice. One stakeholder pointed out that the absence of conclusions about management interventions made it difficult to use the SM. One stated that while the SM was a good starting point, ‘it felt like it was better for researchers who are doing this. So there are no obvious uses for us. It is not sufficiently aggregated’. (Stakeholder). While providing insights as to the state of the knowledge, they did not see the usefulness of the SM in their work: ‘We are still waiting for the rest [subsequent SRs]. We saw that we will not be

Table 5. Expectation of use and stated use of the SRs. The project managers’ and researchers’ expectations of process use was not investigated.

| Process        | Stakeholders | Project managers and researchers | Stakeholders |
|----------------|--------------|---------------------------------|--------------|
| Instrumental   | Stated use (weak): 1 | Expected use: 4 | Stated use (weak): 1 |
| Conceptual     | Stated use (weak): 1 | Expected use: 1 | Stated use (weak): 5 |
| Legitimising   | Expected use: 1 | Stated use (weak): 2 | Stated use (strong): 2 |

...
able to do much of the first [SM report]. But when the others come then we need to take care of them’. (Stakeholder).

Understanding the context in which an evaluation is to be used is crucial for understanding why use occurs or not (Contandriopoulos and Brousselle, 2012). The previous perceptions of the user is one such contextual factor:

when a user’s understanding of the implications of a given piece of information runs contrary to his or her opinions or preferences, this information will be ignored, contradicted, or, at the very least, subjected to strong skepticism and low use. (Contandriopoulos and Brousselle, 2012: 63)

Both researchers and stakeholders expressed that the conclusions drawn in the SRs and SMs included in this study confirmed their previous views. According to Contandriopoulos and Brousselle (2012) this should increase use of the knowledge. However, in our study, the fact that the results were in line with the interviewees’ opinions also meant that it did not lead to any changes. One stakeholder said,

it has not changed our policies and the projects we actually support precisely because it verified what we thought we knew before and the practice and experience we had [. . .] But it has been a good support to say whether we should continue in the future. (Stakeholder)

This confirmation was perceived as positive and it was emphasised that there is considerable value in conclusions that are more informed by scientific evidence.

To conclude, those who mentioned that they had used the SR reports generally highlighted the conceptual and/or legitimising value. When the SR showed that an intervention had proven good effects, it confirmed previous beliefs and legitimised a certain intervention. Also, several of the SM/SRs were deemed more important for proposing further research than for informing practice.

The usefulness of SRs

So how useful are the SRs and what affects their usefulness? While these results are based on the experiences of EviEM, the findings are generalisable beyond EviEM as they concern the SR method itself, rather than the workings of EviEM. It should be noted that our conclusions are drawn from SRs within the natural science disciplines, but the issues are similar but even more complex had the SRs been carried within environmental social science (Miljand, 2020a). However, our results are confined to the environmental science and policy field, due to the very different institutional and professional context within for example medicine and social work where SRs are also frequently used.

Project managers, researchers and stakeholders claimed that the strength of SR is the systematic method and the comprehensiveness in terms of included literature on the topic which contributes to their usefulness. Our study also shows that the time it takes to carry out an SR affects how useful they are. If it takes very long from the time an SR is initialised until the results are ready, this has a negative impact on their usefulness. One stakeholder highlighted the risk that the issue is no longer relevant when the results are presented. The expected use of SR is largely dependent on a speedy process:

When we start a project, there is great interest among the stakeholders . . . but if it takes 2 to 3 years . . . So I believe that if we can reduce the time to one year . . . I think that would be the most important to improve the use. (Project manager).
While it is not possible to give a clear answer to how long it takes to produce an SR, it is well known that they are resource intensive and more time consuming than normal literature reviews. Researchers have tried to estimate the time, and in the medical field some have estimated 6-18 months (Smith et al., 2011) and others find a mean of 67.3 weeks (Borah et al., 2017). Haddaway and Westgate (2019) find that a typical SR published by the Collaboration for Environmental Evidence screened on average approximately 8500 unique records and the ‘total time estimated for an average [SR] was 164 person days at 1 full time equivalent’ (p. 438) albeit normally they are carried out part time and therefore extend over a longer period until finalisation. We found that the SRs carried out by EvieEM were more time consuming than the average time estimated, however they are not uniquely long and the results of this study illustrate that the length of SRs can impede their usefulness.

The comprehensiveness of the SR process also raised concern among researchers. One researcher was doubtful whether the gain in doing an SR outweighs the costs:

I’m not sold on SR being so much better than other reviews . . . It is an expensive process, cf. with other types of overviews . . . and especially if you expect that those who will use these will pay for them then you must be able to show that they give something that is worth that extra. (Researcher)

Another researcher stated that SR involves extensive work in order to answer few questions, but that there is no other alternative. In response to the expected use of SR, one of the project managers noted that in many instances, the expectations are unrealistic and would waste resources, since the stakeholders really only want an overview of the literature, rather than showing all evidence or having very precise measures of variability concerning environmental effects.

There was consensus about SR as best suited to answer clear and well-defined questions. Some questions are more prone to being answered by SR: ‘So far, we have looked at natural sciences questions . . . where there are concrete outcomes that are unambiguous and easily measurable’ (Project manager). Indeed, the applicability of SR was considered best when only one outcome parameter is measured, where the research field is well defined and there is homogeneous data material. This affects the usefulness of the SRs and the risk that the SR becomes very narrowly focused was raised, making it difficult to obtain a broader picture. It was suggested by one researcher that one way of overcoming this problem would be to carry out multiple narrow reports and link them together. Also, SR was generally considered most beneficial when there are pre-existing differences in understanding or even issue polarisation in the research community and/or a need to reduce uncertainties. This would be: ‘to untangle ambiguities or erase uncertainties when we don’t know things. Concrete things you really feel that “there are very different views in a research community”. Then SR can be very good’ (Researcher). However, the uncertainties need to be around the effect of an intervention, since SR cannot mediate a situation where there are divergent understandings of what constitutes the problem to start with, or how it should be investigated.

Another theme that several of our interviewees returned to concerns the conclusions that are produced, which risk becoming simultaneously too general and too narrow. The general-specific continuum refers to how aggregated the results are, and whether the conclusions can be formulated into practical advice. The narrow-broad continuum refers to the scope of the SR question and how many aspects and perspectives are included. SR tends to provide answers on whether an intervention works at a general level, where there is risk that too much important
detail is lost. This might end up lacking relevance at the local level – where most interventions take place. One stakeholder thought that the usefulness of SR depends on the specificity of the conclusions, stating that when compiling data, there is a risk of losing details, which makes it difficult to judge whether the results are relevant. Hence, the usefulness of SR to inform decision making depends on the kind of conclusions that are drawn. This also relates to another aspect brought up, namely how to approach comparisons between countries (or different ecosystems). It can be difficult to synthesise such diverse data and constitutes a challenge for stakeholders to know which conclusions are applicable to their own area when the results are aggregations across multiple contexts.

In brief, comprehensiveness was considered the most distinctive feature in terms of usefulness. The stakeholders who received the SRs appreciated this, but opinions differed as to whether the required resources can be justified. Furthermore, the usefulness was hampered both by long time frames and by the SRs being simultaneously too general and too narrow, making the conclusions less applicable for stakeholders.

**Concluding discussion**

Our study reveals different ways in which SRs are used, as well as some important factors affecting their usefulness. We only identified one instance of process use (conceptual), but found several examples of both instrumental, conceptual and legitimising product use. We are aware that the nature of EviEM as an independent organisation for SRs will affect these results. If the SR activity is developed internally and long-standing, one would expect a greater propensity for instrumental and constitutive use. Nevertheless, our conclusions about the ways in which the SRs may be used, and the reasons for such use, should be generalisable to other institutional and international contexts similar to Sweden within the environmental science and policy field.

As anticipated, for well-known methodological reasons, we struggled to identify instrumental use of the knowledge stemming from the SRs, even if there was some indication thereof in the interviews. The instrumental view on use has been criticised for its lack of realism (Amara et al., 2004), and for its simplified understanding of how policy is made (Weiss, 1980). Yet, it still has great appeal to many, potentially ‘because it retains a certain intuitive appeal and a modicum of explanatory power’ (Owens, 2015: 7).

Testimonies were also found of conceptual and legitimising product use. It is worth noting how such use clearly relates to the specific features of SR. The conceptual use tended to be quite vague and general. One reason might be that the results of the SRs did not alter the interviewees’ understanding of the SR questions at hand, but rather confirmed or reinforced previously held beliefs. First, this may be explained in that SR, by definition, is based on existing evidence, and while the SR in itself can produce new knowledge and even alter previously held understandings, any SR is more likely than not to be in line with the general understanding of the problem. Second, and tied to this observation, was that most of the stakeholders in our study turned out to have a PhD in science. Hence, they were quite informed about the science on the specific topic. A third reason might be that learning had indeed taken place during the SR process among the stakeholders, even if the interviewees were explicitly unaware of this, and that they had difficulty remembering exactly what they thought beforehand. The latter is a methodological challenge when investigating knowledge translation (Freeman, 2009).
The legitimising use that was identified raises the question: whether confirming one’s beliefs should be considered a form of knowledge use and if so, if it is to be equated to legitimising use. As mentioned, legitimising use is often met with scepticism, being seen as cherry-picking knowledge that fits preconceived understandings. Yet here, legitimisation is about confirming to oneself and others that the intervention taken is the right one. The comprehensiveness of the SRs then helped to give weight behind the argument for continuing with the intervention. One interviewee emphasised that confirming that the intervention was effective was also a way of using the knowledge even if it did not imply any change. Another interviewee expressed that the SR gave confidence to present the intervention s/he already believed in to managers and the public. For the stakeholders the most apparent use was thus not to reveal much new knowledge but rather to confirm, or legitimise, existing practice. As SRs are unlikely to alter the beliefs of the stakeholders, they are most probably used to confirm already held beliefs. In line with Majone (1989), this may then support their argumentation for a specific policy.

We identified some characteristics that improve and some that constrain the usefulness of SRs. While the systematic method including its comprehensiveness was considered a major strength, the SR methods do not solve any of the fundamental issues of science-policy interaction. Rather, by making reviews more scientifically robust, they focus on increased credibility rather than timeliness or actionability. Compared to, for example, utilisation focused evaluation where all steps in the evaluation process is focused on utilisation (Patton, 1997), in an SR process the focus is on credibility. It is only when that aspect is established that the use aspect can be addressed, as evident in how EviEM worked. The process to be applied was very strict, and inclusion of stakeholders only occurred in the planning and communication of results. The type of question that can be answered is predetermined from the method and therefore efforts are made to adjust the potential users questions to fit the SR methods demand. This can be expected to affect and decrease the use of the SRs.

The fact that the SR questions were simultaneously too focused and too general can also hamper usefulness. When the knowledge becomes too general (compiled on a global level), it can be difficult to apply in local practice and when the SR becomes too focused and the compiled knowledge narrower, this can limit its usefulness in specific situations.

Moreover, usefulness is not only determined by the SR itself, but also by the policy context. It is important to be aware of at what stage of the policy cycle a certain piece of evidence can be useful. Evidence can be more influential if it is provided at an early stage when an issue is being investigated compared to later, when just the issue is to be decided. At that point most involved actors have already made up their minds and might have invested political capital in the issue, thus being less prone to change their minds as a result of new evidence. At what stage an SR feeds into a policy process can therefore help explain the occurrence of use. We have not investigated the specific policy processes in detail in this study, but the timing in relation to the relevant policy process was brought up in our interviews. Examples were given of when decisions had already been made or the policy problem had been reformulated when the SR was finally published. If the SR is published after a decision is already made, this can help explain both the absence of instrumental use, and why legitimising use becomes more common. At this stage, the SR is used to support the continuation of the intervention. Hence, generally formulated conclusions become useful to illustrate the value of an intervention that is already in place (legitimisation).

The importance of timeliness is emphasised by Rose et al. (2017) who argue that the opportunity to influence policy often occurs within so-called policy windows – that is, windows of
opportunity for policy change creating favourable situations for the uptake of knowledge (Rose et al., 2017). Therefore, how to produce syntheses that can feed into such policy windows becomes crucial. As SRs take significant time to compile and analyse, it could be too late to initiate a review when a policy window opens up, since by the time the SR is finalised it might already have closed. Instead it becomes important to ‘collate evidence proactively in advance, even if the saliency of the evidence might not be significant at the time of collation’ (Rose et al., 2017: 51) and then make use of this evidence when a policy window occurs. In response to the time required to carry out an SR, other variants of SR methods have also developed, such as scoping reviews if what is needed is rather an overview of a wide area than detailed answers to a given question (Moher et al., 2015). Another variant is so called rapid reviews. There is no standardised methodology to carry out rapid reviews, however they are adaptations of traditional SR method aimed at shorten the process by, for example, limiting searches by year, databases, or language or exclude quality assessments of the included studies. At the same time, these limitations could imply that bias is introduced in the review (Ganann et al., 2010).

The fact that we found the SRs to lead to learning among the researchers can prove to be important for the transmission of knowledge from the researchers to policy-makers. The lack of timeliness of a given SR can be mediated if researchers who carry out SRs learn from these and can provide good advice when a policy window opens. It is important to emphasise that use can take place both in direct connection to the presentation of evidence or years after it. An SR carried out today can be ‘dormant seeds, generating little impact at the time of activity but coming to fruition much later’ (Owens, 2015: 84). While being hard to investigate when it is years between the cause and the effect, there is a clear risk of underestimating the impact of evidence on policy. In her study of the UK environmental committee Owens (2015) found that the time from advice to impact on policy can vary between 15 minutes and 25 years.

We conclude that expectations that SR can provide evidence-based decision-making are based on two main arguments: First, that the SR method as such is both comprehensive and stringent, thereby shedding light on all the relevant literature on the topic at hand. Second, that by letting stakeholders into the process of formulating research questions and commenting upon the SR result this would improve its usefulness. Nevertheless, our study shows that these expectations are seldom fulfilled in reality due to various constraints, both as to which topics can be addressed and how the SRs can be used. The limitations concern the required time and resources, as well as the type of questions that could be addressed, which proved not entirely relevant to the stakeholders in the end. Still, it is important to consider the many ways in which scientific knowledge is potentially brought to policy makers, where this study has only focused on the SR projects as such. It remains to be further examined to what extent, and how, the knowledge stemming from SR might penetrate in other forms to reach decision-makers, for example through expert advice, but also to inspire new research in areas where the knowledge base was found to be insufficient.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Swedish foundation for strategic
environmental research under their EviEM programme as part of a PhD project entitled ‘Systematic review of social scientific environmental management questions’ (2015-2018).

**ORCID iD**
Matilda Miljand [https://orcid.org/0000-0002-6902-6415](https://orcid.org/0000-0002-6902-6415)

**References**
Albæk E (1988) *Fra sandhed til information: Evalueringsforskning i USA-før og nu*. Copenhagen: Akademisk forlag.
Albæk E (1996) Why all this evaluation? Theoretical notes and empirical observations on the functions and growth of evaluation, with Denmark as an illustrative case. *Canadian Journal of Program Evaluation* 11: 1–34.
Amara N, Ouimet M and Landry R (2004) New evidence on instrumental, conceptual, and symbolic utilization of university research in government agencies. *Science Communication* 26: 75–106.
Andersen NA (2020) The constitutive effects of evaluation systems: Lessons from the policymaking process of Danish Active Labour Market Policies. *Evaluation* 26: 257–274.
Bijker WE, Bal R and Hendriks R (2009) *The Paradox of Scientific Authority: The Role of Scientific Advice in Democracies*. Cambridge, MA: MIT press.
Bilotta GS, Milner AM and Boyd I (2014) On the use of systematic reviews to inform environmental policies. *Environmental Science & Policy* 42: 67–77.
Borah R, Brown AW, Capers PL, et al. (2017) Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* 7(2): e012545.
Boswell C (2009) *The Political Uses of Expert Knowledge: Immigration Policy and Social Research*. Cambridge: Cambridge University Press.
Collaboration for Environmental Evidence (2013) Guidelines for systematic review and evidence synthesis in environmental management. Version 4.2. Environmental Evidence. Available at: https://environmentalevidence.org/wp-content/uploads/2014/06/Review-guidelines-version-4.2-final.pdf
Collins AM, Coughlin D and Randall N (2019) Engaging environmental policy-makers with systematic reviews: Challenges, solutions and lessons learned. *Environmental Evidence* 8: 2.
Contandriopoulos D and Brousselle A (2012) Evaluation models and evaluation use. *Evaluation* 18: 61–77.
Dahler-Larsen P (2011) *The Evaluation Society*. Stanford, CA: Stanford University Press.
Flyvbjerg B. (2006) Five misunderstandings about case-study research. *Qualitative Inquiry* 12: 219–245.
Forss K, Rebien CC and Carlsson J (2002) Process use of evaluations: Types of use that precede lessons learned and feedback. *Evaluation* 8: 29–45.
Freeman R (2009) What is ‘translation’? *Evidence & Policy: A Journal of Research, Debate and Practice* 5: 429–447.
Gannan R, Ciliska D and Thomas H (2010) Expediting systematic reviews: Methods and implications of rapid reviews. *Implementation Science* 5(1): 1–10.
Haddaway NR and Westgate MJ (2019) Predicting the time needed for environmental systematic reviews and systematic maps. *Conservation Biology* 33: 434–443.
Haddaway NR, Kohl C, da Silva NR, et al. (2017) A framework for stakeholder engagement during systematic reviews and maps in environmental management. *Environmental Evidence* 6: 11.
Land M, Macura B, Bernes C, et al. (2017) A five-step approach for stakeholder engagement in prioritisation and planning of environmental evidence syntheses. *Environmental Evidence* 6: 25.
Ledermann S (2012) Exploring the necessary conditions for evaluation use in program change. *American Journal of Evaluation* 33: 159–178.
Leeuw FL and Furubo J-E (2008) Evaluation systems: What are they and why study them? Evaluation 14: 157–169.

Majone G (1989) Evidence, Argument, and Persuasion in the Policy Process. London: Yale University Press.

Miljand M (2020a) The potential of systematic reviews in environmental social science: An analysis of its use to evaluate and inform policy. Umeå University. Available at: https://umu.diva-portal.org/smash/get/diva2:1466314/FULLTEXT02.pdf

Miljand M (2020b) Using systematic review methods to evaluate environmental public policy: Methodological challenges and potential usefulness. Environmental Science & Policy 105: 47–55.

Mistra (2011) Förslag om inrättande av Mistra Council for Environmental Evaluation. Available at: http://eviem.se/wp-content/uploads/2018/03/Pkt-9-Forslag-om-inrattande-av-Mistra-Council-for-Environmental-Evaluation.pdf

Mistra EviEM (2012) Minutes Executive Committee Meeting 3/2012. Available at: www.eviem.se/wp-content/uploads/2018/03/3-Minutes-EviEM-3.pdf

Moher D, Stewart L and Shekelle P (2015) All in the family: Systematic reviews, rapid reviews, scoping reviews, realist reviews, and more. Systematic Reviews 4(1): 1–2.

Oliver K, Lorenc T and Innvær S (2014) New directions in evidence-based policy research: A critical analysis of the literature. Health Research Policy and Systems 12: 34.

Owens S (2015) Knowledge, Policy, and Expertise: The UK Royal Commission on Environmental Pollution 1970-2011. Oxford: Oxford University Press.

Patton MQ (1997) Utilization-Focused Evaluation. Thousand Oaks, CA: SAGE.

Patton MQ (1998) Discovering process use. Evaluation 4: 225–233.

Petticrew M and Roberts H (2006) Systematic Reviews in the Social Sciences: A Practical Guide. Chichester: John Wiley & Sons.

Pullin AS and Knight TM (2003) Support for decision making in conservation practice: An evidence-based approach. Journal for Nature Conservation 11: 83–90.

Pullin AS, Knight TM, Stone DA, et al. (2004) Do conservation managers use scientific evidence to support their decision-making? Biological Conservation 119: 245–252.

Rose DC, Mukherjee N, Simmons BI, et al. (2017) Policy windows for the environment: Tips for improving the uptake of scientific knowledge. Environmental Science & Policy 113: 47–54.

Sackett DL, Rosenberg WM, Gray JM, et al. (1996) Evidence based medicine: What it is and what it isn’t. BMJ 312: 71–72.

Smith V, Devane D, Begley CM, et al. (2011) Methodology in conducting a systematic review of systematic reviews of healthcare interventions. BMC Medical Research Methodology 11(1): 1–6.

Solomon N, Boud D, Leontios M, et al. (2001) Researchers are learners too: Collaboration in research on workplace learning. Journal of Workplace Learning 13: 274–282.

Sutherland WJ, Pullin AS, Dolman PM, et al. (2004) The need for evidence-based conservation. Trends in Ecology & Evolution 19: 305–308.

Vedung E (2015) Six uses of evaluation. In: Hennefeld V (ed.) Nachhaltige Evaluation? Auftragsforschung zwischen Praxis und Wissenschaft. Festschrift zum 60. Geburtstag von Reinhard Stockmann. New York: Waxmann Verlag, pp. 187–210.

Vedung E (2021) The twelve-part typology of evaluation uses. Zeitschrift Für Evaluation 20: 101–130.

Weiss C (1977) Research for policy’s sake: The enlightenment function of social research. Policy Analysis 3: 531–545.

Weiss C (1979) The many meanings of research utilization. Public Administration Review 39: 426–431.

Weiss C (1980) Knowledge creep and decision accretion. Knowledge 1: 381–404.

Weiss C (1982) Policy research in the context of diffuse decision making. The Journal of Higher Education 53: 619–639.

Weiss C (1998) Have we learned anything new about the use of evaluation? American Journal of Evaluation 19: 21–33.
Weiss C (2013) Rooting for evaluation: Digging into beliefs. In: Alkin MC (ed.) Evaluation Roots: A Wider Perspective of Theorists’ Views and Influences. London: SAGE, pp. 130–143.
Wyborn C, Louder E, Harrison J, et al. (2018) Understanding the impacts of research synthesis. Environmental Science & Policy 86: 72–84.
Yin RK (2013) Validity and generalization in future case study evaluations. Evaluation 19: 321–332.

Matilda Miljand is a Researcher in political science focusing on methods for evaluating environmental policy with a specific interest in systematic reviews. Other research interests are climate change initiatives in Sweden.

Katarina Eckerberg is a Professor of political science, focusing on environmental and natural resource policy. Her research topics include European forest policy, collaboration in forest water governance and voluntary forest protection.