Indicator dashboards in governance of evidence-informed policymaking: Thoughts on rationale and design criteria
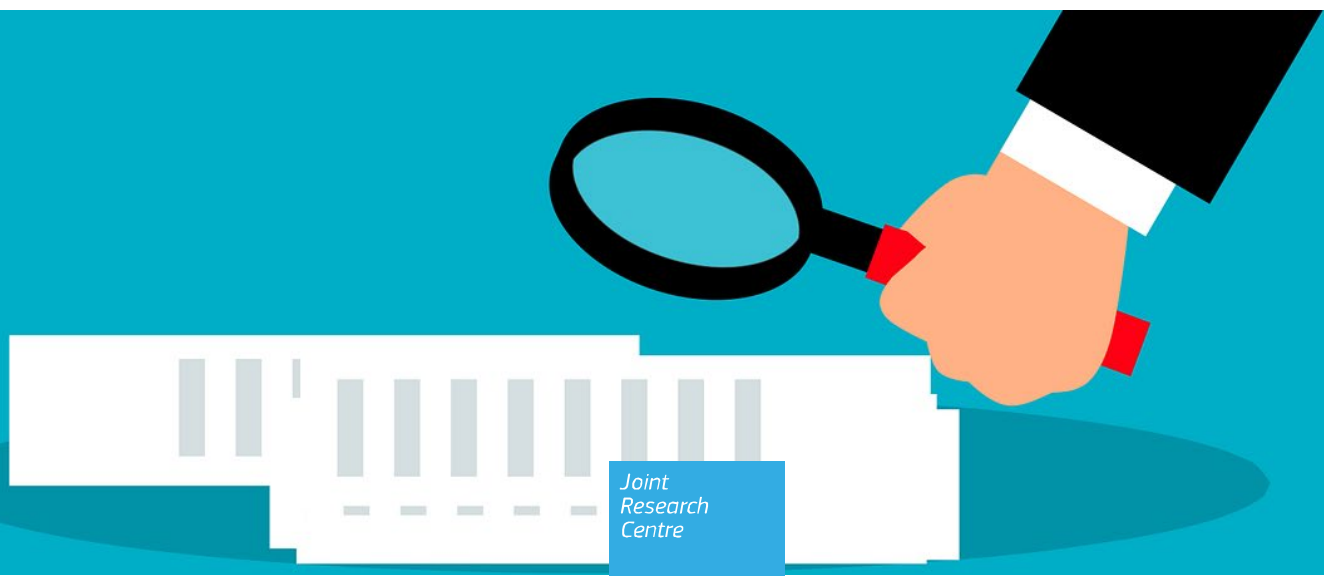
Expert report series: Developing an evaluation framework for science-for-policy ecosystems

Strand, R.

Kristian, K. (editor)
Melchor, L. (editor)

2022

Joint Research Centre

This publication is an input to the "Developing an evaluation framework for science for policy ecosystems" project by the JRC and was previously made available as a draft in preparation of an online workshop on 24th March 2022, organised by the JRC.

Contents

Foreword

In line with its role as the European Commission's science and knowledge service in support of EU policymaking, the Joint Research Centre (JRC) has launched several activities that aim at strengthening and connecting science for policy ecosystems within EU Member States. As policy issues become increasingly complex and interconnected and politics ever more polarised, robust institutions that ensure that scientific knowledge is mobilised, synthesised, translated for, and integrated into the policymaking process become increasingly relevant.

In this work, the JRC benefits from the input provided by professionals working about and/or at the science-policy interface across Europe, from public servants in ministries, government agencies, Parliaments, and government research funding bodies to staff of public and private research institutes, universities, national academies, learned societies, research councils, think tanks, committees, scientific networks, and publishing houses. Through surveys, commissioned studies, and participatory workshops, the JRC seeks to stimulate a vibrant debate about structures, networks, processes, and practices underpinning evidence-informed policymaking across Europe.

To inform and structure the debates about capacity building in support of robust, interconnected science for policy ecosystems, one strand of this work focuses on developing, together with an interdisciplinary group of experts and practitioners, an evaluation framework for the institutional capacity of such ecosystems. Through a combination of commissioned studies, participatory events, and pilot studies, the JRC aims to provide a playbook that support policymakers and other stakeholders at the science-policy interface in designing an evaluation process for the institutional ecosystem that connects scientific research with policymaking processes.

You are about to read one of the studies that we commissioned to develop this playbook. We welcome any feedback that you can share via JRC-E4P-ECOSYSTEM@ec.europa.eu.

We also warmly invite you to join our 1,600+ members strong "Science for Policy Ecosystems" community (https://knowledge4policy.ec.europa.eu/evidence-informed-policy-making/topic/science-policy-ecosystems_en).

Kristian Krieger & Lorenzo Melchor

June 2022

Acknowledgements

Authors

Roger Strand is Professor at the Centre for the Study of the Sciences and the Humanities at the University of Bergen, as well as Co-Director of the European Centre for Governance in Complexity. His research focuses on uncertainty and complexity at the interfaces between science and institutions of governance.

Editors

Kristian Krieger, PhD and **Lorenzo Melchor, PhD** both work as policy analysts at the Knowledge for Policy: Concepts and Methods Unit of the European Commission's Joint Research Centre (JRC). They are responsible for the analysis and capacity building of the institutional foundations of evidence-informed policymaking. They jointly lead and manage the JRC projects on "Strengthening and connecting ecosystems of science for policy across Europe" and "Developing an evaluation framework for science-for-policy ecosystems", under which framework this current discussion paper has been developed.

Summary

The topic of this report is highly convoluted and even complex: namely how to monitor the use of scientific knowledge and evidence in policymaking. The purpose of such monitoring is to contribute to the improvement of evidence advisory ecosystems across Europe. The report takes a philosophical approach by asking about the essential features of evidence advisory ecosystems, or rather the essential features of good evidence advisory ecosystems. It connects that topic to old philosophical debates about what constitutes (good) science and (good) governance, and shows that in European policy discourse, the philosophically obsolete Cartesian Dream of prediction and control and of Science speaking Truth to Power, still finds its expression. Philosophically and empirically updated visions of a good relationship between science and policy are finding their way into policy discourse, however – visions represented by academic concepts such as "honest brokers", "strong objectivity" and "post-normal science", but also vernacular concepts such as transparency and fairness. The report argues for contextuality and pluralism: What constitutes a good relationship and a good evidence advisory ecosystem, depends on the issues at hand and what is at stake, and there is legitimate plurality of views on such questions.

Accordingly, in order to account for and deliberate upon that plurality, it might be useful to introduce a typology and characterise evidence advisory ecosystems in terms of grades, similar to the one-to-five-star system for hotels in Europe:

	Integrity & truthfulness	Relevance & salience	Pluralism, honest brokers	Reflexivity	Coproduction of knowledge
★	✓				
★★	✓	✓			
★★★	✓	✓	✓		
★★★★	✓	✓	✓	✓	
★★★★★	✓	✓	✓	✓	✓

The explanation of these stars and their related norms and virtues is to be found in the main body of text.

Moreover, there are relationships between the desired features of an evidence advisory ecosystem and the suitable approaches to how to govern them. In policy literature on governance, the Cartesian Dream takes the form of command-and-control visions and conventional intervention logic. The report argues that this vision is rather unsuitable for the governance of sophisticated grades of evidence advisory ecosystems. Our discussion of these relationships is summarised in the table below:

	★	★★	★★★	★★★★	★★★★★
Command-and-control	✓	✓	(✓)		
Market governance	✓	✓	✓		
Network governance	✓	✓	✓	✓	✓

A dashboard that could help facilitate governance of evidence advisory ecosystems in European member states, should allow for a plurality of views on the desired type of ecosystem and the choice of governance approach. Such a dashboard could take the form of a checklist. The questions in the checklist can be informed by evidence, and in lucky moments it could even make sense to define indicators for them. The table below gives a taste of what kind of issues the report takes on:

Grade	Examples of questions	Illustrative example
★	— Is the policymaking process informed by scientific knowledge and evidence prior to the choice of policy option? — Is knowledge and evidence presented with integrity and obtained with methods deemed appropriate by scientific peers? — Is there evidence that knowledge inputs can lead to a change in a policy decision? — Do the actors within the ecosystem argue in good faith?	— Were there satisfactory routines and practices to distinguish between peer-reviewed, not peer-reviewed and so-called fast track reviewed evidence during COVID-19?
★★	— Is scientific knowledge and evidence presented, timely, relevant and salient? — Does the policy side actively call for knowledge/evidence that can critically challenge the preferred policy frame or option?	— Was the etiological and epidemiological evidence of the nature of the Omicron variant of SARS-CoV-2 timely introduced into COVID policies in December 2021-February 2022? Was this evidence used to challenge the policy frame of COVID-19 as an exceptionally severe disease?
★★★	— Is the policymaking process organised so as to actively include a plurality of scientific perspectives to minimise bias and partiality? — Do experts act as honest brokers and not as issue advocates? — Do experts actively communicate uncertainty and ignorance? — Do knowledge/evidence inputs lead to an increased range of policy options?	— Was there a balance between knowledge and evidence on short- and long-term consequences to health, society and the economy caused by the COVID-19 disease as much as COVID-19 restricting measures, nationally as well as globally?
★★★★	— Do experts exert reflexivity and identify the impact of their own expertise and the available body of evidence, on how the issues get framed? — Do policymakers actively distinguish between what is known, unknown and unknowable, and consider a variety of sources of justification for their decisions? — Do experts, intermediaries and policymakers express self-doubt and acknowledge the circularity between scientific and political perspectives, and do they deliberate on the doubt together?	— Did COVID modellers account for modelling assumptions with regard to population behaviour, including mechanisms of self-fulfilling or self-destructive prophecies? — Did policymakers consider the possible bias from the knowability of COVID-caused suffering and the relative uncertainty about COVID-measures-induced suffering? — Did experts, intermediaries and policymakers systematically address the impacts of framing COVID mainly as a public health issue?
★★★★★	— Does the policymaking process include the identification and involvement of relevant non-scientific knowledge? — Are relevant non-scientific knowledge holders included in deliberations on the quality and relevance of evidence? — Are methods of co-creation, transdisciplinarity and deliberative democracy employed, to what extent and with which results?	— Were COVID policy decisions made with systematic involvement of those who suffered because of COVID measures (e.g. young adults in good health who chose not to become vaccinated)? — Were the policymaking processes organised so as to prevent stigmatisation, polarisation and social conflict?

When presenting this one-to-five-star typology, it should be emphasized, however, that it is intended as a means for deliberation and not as a ranking system.

1 Why this report? Mandate and background

This report was commissioned as part of the efforts of the European Commission's Joint Research Centre (JRC) to help capacity building processes for evidence-informed policymaking beyond the EU institutions within EU Member States. Specifically, it served as one of three background documents for the JRC workshop "Developing an evaluation framework for science for policy ecosystems", which was linked to the JRC project ENLIGHT-EVIDENCE. The report greatly benefitted from the continuous dialogue with JRC project officers Kristian Krieger and Lorenzo Melchor and experts Ingeborg Niestroy, Kathryn Oliver, and Louis Meuleman. Its content, with the idiosyncrasies, shortcomings and errors it might have, however, is the sole responsibility of the author, Roger Strand.

The specific mandate of this report is to convey insights from critical social research on governance and the science/governance interface, such as they appear in science and technology studies (STS), sociology more in general, and the literature on post-normal science. I have tried to keep the presentation vivid and clear rather than comprehensive. There is no shortage of academic literature on these issues, and I have seen little reason to try to reproduce it in its breadth. Instead, I have tried to state some bold and risky claims that, if they are found acceptable, could make for a change, and if not, might at least spark a lively debate.

The reader should be prepared that the focus of the text is on issues of justification. This is closely connected to my own experience as a scholar but also an expert practitioner giving advice in processes of policymaking and governance. In contexts marked by a sense of urgency and a wish to be useful and practicable, the pressure is often there to focus on "how"-questions. Close to the post-normal science perspective that will be introduced below, I have been convinced that questions **how** to do something have to be accompanied by the corresponding questions about **why** we should do it, with what objective and for what purpose. Not the least this is the case with indicators, which have a tendency of gaining a life of their own and in that process, become proxies that displace or substitute the original goals and targets.

2   What is a "good" science for policy / evidence advisory ecosystem?

2.1   What does "good" mean?

The phenomenon at the very core of our interest in this report is the use of scientific knowledge and evidence in policymaking, and more specifically its **good** (appropriate, legitimate, well-placed) use. This point – that the identification of the phenomenon at hand involves a normative exercise – deserves some initial reflection.

In the natural and social sciences, many topics can be researched and analysed, as well as measured and monitored with the aid of indicators, without the exercise becoming excessively normative. One can measure the abundance of sparrows and magpies in urban ecosystems, or the use of literary references in policy statements, without necessarily taking a strong stance on their desirability or appropriateness. In theory, the same could be done for the use of scientific knowledge and evidence in policymaking. One could empirically describe the inflow of scientific input into policymaking processes and also try to discern and attribute causal effects of such inputs, as one could with any other type of inputs and determinants of these processes, such as lobbying, media attention, educational and class background of policymakers, and so on.

In practice, however, both scholars and practitioners tend to express – or at least relate to – a strong normative interest when it comes to the use of scientific knowledge and evidence in policymaking. Such use is often seen as **good** in itself, or at least good if it is done in the **right** way. If the option is even considered that in a given case, the use of scientific inputs led to a bad outcome or a faulty process, the failure is typically explained by something having gone wrong: The science was of poor quality or of the wrong type; the policymakers were not up to the task to correctly interpret the science, and so on. More than often, it is simply taken for granted that if scientific inputs are duly provided and correctly received and processed, this leads to better policymaking and better outcomes. In EU legislation, this view is explicit, such as in the EC Communication on "Better regulation" (COM(2021)219). The latter even expresses a causal theory to explain the view:

> *Scientific evidence is another cornerstone of better regulation, vital to establishing an accurate description of the problem, a real understanding of causality and therefore intervention logic; and to evaluate impact. (p. 3)*

The philosopher Bernard Williams (2008) noted that we can learn a lot by reflecting on efforts to justify something that is normally taken for granted. First, such efforts might signal an emerging doubt or disturbance of that belief. What is considered necessary and obvious, is rarely in need of explicit justification. Secondly, efforts to articulate a justification may happen to aggravate doubt by involuntarily displaying that the issue is more complex than previously recognised. That which once was perceived necessary, may become demoted to merely being desirable and finally contestable. His example was Aristotle's philosophical efforts to justify slavery.

The issue of indicator dashboards in governance of evidence-informed policymaking is complex in the literal sense of the word: com-plex, **plaited**, that is, with strings that are folded into each other. We have the string already mentioned, that of the use of science in policymaking. Next, we have the string of governing that activity (of using science in policymaking), and doing so by use of indicators, which we can see as a type of evidence, possibly even scientific in nature. This next string is subject to the same type of normative immersion as the first one: Behind the seemingly descriptive concept of governance, there is the normative interest in **good**, well-functioning governance, and the indicators and dashboards-to-be are expected to contribute to that goodness. I shall argue below that ideas and models for the second string (the indicators for the governance) should not be seen in isolation from ideas and models for the first string (the use of science in policymaking). This is also part of the answer why "how" and "why" are related in this case.

Finally, the issues at hand are also linked to unavoidably normative conceptions of science and scientific knowledge itself, and the idea that scientific knowledge in some way is "better" than other types of knowledge, perhaps even the only valid form of knowledge. This means that our already complex issue is entangled into long-standing and notoriously difficult questions in the philosophy of science: What demarcates science from non-science? What is the nature of scientific knowledge? To what extent can science be assumed to provide truth or at least a surrogate for truth? The quote from COM(2021)219 above takes a stand when it claims that scientific evidence contributes to "a real understanding of causality". Taken at face value, this quote could be seen as either an innocent and philosophically ignorant remark or indeed the European Commission taking a strong and radical stand, a minority position, in century-long philosophy debates on scientific realism.

What is important to clarify before we move on to different models of science and policy, however, is that not only philosophers of science but also political scientists, policymakers and governance practitioners have had a variety of opinions on what makes science in itself, and the use of science in policymaking, good. The cited EC Communication takes a substantive, realist position: Science speaks Truth (or its surrogate) to Power; armed with (approximate) Truth, policymakers can device an intervention that, if correctly executed, will have the desired result. So by using scientific input, the outcome is secured. The outcome might perhaps not be perfect in the first attempt if there is uncertainty in the scientific input. However, by continuing the policy cycle and doing rounds of impact evaluation, the process will mimic the scientific method and hence lead to self-correction and progressively better outcomes. This position has been called "**the Cartesian Dream**" of prediction and control (Guimarães Pereira & Funtowicz, 2015). It has a long history in Europe, going back to René Descartes, Sir Francis Bacon and Gottfried Leibniz.

Many would dismiss the Cartesian Dream as indeed a dream or even a dangerous fantasy. Instead, they may argue that the use of scientific knowledge and evidence makes for **neutrality** and better management of bias and conflicts of interest than other knowledge sources. Or if not neutrality, at least **accountability and transparency**. With these less radical positions, the "goodness" of the use of science is not so much an issue of guaranteeing successful outcomes but rather **fair process**, the voice of science becoming less of an oracle and more that of a fair umpire. Sheila Jasanoff (2007) has shown how different countries have different traditions with respect to such positions, a phenomenon that forms part of what she described in terms of different **civic epistemologies**. Scientific knowledge does not play the same role and is subject to different standards in Germany, the UK and the US, according to her analysis. As I shall argue below, such differences have implications for one's model of how – that is, in which ways and to what extent – scientific knowledge should come to use in policymaking.

Moreover, the latter decades of EU and OECD research and innovation policies have advocated for, and also achieved, tighter interaction between academe, industry and government, the so-called "triple helix". The result is a **new social contract of science** in our societies, in which science to a lesser degree can be seen as an ivory tower located at safe distance from greater society with its politicking and commerce. Accordingly, the neutrality of scientific expertise is increasingly often challenged with reference to experts' own combined epistemic and social values, interests and commitments. Before 1962 and the release of Thomas Kuhn's publication "The Structure of Scientific Revolutions" it was philosophically plausible to argue that even if scientists are human beings in flesh and blood, their results are still absolutely objective and can be decided upon by procedures that are entirely disinterested and rational. Since 1962, however, the concept of science as "the view from nowhere" has been increasingly discredited by philosophical analysis as well as historical and sociological evidence. And with science becoming an organic part of the triple helix, **the idea of "the view from nowhere" becomes wholly implausible**. The implication from this learning process is that the question about what it means that science is "good", has to be treated with subtlety.

2.2 Models of science and policy within a logic of justification

One particular source of confusion in questions as grand as that of the role of science in policymaking, is the distinction between **justification narratives** and **questions of performance**. In the excerpt from COM(2021)219 quoted above, the distinction is hard to spot. The reason is that the communication erects a justification narrative that is based in an argument about performance: The right thing to do is to give scientific evidence the place as a "cornerstone" because society will then effectively achieve its goals. We already noted that this is a substantive, realist position, and its most concise expression was provided by Sir Francis Bacon: "Knowledge and power come to the same thing, for where the cause is not known, the effect cannot be produced."

Generally, questions of performance and questions of legitimacy do not fully coincide. Francis Bacon, an outstanding philosopher and statesman of his time, himself fell into public disgrace in 1621 when he was impeached for bribery and removed from his position as Lord Chancellor. Interestingly, Bacon admitted having received gifts but insisted before King James I that his judgements and orders had never been influenced by the bribes, and indeed that he often had ruled against the bribers. This was an argument about performance, one that Bacon repeated several times and, as far as we can know, seems to have firmly believed in.

The problem, however, is that even if it happened to be the case that Bacon's performance as a judge was not influenced by the bribes, to receive bribes was still unacceptable and illegitimate. It is in contradiction with the justification narrative for our courts and other public institutions. At a very general level, there are multiple sources and multiple demands of justification. For instance, some individual decisions are mainly justified in

terms of their moral intent ("good motives") while others are mainly judged in terms of their outcomes. In the modern state, however, the design of public institutions is typically justified by narratives that fall somewhere in between intents and outcomes. Such narratives take the shape of idealised models of how the institution is imagined to function when certain conditions are fulfilled, and how that well-functioning is prone to lead to good outcomes. In this sense, justification narratives within the modern state have the quality of being **imaginaries**. Importantly, these imaginaries serve not only to delineate **how** the state should act, they also serve to justify **that** it can act, giving legitimacy to state action in the vacuum that arose when states became secular and sovereigns no longer could invoke the Grace of God as their ultimate source of justification. In this way, there is a double bond between science and action in the modern state: Scientific evidence gives direction to the decision but more importantly, without support in scientific evidence the decision may fail to be rendered legitimate or even possible.

This reflection allows us to read the COM(2021)219 quote in a more sympathetic light: Not so much as a philosophically ignorant or empirically flawed fantasy but rather as an attempt at setting forth an ideal to strive for, and as a way to argue for the legitimacy of governmental action in the modern state. Still, the question remains if the ideal is the right one. Since Horst Rittel coined the term "wicked problem" in the 1960s, a vast literature has developed on how to govern in the presence of uncertainty, nonlinearity, ambiguity, indeterminacy and other manifestations of complexity. A rethinking of the role of science and scientific knowledge is an essential, integral part of this development. In the words of Silvio Funtowicz and Jerome Ravetz, normal science may have little to offer when facts are uncertain, values are in dispute, stakes are high, and decisions are urgent. Bacon's dictum about knowing the cause and producing the effect is simplistic: It is simply not valid in nonlinear causal networks, not to mention networks where causal agents are conscious beings who themselves interpret and react to acts of governance.

Silvio Funtowicz (2006) summarised the main directions of the debates on science and policy within the European Union into a set of what he (and later we, Funtowicz and Strand, 2007) called models of justification for the relationship between science and policy. The set included five models:

1. The modern model of legitimation.
2. The precautionary model.
3. The framing model.
4. The demarcation model.
5. The model of extended participation.

The **modern model** essentially is what we already called the Cartesian Dream of Science speaking Truth to Power; close to what we still find in a lot of EU official discourse. The three models that follow – of precaution, framing and demarcation – can be seen as representing attempts to save the modern model from challenges of uncertainty, indeterminacy and value-ladenness, respectively.

Discussions about uncertainty developed from the 1970s on, with the German *Bundesrepublik* arguing for the need for a *Vorsorgeprinzip* in environmental management already in 1971. The principle reached the international stage in the Declaration of Rio in 1992, and it can be seen as an attempt to avoid stalemate when scientific evidence is inconclusive. Hardly anywhere is the double bond between science and policy more visible than in the Rio formulation of the precautionary approach. In what has been called the double or triple negative, it never says what governments can or should do, only that in the presence of environmental threats, "lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation". In straight words it concedes that without an additional decision principle, lack of full scientific certainty is indeed a reason for not acting. The problem is, of course, that there is almost always a lack of full scientific certainty in any important socio-economic or socio-ecological issue. The **precautionary principle** carved out a space for the use of scientific evidence in the presence of uncertainty, but the size and power of that space was already curtailed by its own wording as a negative formula, and even more so with its mention of cost-effectiveness. In the presence of uncertainty, the cost-effectiveness of preventive measures is exactly what cannot be robustly estimated. Rather than poor craftmanship by those who worded the principle, it reflects that in contemporary modern societies, the principle of proportionality carries more weight than the precautionary principle. The former is constitutive while the latter is additional and regulative.

The **framing and demarcation models** can be seen as two very different responses to the challenges of indeterminacy and value-ladenness: Of the fact that no scientific result is produced from "the view from nowhere", that different experts and different scientific fields tend to provide different evidence and advice, and that the choice of expertise accordingly in itself is a matter of contestation. The model of demarcation is the attempt to deal with that problem as if it is just a matter of individual vested or ideological interest. It

recommends a purification of expertise by excluding those who have identifiable "conflicts of interest" and ensuring the institutional independence of those who remain, hoping in this way to restore the ideals of absolute objectivity or at least neutrality. The framing model is almost the opposite solution. It acknowledges that every scientific input is partial and carries a perspective, and it argues for the multiplicity and plurality of scientific inputs in order to cancel bias and get a richer picture. Metaphorically, if science enlightens us, what we need are more lamps from more angles to brighten more of the dark spots. This position was taken by the not so frequently cited EC communication, COM(2002)713, that followed the well-known EC white paper on governance. In this communication, the European Commission made bold steps in the direction of acknowledging the necessity to actively frame policy problems as well as the scientific questions to ask about them. It concluded with **pluralism** as a main value in scientific advice. Indeed, COM(2002)713 resonated with philosophical developments such as Sandra Harding's (1993) concept of "strong objectivity", that is, objectivity through the combined strength of a multitude of positions rather than by the illusion of seeing from nowhere.

In spite of its sophistication, the framing model cannot solve the fundamental challenge to the modern model of legitimation, however. Let us again recall how the modern secular state found itself in dire need of an external source of legitimacy for its actions, having severed its ties with an all-knowing and all-loving God. The bet was that Science could take this place as the new supreme knower, if not all-loving at least objective, disinterested and neutral. However, by means of our own scientific faculties (notably in the social sciences and the humanities), experts in the very field of understanding the relationship between science and society came to realize that science is neither all-knowing nor objective in the required sense. It is a partial, fallible, human activity that is culturally, economically, ideologically and politically entangled in the society in which it takes place. **Science is not able to carry the burden of legitimacy of policymaking**. This is a theoretical insight on firm ground that is also mirrored in the world of practice. If we consider several of the grand challenges of our time, such as climate change, ecosystem disruption, loss of wildlife areas and mass extinction of animal species, science has spoken more or less with one voice for decades, but the advice is only marginally translated into action that successfully mitigates the problems. If we take the example of development of antimicrobial resistance, scientific advice could not be clearer and yet is ignored if judged by action. In other areas, scientific advice is chronically ambiguous or contested. Scientific evidence is well received when it fits with the agenda of the receiving end, to the extent that it may be difficult to discern evidence-based policies, in which the policies were shaped and chosen according to scientific inputs, from policy-based evidence, in which sources of evidence were chosen according to the degree to which they supported the already favoured policy option.

The latter 20 years this insight has gradually been flowing into policy discourse, though unevenly so. In explicit, high-level EU policy, it is barely visible. However, concepts such as transdisciplinarity, public participation, public engagement, Responsible Research and Innovation (RRI; see Stilgoe *et al.*, 2013), co-creation and co-production have entered the scene. More than often, they are products of academic literature that take their points of departure in the wickedness of wicked problems and the obsolescence of the modern model and the Cartesian Dream. Funtowicz' fifth model, that of **extended participation**, is an example from the strand of such literature associated with the concept of "post-normal science". Common to most of this literature is the effort to try to envisage solutions that do not postulate counterfactual goals of absolute objectivity or truth (playing the "god trick", as the feminist scholar Donna Haraway (1988) put it) but rather try to work from within the imperfections. Such solutions entail the acknowledgement of uncertainty and value-ladenness as a matter of fact and encourage experiments with deliberative approaches to deal with the imperfections in fair and inclusive manners. Referring again to Haraway, such deliberative approaches should "stay with the trouble".

The immediate implication from such perspectives is that scientific evidence cannot automatically enjoy a special status protected from contestation. Scientific perspectives are often essential to attain a proper description and understanding of the policy problem and its possible solutions, and the facts and values brought to the table by scientists are as such valuable. However, they have to be deliberated upon together with facts and values from other sources. A lot of the recent literature focuses on how to organise such deliberative processes. Other parts of the literature focus on how to present scientific knowledge in a transparent way in order to facilitate the communication of uncertainty and value-ladenness (such as Jeroen van der Sluijs' work on the NUSAP notation). Finally, authors such as Daniel Sarewitz and Roger Pielke Jr have sought to refine the conception of the expert role, for example in the idea of the "honest broker" who does not pretend to know the truth but rather tries to increase the range of policy options to be considered (Pielke Jr, 2007).

As mentioned, a number of these ideas and concepts – above all transdisciplinarity, participation, co-creation and the honest broker – are indeed becoming more present also in policy discourse. For instance, transdisciplinarity is definitely becoming a buzzword after its endorsement in noteworthy OECD publications. Still, as would be expected, the original meaning with its critical diagnosis on the role of science in policy, is rarely retained. More often than not, the modern model is implicitly assumed, now with an addition: *In addition,*

it would be good to have the views of the citizens, the knowledge inputs from non-scientific sources, and so on, perhaps even explicitly stating that this will be good for public acceptance, or to quote a particularly severe formulation of expected impact of a Horizon 2020 Green Deal call: "behavioural change and long-term commitment, trust, social acceptance and buy-in from people, communities and organisations" [1]. And even when the phrases and words imported from the sphere of deliberative democracy are the correct ones, such as participation and co-creation, the practices into which they are introduced may simply be too immersed in an ideology of authoritarian technocracy (Völker & Guimarães Pereira, 2021).

Summing up, it is my belief that these conceptual developments are strongly important when discussing the issue of the use of scientific knowledge and evidence in policymaking. Not only pragmatic questions of performance but the very ideals of why (with which goals and for what purpose) one should use evidence in policymaking are at stake. What becomes very clear from such reflections is that the dictum "the more, the better" does not hold with respect to the use of evidence. We may keep the notion with which I set out: the notion of the good (appropriate, legitimate, well-placed) use of scientific knowledge and evidence in policymaking. However, depending on the context a number of attributes could be proposed as constituting such goodness. From one perspective attributes such as accurate, precise and comprehensive would be emphasized. From another perspective, attributes such as reflexive, dialogical and transparent would be seen as equally if not more important. Some would highlight humility as a main but underrated virtue (Jasanoff, 2003). If we are to devise indicator dashboards to describe and monitor the use of evidence in policymaking, we have to make sure that we do not forfeit these questions by design, by only including indicators that are meaningful from one particular perspective.

2.3   Quality as fitness for purpose

At this stage in the report, I am going to make an important methodological choice. The critical diagnosis is presented, and the reader is reminded that the issue at hand – to describe the use of scientific knowledge and evidence in policymaking, possibly even by use of an indicator dashboard – is complex and that there will be no simple solution to it.

I would like to reject two options. The first option is that of normal science retracting into its shell, stating that the problem definition is immature and that you cannot have indicators for something that is so contested and poorly defined. The second option is that of normal policy work, which is to pretend that the critique never happened, forget about the "why" and "how" questions and frantically look for a "what" instead. I have seen this happen in several sectors, namely that the critical discussion ends in fatigue and stalemate and is replaced by the urge to anyway do and deliver something. "What figures can we actually produce? What statistics are available?" Sometimes this urge results in idiotic indicators such as the number of gender equity plans in research performing institutions or the percentage of research proposals that have been reviewed by an ethics committee. One does obtain numbers, but they carry no sense. At other times, the indicators may appear meaningful – such as circular economy indicators on domestic waste – until one learns that the choice of domestic waste is not because it is the more important category but because the data on industrial waste are unreliable.

The option I am going to choose is grounded in the post-normal science dictum that **quality is fitness for purpose**. Above, I have described a plurality of perspectives on the relationship between science and policy. Both in general and in a given case, an analyst (such as me) should be open to the possibility that this plurality is one of different **legitimate** perspectives. This is what it means to refrain from Haraway's "god trick".

The first dimension of this choice is to allow for **contextuality**. Different member states, different policy fields, and different issues at stake may require different emphasis on the many possible virtues of an evidence advisory ecosystem. Sometimes, there may be a strong need to just get some basic facts straight and carry them through to the decision. At other times, knowledge is called for to increase the range of policy options. And sometimes the need may be felt for "uncomfortable knowledge" (*sensu* Rayner, 2012) that contradicts implicit policy assumptions and challenges institutional divisions of labour.

There are even situations where more information is mostly prone to clutter or derail the decision. For example, in order to inform the European Parliament on the issue of the gender wage gap, the European Commission once presented econometric evidence to support the claim of an arithmetic relationship between reducing the gap and increasing GDP. Critics maintained that such evidence ran the risk of reframing the issue away from a question of discrimination and fundamental rights and into an issue of economic efficiency. The relevant

---

thought experiment would be that it later turned out that the calculations were wrong, and that GDP would increase by increasing the gender wage gap. What would happen then?

The second dimension, however, is even more important, namely that of **plurality**. There may be legitimate disagreement about what is at stake in a given situation, and what virtues of the evidence advisory ecosystem that should be encouraged. To the extent that a dashboard could be constructed to inform that disagreement, it should ideally be able to inform and sharpen the positions across the board and so contribute to a clearer, more enlightened deliberation.

This being said, it would be intellectually dishonest to pretend that the critiques of the modern model of legitimation do not exist and that the Cartesian Dream is as valid a perspective as any other. When the European Commission talks in universal terms about "a real understanding of causality and therefore intervention logic", this is an empirically false description of how things are, and an unsuitable normative ideal for how they might be. An absolute substantive, realist position is not part of the valid range of legitimate positions. At the same time, there are myriads of issues for which there is valid and relevant scientific knowledge is available, and in which the aspect of the value-ladenness of the knowledge is not so salient. If we avoid the big capital letters in Science and Truth, it is definitely possible to reconceptualise the modern model as one of science speaking *truthfully* and with integrity, to power.

2.4   Vision: Evidence stars

I will proceed to play with ideas about what actually could go into a dashboard. In what follows, I will elaborate on an idea that was originally suggested by Silvio Funtowicz, which is to distinguish between quality and grade. Grade is an aspect of quality and should not be confused with quality in general. Let me explain:

In several European countries, many hotels are associated with the Hotelstars Union. Depending on the services and commodities they offer, these hotels can boast with a number of stars, from one to five stars. On average, one would expect a five-star hotel to be more expensive and luxurious than a four-star hotel, and so on. The Hotelstars Union governs this classification scheme by use of detailed sets of criteria[2]. Every individual hotel is then awarded its number of stars according to the commodities of the rooms (TV set, private bathroom, hair drier and many others) and the reception services (such as having a sofa in the reception area).

Combined with (own and others') customer experience, the star system nicely allows for contextuality and plurality while still providing some rigour. A two-star hotel can be excellent (in its grade) but does of course not offer all the services and commodities of a four-star hotel. Still, perhaps for a particular guest on a particular trip, say a short business trip without the need to have meetings at the hotel, the two-star hotel is exactly what the guest needs. For other purposes, a four-star hotel may be desirable. While planning a family trip, family members may initially disagree on which services and commodities they need. The Hotelstars Union descriptors might then serve to clarify the issues at stake and narrow down the choices.

Importantly, however, there is an experiential dimension to quality that eludes that of grade. I already touched upon this: a two-star hotel may be excellent within its grade, but the nature of that excellence might be hard to formalise. It could be a matter of location, friendliness of staff, architectural beauty or good feng shui. The stars cannot tell you this.

I propose a star system that distinguishes different grades of evidence advisory ecosystems. It would be similar to the Hotelstars Union in that the one-star grade is the most basic one (a room with a bed) and then we gradually add features and attributes up to five stars. To repeat: Quality is here conceived as fitness for purpose, so it is not taken for granted that five stars are "better" than two stars. It depends on the purpose.

---

2 See https://www.hotelstars.eu/criteria/

Table 1. Features and attributes of evidence advisory ecosystems at different grades (stars).

	Integrity & truthfulness	Relevance & salience	Pluralism, honest brokers	Reflexivity	Coproduction of knowledge
★	✓				
★★	✓	✓			
★★★	✓	✓	✓		
★★★★	✓	✓	✓	✓	
★★★★★	✓	✓	✓	✓	✓

A **one-star evidence advisory ecosystem** would satisfy the requirements of our modified, lower-case modern model: Science speaks truthfully and presents valid results obtained with methods that are deemed appropriate by scientific peers. The experts transparently declare their potential conflicts of interest, in line with the above mentioned demarcation model. Intermediaries and boundary organisations do not see themselves as issue advocates and ascribe to an ideal of transparent, non-biased, non-distorting transfer of information from the very best science (as judged by the scientific community itself) to the decision-makers. Moreover, there is integrity also at the receiving end of the advice. Policymakers ask themselves if the evidence contradicts their prior policy conclusions and take appropriate action if it does. Actors argue in good faith and are willing to change their minds in the presence of a better argument.

A **two-star ecosystem** would additionally be equipped – institutionally and by the work of its intermediaries – so as to promote timeliness, relevance and salience of advice, in line with what Pielke (2007) called "science arbiters". In the example above, someone would have filtered out the econometric calculation of the effect on GDP on the gender wage gap. A main motivation for requesting and providing advice is not to gain support for an already chosen policy option but to critically challenge and test it against empirical evidence. The ecosystem would embody Karl Popper's slogan of "I may be wrong and you may be right, and by an effort, we may get nearer to the truth" and aim for robust decisions by means of this ideal of organised scepticism, also by employing continuous learning through impact assessments in the course of the policy cycles.

A **three-star ecosystem** would additionally take on the challenge that scientific knowledge is uncertain and value-laden and that experts do not see from "the view from nowhere". Experts would not identify (explicitly or implicitly) as issue advocates but see their role as honest brokers who help policy also by widening and enlightening the range of policy options. Intermediaries would enact the framing model, actively ensuring a plurality of different scientific perspectives. Experts would also be able to understand and communicate the sources of uncertainty in the evidence as well as its implications for decision-making. Policymakers would consider the need for precaution in the absence of scientific certainty and would be able to deliberate policy options in light of the whole body of partial, uncertain and possibly incommensurable evidence.

A **four-star ecosystem** would additionally institutionalise reflexivity and humility. Experts would be epistemologically reflexive and acknowledge that however much they try to be honest brokers, they cannot escape the profile and content of their own expertise and their epistemic communities, and that they run the risk of being stealth issue advocates without realizing it themselves. This insight would make them humble in moral and political terms, being aware that the nature and limits of their own expertise contribute to framing the set of policy options in ways that can even be harmful without their knowing. As reflexive experts, they would aim to articulate not only facts but also the values held by themselves and their epistemic communities, and try to indicate their blind spots. Intermediaries would help experts to admit ignorance and to explain how implicit values and background assumptions of their discipline conditions the evidence they bring to the table. Boundary organisations would have their own expertise in critiquing expertise and would be able to perform reflexive analysis when the experts themselves are unwilling or incapable of doing it. On the receiving end, there is a similar sensitivity to the limits of scientific evidence and the need for political judgement on how to frame the issue and thereby the range of relevant expertise. Policymakers are aware that the consequences of a given decision may be impossible to predict and are able to justify decisions by other principles (precaution, fundamental rights, virtues, etc). Finally, they are themselves reflexive and humble, being aware that their

framing choices can be of great importance and at the same time lack foundation or even be built on implicit, unsupported epistemic beliefs. In an ecosystem characterised by reflexivity and humility, experts, intermediaries and policymakers would be allowed and encouraged to express and humbly deliberate on their self-doubt.

A **five-star ecosystem** institutionalises inclusiveness and extension of the peer community to facilitate co-production of knowledge and values. It recognises the need to leave the echo chamber of scientific and political elites in order to ensure a broader, democratic legitimacy of the framing of issues and the corresponding framing of what counts as relevant evidence. It identifies and employs non-scientific sources of knowledge to the extent that they are timely, relevant and salient, and facilitates transdisciplinary (in the OECD sense of going beyond academe) and trans-sectorial deliberation on evidence when needed. Importantly, such questions about timeliness, relevance and salience are answered by processes that involve deliberative democracy and public participation, rather than leaving them to scientific and political elites. This means in practice that the self-doubt from the four-star ecosystem is actively exposed to the public (to the extent that the public did not already know) and brought to the table. Participants in the extended peer community are by default acknowledged as both knowledge- and stakeholders. Organisers of deliberative processes do not push for consensus or expediency. Participatory processes to frame issues are acknowledged as meaningful institutional venues for the expression and potential resolution of controversy and social conflict.

Grade does not exhaust the question about quality. A hotel may have a sophisticated sauna, but the sauna will be of little value if it is uncomfortable, noisy and filthy. Likewise, sophisticated features such as pluralism, reflexivity and transdisciplinarity can easily become perfunctory and take the shape of window-dressing or even parody. Just as with the Hotelstars, there will always be a need to consider the dimension of intrinsic quality in addition to the grade. This inescapably requires the use of judgement. It also raises questions as to what kind of information can be gathered to characterise and evaluate an evidence advisory ecosystem. The number of stars in the doorway says something about hotels. Trip Advisor is a useful supplement to learn about the intrinsic qualities, especially if one wishes to act at a distance and without first-hand knowledge. If the family members have own experience, however, they can deliberate themselves on the quality. This leads us directly to the relationship between quality and models of governance.

3   Models of governance of the evidence advisory ecosystem

I remarked above that the issue of indicator dashboards in governance of evidence-informed policymaking can be thought of as complex in the literal sense, of having strings folded into each other. So far, I have mostly discussed only one string in the braid, namely that of the phenomenon of evidence-informed policymaking. I shall now turn to the second string of **governing that activity** (of using science in policymaking).

As with the first string, this topic is also frequently studied and engaged with from a perspective that combines descriptive and normative interest. The intrinsic value load in the concept of governance is deep and on a par with science and knowledge, and I suspect it to have linguistic roots. Poor, invalid, false knowledge is not knowledge. The same can be said for the etymological root of governance, which is navigation. To lose control of a ship and run it aground is not to navigate. It is failing to navigate.

The literature on governance is even vaster than that on science for policy, and the respective issues are entangled into each other. Again, it would be futile to try and comprehensively represent the developments in that literature in a report such as this. However, I believe it is of relevance to our topic that there are different visions and approaches to what governance is and should be, and I will make some elementary remarks about these approaches in what follows. The purpose of this exercise is to show that some visions of governance are more suitable to deal with the full range of "evidence stars" presented above.

3.1   Governance as command-and-control

When one analyses the policy discourse in the European Commission, one finds frequent claims that indicate a notion of governance as a matter of command and control. The references may be implicit or, as in our infamous quote from the Communication on Better Regulation, even explicit:

…establishing an accurate description of the problem, a real understanding of causality and therefore intervention logic (p. 3)

This conception of governance is the classic one, strongly related to the Cartesian Dream: There is a governance actor (an institution or a person) who is able to objectively define the problem and the system in which the problem resides. Next, the actor decides on an intervention on the system that fixes the problem. This style of reasoning and acting has many names – conventional intervention logic, command-and-control, predict-and-control – and has a long history in the military sector as in engineering, as well as other social and technical systems where complexity is constrained by use of discipline or other means to control actors and their environment. In recent years, the popularity of the acronym SMART ("specific, measurable, attainable/achievable/assignable, realistic/relevant, timely") is yet another manifestation of the command-and-control approach to governance.

Mike Power (1999) warned against the increase in "rituals of verification" due to the strengthened emphasis on accountability and control. Evidence advisory ecosystems are populated by high-skilled knowledge workers and civil servants whose job description is to exert a high degree of autonomous judgement. Subjecting their practices to a command-and-control type of governance, which typically results in prescriptive rules and procedures, seems like an unwise choice. The reason is that expert judgement is a type of faculty that cannot be reduced to rule-following without compromising its quality (Dreyfus & Dreyfus, 1986). Moreover, I have argued above not only for the role of contextuality but also of plurality when dealing with the issue of quality in evidence advisory ecosystems. There may be no uniquely defined, authoritative definition of the governance problem to be addressed, and even if there were, it may become contested together with the authority that defined it. Indeed, deciding on the governance problem may be a major part of the governance challenge. Moreover, audit cultures are vulnerable to the problem of self-reference and infinite regress that was stated in its classical form by Juvenal almost two thousand years ago: *Quis custodiet ipsos custodes?* Who guards the guardians, who audits the auditors? All of these considerations lead us directly into other conceptions of governance.

3.2   Market and network approaches to governance

Jones *et al.* (1997) defined network governance as "coordination characterized by informal social systems rather than by bureaucratic structures within firms and formal contractual relationships between them". While this is not the only definition or usage of the term, network approaches to governance have their descriptive and/or normative focus on collective efforts within networks to the effect of setting and pursuing goals and steering action. For instance, the EU-funded project SIAMPI studied what they called productive interactions between

science and society in order to reach an understanding of how to improve governance of publicly funded research and innovation. It is useful to contrast this vision of governance with "SMART". It may sound convincing to call for goals to be "specific", "realistic" et cetera. However, the problem is that a command-and-control approach presupposes that it is possible (and wise) for a governance actor outside or above the system to be governed, to set specific goals in advance and to know what is realistic. In a way, the mantras of SMART carry the premonition of their own failure: If the system to be governed is one of creative and competent human actors, preset, specific goals may actually render the endeavour less realistic and timely and also less ambitious. From the perspective of a regulator, the choice of a network approach to governance is accordingly to some extent the choice to trust the integrity and competence of the actors in the network. Or alternatively, for the less trusting ones, to admit that command-and-control approaches risk so much resistance and subversion that it is better to grant the actors some autonomy.

Market governance approaches are as diverse as network approaches, and may for our purpose be seen as a category somewhere in between the Yin and Yang of networks and intervention logics. Furthermore, they can be recognised as individual elements introduced into a larger governance scheme. At their core, they adopt market heuristics in order to understand the system to be governed and use those heuristics to define rules or procedures that encourage or direct behaviour towards improved performance. In its broadest definition, market governance can encompass attempts at installing market mechanisms of competition (such as the use of tenders in the evidence advisory ecosystem) or equivalents to market regulation, such as incentives for collaboration, division of labour, voluntary standards, et cetera. Market governance approaches can be adopted with a "SMART" conceptual frame, setting goals in advance and then adjusting governance instruments with reference to their results, or within a more decentralised network perspective that encourages the system to be governed, to develop and grow on its own.

3.3 Governance *in* complexity as opposed to governance *of* complexity

A theoretical insight that permeates this part of the report and will be crucial for my final conclusions, is the distinction between governance in complexity as opposed to governance of complexity. The distinction is due to Arie Rip (who, in the name of accuracy, called it governance in and of reflexive modernity rather than complexity). Rip (2006) wrote:

> *…political actors, and more generally, actors with a governance responsibility, will see themselves as somehow outside the system that they have to govern. This is almost unavoidable: to articulate a strategy, one has (so it appears) to diagnose a situation 'out there' and formulate a response. In addition, governance actors will be held accountable for what they set in motion, and for that reason they will be positioned (and will position themselves) as independent of the system that is being governed. The effects of their strategies, however, are determined by ongoing dynamics outside their influence, and by the response of other actors to the strategies of the governance actors. (p. 82)*

At this stage it may be wise to stop and reflect again on our combined descriptive and normative interest in the issues of governing evidence advisory ecosystems. The quote from Arie Rip's text mainly states a descriptive point: Unless you happen to enjoy a monopoly of violence and you are the commander of a highly disciplined boarding school, army or concentration camp, the governance you exert will as a matter of fact be governance in complexity and not just governance of complexity. You will not be in control, and the Cartesian Dream is a fantasy.

This matter of fact does not imply, however, that it necessarily is a faulty normative choice to opt for SMART objectives or set out to define an intervention logic in a given case. That depends on the context. It only implies that the choice is fallible and that one has to expect that things will not go entirely according to plan.

There is a subtler normative implication here, though, that can be drawn as we shall know revisit our Evidence-stars grades and connect them to governance approaches. Attempts at governance of complexity, that is, attempts at governing a system from its outside – and in particular those related to command-and-control and SMART objectives – are usually attempts at reducing, curtailing or controlling complexity. Nonlinearities, higher-order causal interactions and emergent phenomena at system scale are usually undesirable from a governance of complexity perspective because they are in principle unpredictable. While also a distant governor-of-complexity in principle could encourage emergence and "let the hundred flowers bloom", to paraphrase Chairman Mao, SMART is incompatible with such strategies because in order to be SMART one has to specify in

language (or numbers) in advance what one wants. This is indeed the deepest reason why SMART implies a type of governance that will try to reduce complexity: SMART forces action to be precisely defined in what Daoist philosophers called the realm of the **names**, which is in cognitive categories such as words and numbers. The actual system, however, resides in the **Dao**, that is, the real, material world itself, which invariably has features that remain unknown to us. This is also why such strategies are fallible, or to use the words of the 2,500 years old classic *Dao De Ching*: "Those who control, fail." (Lao-Tzu)

From the Daoist perspective, governance should accordingly be light-handed, searching for synergies with already existing forces and courses of development. In this tradition, good governance is often explained with water metaphors. It is better to let the river flow (while perhaps subtly diverting its original direction) than to try to push the water upstream. Also, good governance is sometimes described as having the quality of water. Water can penetrate and spread without creating shocks and impacts, and it can nourish rather than force.

Revisiting our Evidence-stars, the four-star ecosystem is defined by reflexivity and humility, that is, by norms and virtues that show themselves in experts' and policymakers' use of judgement. I believe it is going too far to say that such judgement resides only in the Dao and cannot be represented in the realm of the names; that would imply that expert judgement in principle is inaccessible to explication and accordingly cannot be held accountable (although this position has also been argued for in the philosophical study of tacit knowledge as an element of expertise). What seems clear, though, is that the quality of an act of judgement cannot be separated and assessed independently from the content matter of that act of judgement. If an expert concludes that his particular predictive model of virus spread is too unreliable to apply in a decision on a prolongation of COVID-19 measures, and a policymaker decides to use it anyway with reference to the precautionary principle, the quality of these two acts of judgement cannot be meaningfully assessed at a distance without entering into the actual content of the argument. Accordingly, one cannot hope to govern such an ecosystem well from a distance by defining an intervention logic or procedural rules aimed at installing reflexivity and humility. What one could hope to achieve, is a procedural mandate to perform some act of deliberation. Whether it would actually be reflexive and humble, and not just a perfunctory exercise to comply with the mandate, would be inaccessible from the formal, distanced point of view.

If I am right in this argument, command-and-control and market governance approaches are unsuitable to govern well a four-star evidence advisory ecosystem, and even less so a five-star ecosystem which in addition is defined by encompassing plurality as part the practice itself. The five-star ecosystem is in that sense an instance of network governance, and it would seem that the Daoist ideal of being like water is not the worst way in which a regulator could address it.

The same argument could perhaps also be made for a one-star ecosystem, which indeed is defined in terms of the integrity and truthfulness of the actors. On the other hand, it may be argued that the scientific institutions have more to offer in terms of checks and balances on integrity and truthfulness, in terms of peer review and expert consensus. At least this is considered to be part of the social contract of science: That science is self-correcting and self-governing to the extent that citizens, policymakers and external governors should be able to verify signs of integrity without having to go all into the content matter. In the table below I summarise the discussion of the suitability of governance approaches with respect to Evidence-stars grades. In it, I give command-and-control and market governance approaches the benefit of the doubt in the case of one to three stars. I shall not argue the decision for every square in that table; a consideration such as this is of course subject to contextuality and plurality and in no way to be taken as exact science.

Table 2. Suitability of governance approaches with respect to Evidence-stars grades of evidence advisory ecosystems.

	★	★★	★★★	★★★★	★★★★★
Command-and-control	✓	✓	(✓)		
Market governance	✓	✓	✓		
Network governance	✓	✓	✓	✓	✓

4 Indicators dashboards

At this stage, some readers may not only feel a growing concern but also a strong sense of déjà vu. The task to be solved, was to help define indicators for use in governance of evidence-informed policymaking. However, indicators, as an element of governance, most of all belong to the conceptual universe of command-and-control and conventional intervention logic. And above I just concluded that that conceptual universe is definitely unsuitable for governing the more sophisticated grades of evidence advisory ecosystems, and only by exercising the benefit of the doubt did I find it suitable for the more basic grades.

So how to deal with the question of indicators then? There is even a déjà vu here to a previous remark within the report, about reverting to becoming an operator within what I called normal policy work: Pretend that the critical analysis never happened and scramble together some data that can be dressed up as indicators.

My guess is that readers will recognise the following three possibilities in this kind of situation:

First, there may be a denial of the critique or simply a failure to understand it. In both cases, there may be a reversion to the hegemonic discourse, along the following lines: The European Union is paving the way for Better Regulation and the Member States will be greatly served by indicators that can help them develop their evidence advisory ecosystems to the highest of standards where Science always is allowed to speak Truth to Power and where irrationality and arbitrariness is conquered. This type of discourse can then be accompanied with numerical indicators about the volume of scientific evidence presented and the degree of compliance in policy with expert recommendations. Command-and-control governance can be deployed to maximise such indicators and with luck, establish and improve the quality of a one-star or two-star evidence advisory ecosystem. It will just be a matter of asking for more science and more evidence, everywhere, perhaps with filters for relevance and salience. As for scientific quality, one may simply rely on the mechanisms of quality assurance within scientific institutions themselves. One is likely to be confronted with contestation, dissent and all that is known from science and technology studies and post-normal science since the 1980s and 90s, but from inside the hegemonic discourse it will just be dismissed as bias and irrationality. And in the meantime, more indicators can be deployed.

Secondly, there may be a response that resembles the first one but only on the surface. I have previously described this phenomenon as **desperate modernity** (Strand, 2002). Experienced scientists and policymakers are often very well aware of uncertainty and complexity as well as of the limitations of conventional intervention logic. They know that the Cartesian Dream is a fantasy. Yet, exactly because of their awareness of complexity, they hesitate to criticize and destabilize the justification narratives in the modern state. The narratives may be incoherent, false and intellectually bankrupt but still they play an important role in preventing the collapse of modern institutions with a resulting social and political chaos. Accordingly, they consider it unsafe to publicly admit that science does not provide "a real understanding of causality", or that economic growth cannot be decoupled from resource use, or that the economy never can become circular, to use some of the examples I have encountered over the years. It is known that the Emperor has no clothes on, but it must not be said. Instead, such scientists and policymakers choose to work within the institutions, actually not without a Daoist sense of reflexivity and humility, going with the flow of the river but softly trying to divert it by piece-meal additions of transdisicplinarity, co-creation, precaution and so on. By their words, they may sound like they promote command-and-control governance but these words are a tactical element in deeds that amount to network governance. In terms of indicators, there would be the tactical inclusion of what is desired by those who deny or fail to understand the critique – give them what they want! However, there would be additions that do not really serve a SMART type of intervention logic but rather are designed to give visibility and symbolic value ("performativity" in STS jargon) to the good cause even if the indicator is statistically unsound or not even measurable. A nice example is the key performance indicator of Responsible Research and Innovation in DG RTD, which is **institutional change** in research performing organisations. Similarly, one could throw into our dashboard indicators such as "presence of co-creation", "degree of expert reflexivity", "range of policy options" et cetera, to give tags and institutional value to what one believes to be important norms and virtues.

Finally, there is the position of the radical optimist who believes that indicators (in the widest sense of the word) might be useful as knowledge inputs in conscious effort of network governance to improve evidence advisory ecosystems. The purpose of such indicators (again, taken in the widest sense) would be very close to the mandate of the JRC ENLIGHTEN-EVIDENCE project: not only to inform but also to facilitate dialogue and sustain processes of mutual learning and ongoing self-evaluation. Such indicators can be ironic, in the sense that they at first sight appear to serve the hegemonic discourse of command-and-control, but when properly used, the user (or spectator) discovers that the discourse or practice is implausible, unfair or ill-informed. An example could be counting the number of stakeholders in participatory processes as an indicator of participation. Or they

could be serious attempts at accurate descriptions, such as the impact narratives developed in the FP7 SIAMPI project.

4.1 Indicate, monitor, assess or narrate?

A tactical choice in such considerations is whether to use the word "indicator" at all. In the monitoring and evaluation guide for the Horizon 2020 project TRANSFORM, we drew heavily on Theodore Porter (2015) and wrote:

> *Indicators are only one type of tools, and it is in principle wholly possible to monitor and evaluate without using indicators.*

Indicators are guides for certain types of action. To quote Theodore Porter, a world-leading historian of statistics:

> *"Etymologically, an indicator, like an index, has to do with pointing. [...] indicators detect, point or measure, but do not explain. [...] A quantitative index or indicator typically cannot measure the very thing of interest, but in its place something whose movements show a consistent relationship to that thing. Since its purpose is merely to indicate as a guide to action, ease of measurement is preferred to meaning or depth." (Porter, 2015, p.34)*

Often, indicators (Porter continues) are:

> *"pursued to promote informed action and decisions of a decentralized sort." Historically, they were created to efficiently provide information at a distance to decision-makers who lack first-hand knowledge. (Strand et al., 2020)*

In conclusion, indicators are not really the knowledge input needed in network governance where the actors indeed do have first-hand knowledge but need to develop their **understanding** of the system within which they operate.

In the real world, however, there will be a plurality of views, ranging from those who do not understand these considerations and profoundly wish for some quantitative indicators to govern by; to the desperate modernists who know that they cannot fight for their good cause without some seemingly legitimate quantitative indicators to show for them; to the radical optimists who want ironic indicators and qualitative narratives. And again, we should admit for contextuality and that we as analysts cannot know by default who has the "right" position in a given case.

This means that **the dashboard under consideration must allow for all of this**. It should be able to provide all relevant end-users with the type of perspectives they need, across the range of Evidence-star grades and modes of governance.

At this point, however, the author of this report has to come out clean: I have been part of such a venture before, namely the EC Expert Group that was supposed to define indicators for Responsible Research and Innovation (Strand *et al.*, 2015). The line of thought of that work was very similar to the one in this report, and we ended up proposing a whole suite of indicators of all sorts: measurable and unmeasurable, naïve and ironic, quantitative and qualitative, conventional and radical. To my knowledge, none of these indicators are in actual use. DG RTD shelved the report and commissioned (in parallel) another study that produced a set of quantitative indicators, the so-called MORRI indicators, which did satisfy the desires of the Cartesian Dream. Instead, the report we produced, ended up as a source of intellectual arguments to be cited by progressive academics and policymakers but not as a source of something that actually could be implemented.

There are probably many explanations that can be given to this example of failure in practical terms (Strand and Spaapen, 2021). One of them is that the mentioned expert group, however well-versed in the theory of co-creation and participation, did not have the opportunity to actually engage and co-create with the real end-users of the potential product. I suspect that the readers of this report can confirm that in their fields too, lots of indicators have been proposed, very few are populated with data and there are not many who actually use many of the proposed indicators, at least not by their own free choice. In some EU-funded projects (e.g. NEWHORRIZON and SUPERMORRI) one can see one type of development to ameliorate this situation, by placing less focus on indicators and rather develop dashboards that serve as self-assessment tools that can facilitate learning and dialogue.

4.2   What about a list of questions?

In spite of my poor track record with getting indicators into use, let me end this report with a suggestion that I believe in but that is going to be met with the argument that it is not technical and is too simple to sound as real (social) science: Perhaps the dashboard could be structured as a checklist, a list of questions related to the Evidence-star grades. The idea would be that a member state, or a boundary organisation in a member state, or an individual expert or policymaker, could run the checklist and ask questions such as: Do we need (qualitative or quantitative) studies to answer the questions? Who should we discuss with? What implications should be drawn from our answers to the questions, and who should be invited to take part in deciding on those implications?

I conclude with a table (See Table 3, next page) that illustrates the type of checklist one could have. Since the Commission already in 2021 decided to celebrate Europe's policy performance during COVID-19 (COM(2021)380), I thought it could be good exemplify such a checklist with some possible interesting questions about the interface between scientific knowledge and policymaking during the pandemic.

Table 3. The simplest possible dashboard: A list of questions

Grade	Examples of questions	Illustrative example
★	— Is the policymaking process informed by scientific knowledge and evidence prior to the choice of policy option? — Is knowledge and evidence presented with integrity and obtained with methods deemed appropriate by scientific peers? — Is there evidence that knowledge inputs can lead to a change in a policy decision? — Do the actors within the ecosystem argue in good faith?	— Were there satisfactory routines and practices to distinguish between peer-reviewed, not peer-reviewed and so-called fast track reviewed evidence during COVID-19?
★★	— Is scientific knowledge and evidence presented, timely, relevant and salient? — Does the policy side actively call for knowledge/evidence that can critically challenge the preferred policy frame or option?	— Was the etiological and epidemiological evidence of the nature of the Omicron variant of SARS-CoV-2 timely introduced into COVID policies in December 2021-February 2022? Was this evidence used to challenge the policy frame of COVID-19 as an exceptionally severe disease?
★★★	— Is the policymaking process organised so as to actively include a plurality of scientific perspectives to minimise bias and partiality? — Do experts act as honest brokers and not as issue advocates? — Do experts actively communicate uncertainty and ignorance? — Do knowledge/evidence inputs lead to an increased range of policy options?	— Was there a balance between knowledge and evidence on short- and long-term consequences to health, society and the economy caused by the COVID-19 disease as much as COVID-19 restricting measures, nationally as well as globally?
★★★★	— Do experts exert reflexivity and identify the impact of their own expertise and the available body of evidence, on how the issues get framed? — Do policymakers actively distinguish between what is known, unknown and unknowable, and consider a variety of sources of justification for their decisions? — Do experts, intermediaries and policymakers express self-doubt and acknowledge the circularity between scientific and political perspectives, and do they deliberate on the doubt together?	— Did COVID modellers account for modelling assumptions with regard to population behaviour, including mechanisms of self-fulfilling or self-destructive prophecies? — Did policymakers consider the possible bias from the knowability of COVID-caused suffering and the relative uncertainty about COVID-measures-induced suffering? — Did experts, intermediaries and policymakers systematically address the impacts of framing COVID mainly as a public health issue?
★★★★★	— Does the policymaking process include the identification and involvement of relevant non-scientific knowledge? — Are relevant non-scientific knowledge holders included in deliberations on the quality and relevance of evidence? — Are methods of co-creation, transdisciplinarity and deliberative democracy employed, to what extent and with which results?	— Were COVID policy decisions made with systematic involvement of those who suffered because of COVID measures (e.g. young adults in good health who chose not to become vaccinated)? — Were the policymaking processes organised so as to prevent stigmatisation, polarisation and social conflict?

References

COM(2002)713, 2002, Communication from the Commission on the Collection and Use of Expertise by the Commission: Principles and Guidelines "Improving the knowledge base for better policies", 11 December 2002, COM(2002) 713 final, available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2002:0713:FIN:EN:PDF%20, (accessed 8 June 2022).

COM(2021)219, 2021, Communication from the Commission to the European Parliament, the European Council, the European Economic and Social Committee and the Committee of the Regions. "Better regulation: Joining forces to make better laws", 29 April 2021, COM(2021)219 final, available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2021:219:FIN, (accessed 8 June 2022).

COM(2021)380, 2021, Communication from the Commission to the European Parliament, the European Council, the European Economic and Social Committee and the Committee of the Regions. "Drawing the Early Lessons from the COVID-19 Pandemic", 15 June 2021, COM(2021) 380 final, available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM%3A2021%3A380%3AFIN, (accessed 8 June 2022).

Dreyfus, H. L., and Dreyfus, S. E., 1986, Mind over Machine. The Power of Human Intuition and Expertise in the Era of the Computer, Free Press, New York, NY.

Funtowicz, S., 2006, 'Why Knowledge Assessment?' in Pereira, A. G., Vaz, S. G., and Tognetti, S. (eds): *Interfaces between Science and Society*, chapter 8, pp. 138-145, Routledge, London (UK).

Funtowicz, S., and Strand, R., 2007, 'Models of Science and Policy', in Traavik, T. (ed): *Biosafety First-Holistic Approaches to Risk and Uncertainty in Genetic Engineering and Genetically Modified Organisms*, chapter 16, pp. 263-78, Tapir Academic Press, Trondheim (Norway), JRC42711

Guimarães Pereira, A., and Funtowicz, S. (eds.), 2015, *Science, Philosophy and Sustainability. The End of the Cartesian dream,* Routledge, London (UK).

Haraway, D., 1988, 'Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective', *Feminist Studies*, Vol. 14, No. 3 (Autumn 1988), pp. 575-599, doi: 10.2307/3178066.

Harding, S., 1993, 'Rethinking Standpoint Epistemology: What is 'Strong Objectivity'', in Alcoff, L., and Potter, E. (eds): *Feminist Epistemologies*, Routledge, London (UK), pp. 49-82.

Jasanoff, S., 2003, 'Technologies of Humility: Citizen Participation in Governing Science', *Minerva*, Vol. 41, pp. 223-244, doi: 10.1023/A:1025557512320.

Jasanoff, S., 2007, *Designs on Nature,* Princeton University Press, Princeton (NJ).

Jones, C., Hesterly, W. S., and Borgatti, S. P., 1997, 'A General Theory of Network Governance: Exchange Conditions and Social Mechanisms', *The Academy of Management Review*, Vol. 22, No. 4 (Oct. 1997), pp: 911-945, doi: 10.2307/259249.

Lao-Tzu, ~6th century BC/1993, *Tao Te Ching*, Hackett Publishing Company, Cambridge (MA).

Pielke Jr., R., 2007, *The Honest Broker Making Sense of Science in Policy and Politics,* Cambridge University Press, Cambridge (UK), doi: 10.1017/CBO9780511818110.

Porter, T., 2015, 'The flight of the indicator', in: Rottenburg, R., Merry, S. E., Park, S-J., and Mugler, J., (eds): *The world of indicators: The making of governmental knowledge through quantification*, chapter 2, pp. 34-55, Cambridge University Press, Cambridge (UK), , doi: 10.1017/CBO9781316091265.002

Power, M., 1999, *The Audit Society: Rituals of Verification*, Oxford University Press, Oxford (UK).

Rayner, S., 2012, 'Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses', *Economy and Society*, Vol. 41, No. 1: Strategic unknwons: towards a sociology of ignorance, pp. 107–125, doi: 10.1080/03085147.2011.637335.

Rip, A., 2006, 'A co-evolutionary approach to reflexive governance and its ironies', in Voss, J-P., Bauknecht, D., Kemp. R., (eds): *Reflexive Governance for Sustainable Development*, pp. 82-100, Edward Elgar, Cheltenham (UK).

Stilgoe, J., Owen, R., and Macnaghten, P., 2013, 'Developing a framework for responsible innovation', *Research Policy*, Vol. 42, No. 9, pp: 1568-1580, doi: 10.1016/j.respol.2013.05.008.

Strand, R., 2002, 'Complexity, Ideology and Governance', *Emergence*, Vol. 4, No. 1-2, pp. 164-183, doi: 10.1080/15213250.2002.9687743.

Strand, R., Spaapen, J., Bauer, M., et al., 2015, Indicators for promoting and monitoring Responsible Research and Innovation. Report from the Expert Group on Policy Indicators for Responsible Research and Innovation, Publications Office, Luxembourg, (EUR 26866 EN), doi: 10.2777/9742.

Strand, R., and Slaattelid, R., with contributions from Simone, A., Mazzonetto, M., and Fernández, T., 2020, *D7.1 Monitoring and Evaluation Guide*, TRANSFORM, Bergen (Norway), available at: https://www.transform-project.eu/wp-content/uploads/2020/11/TRANSFORM_D7.1.pdf, (accessed 8 June 2022).

Strand, R., and Spaapen, J., 2020, 'Locomotive breath? Post festum reflections on the EC Expert Group on Policy Indicators for Responsible Research and Innovation', in Yaghmaei, E., and van de Poel, I., (eds.): *Assessment of Responsible Innovation: Methods and Practices*, Routledge, London (UK), available at: https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9780429298998-4/, (accessed 8 June 2022).

Völker, T., and Guimarães Pereira, A., 2021, ''What Was That Word? It's Part of Ensuring Its Future Existence' Exploring Engagement Collectives at the European Commission's Joint Research Centre', *Science, Technology, & Human Values*, September 2021, doi: 10.1177/01622439211046049.

Williams, B., 2008, *Shame and Necessity*, 2nd ed., University of California Press, Berkeley (CA).

List of tables

The European Commission's science and knowledge service
Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office
of the European Union