ARTICLE

Which Methods Are Useful to Justify Public Policies? An Analysis of Cost–Benefit Analysis, Multi-Criteria Decision Analysis, and Non-Aggregate Indicator Systems

Gertrude Hirsch Hadorn[1]

Abstract
Science-based methods for assessing the practical rationality of a proposed public policy typically represent assumed future outcomes of policies and values attributed to these outcomes in an idealized, that is, intentionally distorted way and abstracted from aspects that are deemed irrelevant. Different types of methods do so in different ways. As a consequence, they instantiate the properties that result from abstraction and idealization such as conceptual simplicity versus complexity, or comprehensiveness versus selectivity of the values under consideration to different degrees. I hold that none of these methods is best in general. Instead, I opt for the valuation method that is useful for the policy issue in question both in terms of its relevance and in terms of its practicability. Relevance requires that the method can represent and account for what is at stake in the policy issue. Practicability refers to aspects such as easy versus difficult handling of the method. To argue for the claim, I evaluate three types of valuation methods: (1) cost–benefit analysis that rests on unidimensional measurement and ranking, (2) multi-criteria decision analysis that applies multi-dimensional measurement but unidimensional ranking, and (3) non-aggregate indicator systems that operate with multi-dimensional measurement and sometimes also multi-dimensional ranking. Second-order justification indicating whether and how the valuation method chosen is capable of accounting for the substantive value considerations that constitute the real-world policy issue in question renders the conditions on which the results of a proposed policy evaluation rest transparent.

Keywords Cost–benefit analysis · Multi-criteria decision analysis · Non-aggregate indicator systems · Rational choice · Multiple values · Policy justification

✉ Gertrude Hirsch Hadorn
  hirsch@retired.ethz.ch

[1] ETH Zurich, Institute for Environmental Decisions, 8090 Zurich, Switzerland

1 Introduction

By developing and implementing policies, governments intend to meet the needs of people in a better way while, at the same time, maintaining the conditions of life in a changing world. In the context of public or political practice, the term "policy" refers to "[a] principle or course of action adopted or proposed as desirable, advantageous, or expedient".[1] In democratic societies, policy issues such as the question as to whether the status quo of, for example, health policy, security policy, or climate policy needs to be changed and, if so, as to how to achieve this objective are intensely debated. Such debates call for justification of suggested policies, the purpose of which is to provide guidance for decision-making. Generally speaking, a policy should work and promote the common good, and it is in want of democratic legitimacy.[2]

There are various types of methods in use for justifying the epistemic rationality of claims that a proposed policy will work, for assessing the practical rationality of claims that a policy will promote the common good, and for assuring the democratic legitimacy of policy decisions. The application of these methods may result in different assessments and involve problems that are well-known. These problems arise for several reasons. One reason is the lack of knowledge that would be required to apply a method to a specific case, which turns policy choices into decisions under great uncertainty (Hansson and Hirsch Hadorn 2016; 2018). Another reason relates to the restrictions that the abstractions and idealizations of science-based methods impose on the attempt to account for substantive values and, furthermore, to causal assumptions that constitute a real-world policy issue. Science-based methods for policy assessment typically share a consequentialist perspective on policies. They represent assumed future outcomes of policies and values attributed to these outcomes in an idealized, that is, intentionally distorted way and abstracted from aspects that are deemed irrelevant. In what follows, I set out to show that different types of methods do so in different ways. As a consequence, methods instantiate the properties that result from abstraction and idealization such as conceptual simplicity versus complexity, or comprehensiveness versus selectivity of outcomes and the values under consideration to different degrees.[3] These properties determine the usefulness of a method in terms of

---

[1] OED entry "policy": https://www.oed.com/search?searchType=dictionary&q=policy&_searchBtn=Search (accessed 22 March 2021).

[2] In today's democratic societies, the model of the common good is used as a normative reference for political debate. There is no generally accepted definition of the "common good", however. Rather, it is a contested concept. Since the common good serves as a socio-political ideal, the notion has been shaped from different perspectives. Conceiving of the common good as sustainable development, for example, closely relates the common good to social justice. In political philosophy, the concept of the common good is used as a functional concept to distinguish common from private interests. Such a distinction can already be found in Plato's conception of the Republic, see Hussain (2018). Political theories of modern liberal democracy relate the common good to democracy and justice. Hussain (2018, para. 2), for instance, explains the common good "as part of an encompassing model for practical reasoning among the members of a political community. The model takes for granted that citizens stand in a 'political' or 'civic' relationship with one another and that this relationship requires them to create and maintain certain facilities on the grounds that these facilities serve certain common interests. The relevant facilities and interests together constitute the common good and serve as a shared standpoint for political deliberation." There are different stances on whether something like common interests is really required in order to promote the common good, and, if so, on how to distinguish these interests from private ones. These general stances bear on how they conceive of the common good, democracy and social justice more specifically.

[3] There has been extensive debate about the roles of these properties in the sense of virtues or cognitive values of scientific representations in the natural sciences and humanities. For a systematic account, see

its relevance, that is, in terms of whether the method can represent and account for what is at stake in the policy issue. Furthermore, these properties determine the usefulness of a method with respect to its practicability, for example easy versus difficult handling of the method. Moreover, they bear on the results of policy evaluation. Therefore, a justification of the usefulness, that is, the relevance and the practicability of a type of method with regard to the policy problem in question is required.

So as to argue for the claim that the second-order decisions concerning the choice of a method for policy justification have to be justified themselves, I focus on the abstractions and idealizations of three frequently used types of valuation methods for justifying the practical rationality of policies. In Sect. 2, I discuss cost–benefit analysis (CBA) that rests on unidimensional measurement and ranking. In Sect. 3, I turn to multi-criteria decision analysis (MCDA) that applies multi-dimensional measurement but unidimensional ranking. In Sect. 4, I discuss non-aggregate indicator systems that operate with multi-dimensional measurement and sometimes also multi-dimensional ranking. In all of these three sections, I point out the conditions under which it is (not) justified to use the respective type of method. In Sect. 5, I provide a summary and a conclusion. In closing, I comment on relations between practical rationality, epistemic rationality, and democratic legitimacy that are to be taken into consideration when a particular method is to be chosen and applied.

2 Cost–Benefit Analysis

The basic idea of CBA was introduced by the French engineer and economist Dupuit in the nineteenth century (Dupuit 1844). The development of CBA as a practical economic method has begun about 100 years later with the US Flood Control Act of 1939. Because flood control was seen to be in the interest of the general welfare, this Act stated that "[…] the Federal Government should improve or participate in the improvement of navigable waters […] if the benefits to whomsoever they may accrue are in excess of the estimated costs" (English translation of Dupuit 1884 as cited in Pearce 1983, 14). This vague idea of justifying the adoption of a policy or project by using the ratio of future benefits and the costs thereof as a decision criterion has evolved into a family of methods for measuring the expected economic value. The expected economic value is attributed to a policy as a—however questionable—indicator of social welfare (for the history of CBA in economics see Pearce 1983). In many countries, CBA has been applied to policy choices in various fields. For the US, for instance, executive orders by the presidents Reagan,[4] Clinton,[5] and Obama[6] have established what Cass Sunstein has called the "Cost–Benefit Revolution", namely the requirement that the decisions by the US administration on regulatory issues have to be informed by a CBA (Sunstein 2018, 3–26). At the international level, CBA has been used, among other studies, in the Stern Review "The Economics of Climate Change"

Footnote 3 (continued)

Hirsch Hadorn (2018). These value concepts are typically unclear and used differently in various disciplines. For the concepts of "simplicity" and "complexity", see Chall and Martens (2020).

[4] https://www.archives.gov/federal-register/codification/executive-order/12291.html (accessed 22 March 2021).

[5] https://www.archives.gov/files/federal-register/executive-orders/pdf/12866.pdf (accessed 22 March 2021).

[6] https://obamawhitehouse.archives.gov/the-press-office/2011/01/18/executive-order-13563-improving-regulation-and-regulatory-review (accessed 22 March 2021).

to justify the conclusion that "the benefits of strong and early action far outweigh the economic costs of non acting" (Stern 2007, xv). The worldwide popularity of CBA is not least to be put down to the reason that CBA seems to be unique as a practical but reasoned technique, capable of reducing a very complex problem to a more manageable task while delivering a definite quantitative result (Pearce 1983, 21).

As Hansson (2007, 163–164) has rightly noted, "many steps are needed to take us from this basic principle [of weighing advantages against disadvantages, costs against benefits] to any of the forms of cost–benefit analysis that are currently in use. Some of these steps do not share the immediate intuitive appeal of the fundamental principle". Grounded in welfare economics and ethical consequentialism, justification for policy proposals provided by CBA rests on a specific set-up for comparing alternative policy options in terms of their value. As a comparative method, CBA presupposes a given set of alternative policy options that are characterized by the probability of their respective future outcomes which are to be taken into consideration. The Stern Review, for instance, compares the consequences of action and of inaction on climate change in terms of consumption, education, health and the environment (Stern 2007, 32). These are typically used as dimensions of human well-being in development and sustainability assessments. As far as the procedure is concerned, standard CBA then calculates an ordering of the policy options based on their total expected value in four steps. First, general empirical preferences of cardinal format are assigned to the outcomes. Therefore, CBA claims to be based on general facts, not on ideological positions (Sunstein 2018, IX–XII). These preferences are typically measured in monetary terms. For this purpose, preferences are either inferred from behavior (revealed preferences) or explicitly formulated (stated preferences), and they are measured by attributing a cardinal value to each of the future consequences to be considered in conducting the CBA. In the case of benefits, this is willingness to pay for improvement or willingness to accept compensation for worsening. Costs are measured as opportunity costs and changes in price or income. Contingent valuation serves as a substitute for real-market prices to include non-market consequences of public policies. Since preferences may depend on any implicit substantive considerations of those individuals who have been investigated, CBA claims neutrality in terms of substantive value considerations and comprehensiveness regarding the outcomes that can be included in valuation. An example is accounting for moral principles such as respecting human rights or preserving biological species by representing these principles as monetized preferences for morally relevant outcomes of a policy. The Stern Review, for instance, uses, among others, the ecosystem services framework (Costanza et al. 1997) that conceives of the environment and environmental processes as natural capital providing life supporting services.[7] In the second step of the procedure, the expected value of each of the outcomes is determined by multiplying the cardinal preferences for each of the outcomes with the probability of the respective outcomes. Thereafter, in the third step of the evaluation procedure, the expected values of

---

[7] The Millennium Ecosystem Assessment (2005), for instance, has proposed "The Total Economic Value Framework" for "assessing the impacts—the gains and losses—of alternative ecosystem management regimes" (p. 132). The framework encompasses various economic valuation methods including contingent valuation in order to determine not only the use and the option value of the ecosystem services for human welfare but also the existence value of ecosystems. The authors of the Millennium Ecosystem Assessment argue that a comparative evaluation of the existence value of ecosystems is needed because of trade-offs between impacts. They concede, however, that the market value does not capture the existence value. They suggest establishing, in addition, a politically determined safe minimum standard that would serve as a threshold for contingent valuation results.

the outcomes are aggregated for each policy option. The fourth and last step finally consists in an ordering of the policy options according to their total expected value. The policies that have been assigned the maximal total expected value in the CBA would be rational to choose. Therefore, CBA claims to be a quantitative method that informs decision-makers about the practical rationality of alternative policy options.

As a unidimensional cardinal valuation method, CBA abstracts from multiple substantive value considerations as set forth in the above. Furthermore, idealization is required for generalizing the data concerning the empirical preferences of individuals that have been measured in a representative sample. Assuming that these data on empirical preferences are externally valid necessarily entails ignoring possible influences on preferences in the sample that had been included in data collection, for example context sensitivity. The rationale for abstraction and idealization in CBA is to structure a decision on public policies in a way that is analogous to the model of rational choice in a decision under risk in decision theory. I discuss three issues that make this analogy questionable. (1) As for abstraction, well-known problems show that CBA as specified for empirical application is neither neutral regarding substantive value considerations nor comprehensive in terms of outcomes that can be included in valuation. (2) With respect to idealization, it is well-known too that the empirical preferences constructed by means of some kind of CBA do not meet the requirements of decision theory. Sunstein has referred to these two issues as the "knowledge problem" of CBA (Sunstein 2018, 79). These issues may have serious impacts in the sense that CBA results are neither relevant nor accurate. (3) In addition to these two issues of the method that relate to descriptive aspects, there is also a normative problem, since one may ask, whether the principle of rational choice that provides the rationale for abstractions and idealizations in CBA is an appropriate normative principle for decisions on public policies.

(1) A considerable part of the critique of CBA has focused on the use of monetization for measuring the expected economic value. Although problems that are specific to monetization can be avoided by drawing on a different standard for measurement, for example disability-adjusted life years (DALYs), fundamental problems remain. These are a consequence of abstracting from multiple substantive value considerations and replacing them by unidimensional comparative value judgments and aggregation (Hansson 2007, 164). I focus on the fundamental problems relating to CBA that uses monetization as a standard while justifying specific standards is beyond the scope of this paper. The formal concept of value in decision theory as introduced by Luce and Raiffa (1957) is "utility". "Utility" is an empty concept without substantive meaning. It only serves to construct a cardinal scale of fictional utiles that act as the units for representing preference orderings.[8] Constructing the empirical preferences requires some empirical technique such as monetization. As a consequence, empirical preferences no longer simply describe a formal relational value judgment. Rather, they have a substantive interpretation, which, in the case of monetization, is proportional demand.[9] Modelling comparative value judgments in relation to

---

[8] For the various conceptions of "utility" in decision theory and as an ethical position in utilitarianism see Briggs (2017), Hansson (2016) and Luce and Raiffa (1957).

[9] It is interesting to consider Aristotle's argument for monetization concerning this problem. According to Aristotle, monetizing is a sufficiently accurate common measure that is required in equalizing demands for trading commodities because "[…] all commodities exchanged must be able to be compared in some way. It [i.e., this measure] is to meet this requirement that men have introduced money […] for it is a measure of all things, and so of their superior or inferior value, that is to say, how many shoes are equivalent to a house or to a given quantity of food. […] It is therefore necessary that all commodities shall be measured by some one standard […]. And this standard is in reality demand […], which is what holds everything together, since if men cease to have wants or if their wants alter, exchange will go on no longer, or will be on differ-

demand as the standard turns the alleged neutrality of CBA regarding the diversity of substantive value considerations into a replacement of the diverse substantive considerations by demand measured as a degree of exchange value such as a particular amount of money. This replacement is problematic, however, for instance when a price needs to be attributed to a human life. The costs of saving lives differ between countries because of national circumstances including opportunity costs. Therefore, Working Group III of the Intergovernmental Panel on Climate Change concluded in its second report that human lives vary in their monetary value (IPCC 1995). This inequality in the values of human lives has been strongly contested in public for moral reasons (Brun and Hirsch Hadorn 2008). What is no less morally questionable is the principle of compensating some people's losses by other people's gains in aggregation (Hansson 2016, 40–45). This issue also arises when future costs and benefits of climate policy are discounted (Broome 2008). This means that monetization does not cover moral considerations while moral considerations make it impossible to treat moral values fully as economic values. The reason for this is that assigning prices to human lives does not entitle anyone to "buy" a person and to do with this "purchase" whatever he or she wants. Against this background, the economic evaluation of the ecosystem services in policy evaluation, too, needs to be questioned. Ecosystem services "are not typical commodities, but instead are characterised by limited degrees of substitutability, non-linearities and critical thresholds that imply that they might not always change in marginal ways. When an ecosystem is approaching a critical threshold […] some of the services it provides may become scarce in absolute rather than relative terms, i.e., if these services are fundamental to the satisfaction of critical human needs and rights, then their further loss can no longer be compensated through trade-offs with other goods and services" (Wegner and Pascual 2011, 495). Because CBA presupposes that such trade-offs are legitimate, it is not appropriate for such cases. In consequence, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) has proposed valuing biodiversity and ecosystem services not only in terms of their instrumental or economic value but also in terms of their intrinsic value and their relational value. Both, intrinsic value and relational value, the latter of which refers to contemplative relationships of humans with nature, are understood to be not substitutable. Hence, the application of CBA for measuring these values is not appropriate (Chan et al. 2016).

In order to clarify the function of demand as a choice value and to discuss the principal critical points raised against CBA as a unidimensional valuation method, I use terminology proposed by Chang (1997). She calls any substantive value consideration that serves as a standard (criterion for evaluation) a "covering value" since comparing the outcomes of options with respect to this value is only meaningful if they are covered by this value. A covering value that acts as a choice value should comprise everything that is relevant to a given choice. Choice values are either simple or complex, encompassing a range of so-called "contributory values" in the latter case. A complex value can be conceived of either as a list of contributory values only or as a new value itself.[10] I regard a complex value as a

---

Footnote 9 (continued)

ent lines. But demand has come to be conventionally represented by money […]. Money then serves as a measure which makes things commensurable and so reduces them to equality. […] Though therefore it is impossible for things so different to become commensurable in the strict sense, our demand furnishes a sufficiently accurate common measure for practical purposes" (Aristotle 2003, 1133a17–b21).

[10] Chang (1997) moreover suggests terminology to distinguish between different problems of evaluative comparisons: (1) non-comparability: there is no common covering value to compare two options; (2) incomparability: no traditional value relation holds between two options, that is, they are not equal and neither of them is better than the other; (3) incommensurability: two options cannot be compared on

new value if a new scale is constructed to represent the scores on the contributory values in some integrated way.

In Chang's terminology, CBA can be characterized as follows: CBA assumes that demand can function as a covering value and as a choice value in policy justification since it claims to cover all relevant outcomes of the suggested policies because of its neutrality regarding substantive value considerations. I submit that demand functions as a simple but not as a complex choice value. The reason for this is that demand is taken to cover everything that matters in a choice without including a range of contributory substantive values—neither as a mere list nor by constituting a new value. Instead, monetization abstracts from the diversity of substantive value considerations, replacing them by constructing a degree of demand for every outcome. If this reconstruction of the function of demand in CBA is correct, the claim that CBA is neutral and comprehensive is not tenable. Rather, CBA is a biased approach because its focus on demand is not made explicit as a restriction implied by the method. Another measurement technique using a different simple choice value would entail a different restriction on valuation since unidimensional valuation is a fundamental type of restriction that CBA as an empirical valuation method involves. Thus, when CBA is employed to justify real-world policies, the simple choice value used to specify a CBA for application needs to be justified in terms of how critical related restrictions are to what is at stake in the policy issue in question.

There is yet a further reason why CBA builds on substantive value considerations that require justification. Constructing the set-up of a CBA and characterizing the policy options is always done from a specific perspective on the problem that is up for decision (Hansson 2007). This perspective with its related contextual factors bears on the topic that is selected, on how this topic is demarcated, on the structure of the decision (one-shot or dynamic), on the alternative policy options, and on how these options are framed and characterized, that is, which of their possible future consequences are considered. This influences people's preferences and, eventually, the resulting expected values. Thus, the construction of a CBA and policy evaluation in general are guided by specific substantive value considerations. A case in point is the evaluation of nutritive technologies in the report of the United Nations Framework Convention on Climate Change (UNFCCC 2008). Such technologies have been proposed to reduce greenhouse gas emissions from livestock by way of modifications to feed, including composition, plants, quality, and supplements such as tannins or saponins. Each of these options can contribute to reducing emissions of methane and other greenhouse gases, but with side effects. In the UNFCCC report, the values that followed from focusing on nutritive technologies while ignoring other sorts of measures were made explicit. This appraisal presumed, however, that production levels should not be affected because of the common regulations in EU countries and in view of natural and technological agricultural conditions in Europe (Hirsch Hadorn et al. 2015). Reflection on the ultimate values may thus have resulted in other methods that could have been considered to be more favorable. The Swiss Federal Government, for instance, decided in 2013 to change its scheme of direct subsidies to farmers. The system has shifted from granting subsidies according to the number animals to limiting the number of animals. Moreover, if farmers want to receive direct subsidies, they are now required to fulfill several conditions

---

Footnote 10 (continued)

a cardinal scale. These problems of evaluative comparisons are analyzed in more detail in Brun and Hirsch Hadorn (2008).

that relate to the management of cattle and the quality of their agricultural land.[11] This example illustrates that an explicit definition of the values that have guided the set-up of policy evaluations provides decision-makers with important information. Against this background, Resnik (2011, 8–13) has conceded that second-order decisions determine how a decision is specified, that is, how the options for choice are determined and framed as well as what the relevant outcomes are. He refrains from justifying them explicitly, however, since he takes rational choice versus gut feelings as the only alternatives available and excludes second-order rational choice to avoid regress. Being explicit about and justifying these decisions by means of arguments is what is asked for in real-world policy decisions, however (Hansson and Hirsch Hadorn 2016).

(2) CBA does not meet the requirements of modeling real-world policy decisions as rational choice because the model of rational choice builds on a technical notion of "preference" that is different from colloquial speech. Preferences in this technical sense are states of mind that express relational value judgments between at least two propositions or bundles of goods A and B, namely one being better than the other (A > B), or both being equal in value (A = B) (Grüne-Yanoff and Hansson 2009, 161–162). In order to serve the purpose of establishing a value-ordering of the policy options to choose from according to the total expected utility of each option, preferences in the technical sense have to fulfill several formal requirements. The axioms of utility theory such as transitivity and completeness require that preferences be stable and generalizable over time, policy issues, and people. Thus, they are idealized and presumed to be insensitive to variable context factors. There is vast empirical evidence suggesting that elicited and revealed empirical preferences do not fulfill these formal requirements (Lichtenstein and Slovic 2006). Empirical preferences are affected by various problems that have been termed "bounded rationality" (Grüne-Yanoff 2007). For instance, a different framing of logically equivalent policy options may have the effect of a change in preference. Moreover, context sensitivity of empirical preferences questions their external validity when they are taken as generalized preferences in another CBA that is conducted in a different context as, for instance, for another sort of policy choice. Ackerman and Heinzerling (2004) discuss the difficulties that arise when the statistical value of a human life is calculated in detail. In particular, they raise doubts about the validity of the results of such calculations in the context of typical uses of CBA in which the benefits and risks of policies are assessed with respect to health and the environment (Ackerman and Heinzerling 2004, 61–90). As regards the valuation of ecosystem services in policy evaluation, Wegner and Pascual (2011) point to general problems that question the validity of price information used in CBA. They hold that this sort of information "is easily distorted by externalities, imperfect forms of competition, subsidy schemes, distributional biases and imperfect flow of information" (p. 493). More specific problems regarding the attempt to evaluate ecosystem services are entailed by the fact that individuals "tend to value elements of nature according to different rules, depending on the institutional context within which they act [… Moreover,] the unequal distribution of wealth in society may significantly influence the determination of market prices, biasing the analysis towards the preferences of higher-income groups and resulting in the under-provision of important ecosystem services to lower-income groups" (p. 502). Because preferences used in a CBA do not fulfill the requirements of decision theory, their use for determining rational choice must be questioned. Thus, opting for the best real-world policy according

---

[11] https://www.fedlex.admin.ch/eli/cc/2013/765/de (accessed 14 March 2021).

to a betterness-ordering determined by CBA is not justified as a rational choice but rather in the sense of choosing the best option given that the restrictions that are imposed by the specific unidimensional valuation and the empirical preferences on which the policy choice under consideration builds are appropriate. The question as to whether there are cases in which CBA can be deemed suitable is the subject of controversy among experts. According to Sunstein, who is a proponent of CBA, this type of analysis is appropriate for easy cases, that is, cases in which those who benefit from the policy also pay its costs, but not for "hard" cases in which the constellation is more complex (Sunstein 2018, 40–42). Sen, by contrast, doubts both the relevance and the validity of CBA results in principle: "When all the requirements of ubiquitous market-centered evaluation have been incorporated into the procedures of cost–benefit analysis, it is not so much a discipline as a daydream. If, however, the results are tested only in terms of internal consistency, rather than by their plausibility beyond the limits of the narrowly chosen system, the glaring defects remain hidden and escape exposure. Daydreams can be very consistent indeed" (Sen 2000b, 952).

(3) I now turn from the description-related issues to the normative claim of CBA. It is questionable whether the practical rationality of rational choice would provide appropriate justification for decisions on real-world policies that promote the common good. Because "rational choice" as a formal conception of rationality is based on the axioms of decision theory, one has to ask why it is rational to follow these axioms. Velleman (2014) has argued in this connection that "insofar as obeying the axioms of decision theory is rational, it's rational because it makes our preferences formally coherent, thus ensuring that they are intelligible" (p. 37). Nevertheless, if intelligible preference orderings come at the price of unidimensional idealization, they provide insufficient justification for real-world policies that aim to promote the common good. Thus, the problems with CBA do not arise from building it on preferences as such but rather from constructing a betterness-ordering of preferences based on a unidimensional measurement of preferences. Since the UN Conference for Environment and Development (UNCED) in 1992, most countries and many global organizations have signed commitments to promote sustainable development. If the common good is conceived as sustainable development, the practical rationality of a policy rests on substantive assumptions such as increasing distributive justice between and within generations while respecting multiple substantive criteria with critical limits, as discussed in Sect. 4.

In sum, CBA performs well regarding practicability because of an easy handling of the method compared to other methods. This advantage follows from the application of unidimensional measurement techniques and the use of existing data sets. The relevance of the method, however, is restricted to those cases in which use of unidimensional measurement techniques and of existing data sets is appropriate. These restrictions limit the usefulness of CBA considerably. Nevertheless, there has been a tendency to ignore this limitation, not only because of the practicability of the method. CBA delivers a definite quantitative result that can be communicated in simple terms. This is why CBA meets expectations of those decision-makers who prefer this straightforward type of results and suppose policy assessment to deliver them.[12] This *prima-facie* advantage of CBA may well turn into a disadvantage, however, if the policy that performs best in a CBA proves to be inefficient later on because the conditions concerning relevance had not been fulfilled by the analysis.

---

[12] I am indebted to an anonymous reviewer for drawing my attention to this important point.

3 Multi-Criteria Decision Analysis

MCDA includes a broad range of very heterogenous methods that are applied, for instance, in environmental impact assessments (EIA) to assess the unintended impacts of a proposed policy or project on flora, fauna, ecosystem, soil etc. (Rauschmayer 2001), in life-cycle assessments (LCA) of products and processes (Finnveden et al. 2009), or in many indicator systems for sustainable development (Singh et al. 2012). I discuss MCDA as an example of a hybrid method that uses aggregation to combine multi-dimensional measurement with a unidimensional betterness-ordering. Sometimes, also non-aggregate techniques are referred to as "MCDA". I, by contrast, use the term for aggregate methods only and go into non-aggregate methods in Sect. 4.

MCDA starts by determining the multiple criteria that are relevant to the subject of the policy choice to be dealt with. These criteria are operationalized by converting them into empirically measurable single indices and assigning the respective scores for each of the outcomes of a proposed policy on these indices. Both the determination of these indicators and the assurance of data quality are the subject of intense methodological debate (Singh et al. 2012, 287). In the next step, MCDA synthesizes this diversity into a composite index by means of an additive procedure. The aim of this step is to establish a one-dimensional ranking of the options that are open to choice. For this purpose, a broad range of synthesis methods are in use. Roy (2005) has proposed distinguishing between methods that rest on an overall synthesizing criterion (e.g., MAUT, SMART) and tools that work with outranking methods of pairwise comparison (e.g., ELECTRE, PROMETHEE). In Chang's (1997) terminology, the multiple single indices or criteria used for assessment purposes are multiple contributory values of a complex covering value. This complex covering value is a new value because a composite index is constructed to establish a unidimensional betterness-ordering of the alternative options, projects, products, or states. The mathematical procedures that are employed to construct the composite index abstract from the specific meaning of each of the contributory values. Weighting factors ensure that the relative importance of each contributory value is given due consideration.

MCDA is a hybrid with respect to practical rationality and not a purely consequentialist approach. MCDA claims to account for the relevant aspects of the subject for policy choice. On the one hand, justification for abstraction in the selection of the diverse criteria and for idealization in the process of constructing the respective indices has to be grounded in theoretical considerations called "frameworks" (top-down approaches). On the other hand, procedures are required that ensure context sensitivity and democratic legitimacy (bottom-up approaches), such as participation of stakeholders in determining the more specific indices and subindices (Rauschmayer and Risse 2005; Stirling 2006). MCDA determines the best of alternative policies, using unidimensional ordering as a rationale for abstraction and idealization in the construction of the composite index by means of a particular synthesis method. This entails issues concerning the technical level, for example regarding the question of how to construct the scale of the composite index and of how to represent the scores on a diversity of scales as one score on one scale. The type of scale that is used to measure the variable determines the aggregation procedure: interval-scale non-comparability applies dictatorial ordering, interval-scale full comparability builds on the arithmetic mean, and ratio-scale non-comparability rests on the geometric mean whereas ratio-scale makes use of any homothetic function (Singh et al. 2012, 286–287). Further issues concern a conceptual level and include the question of how to weigh the various criteria when the purpose consists in abstracting from their meaning to construct a composite

index. Tackling this question, Marttunen et al. (2019) discuss the benefits and challenges of four methods: means-ends networks, relevancy analysis, correlation analysis/principal component analysis, and local sensitivity analysis of weights. They use data from two case studies: one dealt with sustainable water infrastructure planning in Switzerland while the other focused on regional land use planning in Finland. The importance scale that is common in relevance analysis distinguishes for each of the included variables, which are called "objectives", between five scores ("low", "moderate", "high", "very high", and "unable to determine the importance"). In the reviewed case studies, each score had been associated with a set of qualitative criteria in order to classify the objective. Classification with the score "high", for example, required that at least one of the following four criteria had been met: (i) Economic, social, cultural, or natural values that are related to the objective under consideration are high. (ii) The objective is sensitive to changes in the external environment, and recovery takes a long time. (iii) There are binding regulations (e.g., legislation) concerning the use or the state of the objective. (iv) Local people or other stakeholders are worried about the changes in the state of the objective (Marttunen et al. 2019, 618).

MCDA is a conceptually more complex method compared to CBA because of the multi-dimensional measurement and synthesis methods, and it provides a less easy handling in application. Furthermore, MCDA does not claim neutrality regarding substantive considerations but seeks to ensure relevance of the selected criteria as regards the given policy goals and context. These properties of MCDA may well be an advantage in justifying real-world policy choices. Information about what criteria have been considered and how scores on these criteria contribute to the overall result not only ensures transparency of the results. This information may also serve to justify the set-up of a policy choice, because it is primarily the first step of an MCDA that informs about how a proposed policy scores on the various selected indices. Since MCDA constructs the complex index as an additive account, it assumes that gains and losses in terms of different criteria may compensate for each other in principle. In the overall valuation result of MCDA, as in CBA, these compensations are hidden. It is possible, however, to reconstruct information about what criteria had been included in the evaluation of a policy, about the scores on these criteria and about the contribution that these scores make to the overall result in the aggregation step. Trade-offs may be questionable if thresholds are not met for several important criteria while, due to high positive scores on other criteria, the overall performance resulting from the second step is good. A case in point are sustainability rankings of companies for the capital market. These rankings typically rest on ESG indices, that is, indices that comprise environmental, social and governance-related criteria. The SIX group, for example, uses an aggregated ESG index to provide a sustainability rating of companies for the Swiss capital market. This index consists of twelve levels and ranges from D to A. In order to be included in the ranking a company first of all has to meet an ESG impact rating of at least C+. In addition, the company "must generate less than 5% of its sales in a critical sector. These critical sectors are adult entertainment, alcohol, armaments, betting, genetic engineering, nuclear energy, coal, oil sands and tobacco. Furthermore, an index candidate must not be on the exclusion list of the Swiss Association for Responsible Investments."[13] Such practices merely serve to avoid the possibility of questionable concealed compensation by criterion-based exclusion of certain companies from sustainability ranking. It is not

---

[13] https://www.six-group.com/en/products-services/the-swiss-stock-exchange/market-data/indices/esg-indices.html (accessed 12 April 2021).

a means that allows a transparent handling of compensations between scores on different ESG indices for those companies that are included in the sustainability ranking.

All things considered, MCDA performs well regarding relevance if the multiple criteria that are included are suitably determined. As for practicability, however, the aggregation step usually proves to be rather demanding. Moreover, transparency of information with regard to trade-offs between criteria and thresholds may serve as a reason for conducting a MCDA without aggregation. Hawkins et al. (2012), for example, chose to employ a non-aggregate environmental LCA to compare conventional and electric vehicles (EVs). They summarized some of their findings as follows: "EVs powered by the present European electricity mix offer a 10% to 24% decrease in global warming potential (GWP) relative to conventional diesel or gasoline vehicles assuming lifetimes of 150,000 km. However, EVs exhibit the potential for significant increases in human toxicity, freshwater eco-toxicity, freshwater eutrophication and metal depletion impacts, largely emanating from the vehicle supply chain" (Hawkins et al. 2012, 53). On the one hand, explicit and comprehensive information on specific advantages and disadvantages of EVs and, more generally speaking, the policy under consideration provides a solid basis for informed debate on alternative policies and, eventually, for decision making. On the other hand, administrators, the public, and decision-makers may easily feel overwhelmed when they are confronted with a bulk of complex information. This is a consequence of opting against unidimensional aggregation of multi-variable results, which, in turn, makes it impossible to establish a betterness-ordering of the alternative options and thus to provide simplified but easily communicable information.[14]

4 Non-Aggregate Indicator Systems

Besides approaches that work with unidimensional measurements, there are multi-dimensional valuation methods without aggregation, for instance some non-aggregate indicator systems of sustainable development such as the one proposed by the United Nations (CSD 2007) as a blueprint for developing national indicator systems. The Swiss system called "Monitoring Sustainable Development" (MONET) (FSO et al. 2016), which I use as an example, builds on the UN indicator system too. The principles of sustainable development form the basis of the framework for defining the currently 73 indicators of sustainable development and for determining the relevance and the direction of development for each of these indicators. In Chang's (1997) terminology, such indicator systems provide a list of contributory values that constitute a complex covering value, namely sustainable development. As a complex covering value, sustainable development is not a new value in itself since the contributory values are not aggregated into a new value for unidimensional ranking. The reason for this is that sustainable development as a social mission includes a variety of diverse values that are not hierarchically ordered and cannot be freely compensated either.

As Articles 2 and 73 of the Swiss Federal Constitution set down that the Confederation and the cantons are obliged to promote sustainable development, MONET served as the starting point for devising Switzerland's Sustainable Development Strategy 2016–2019

---

[14] I am grateful to an anonymous reviewer for drawing my attention to problems concerning the way in which politics and the business world take notice of the results of science and concerning the form in which these stakeholders expect these results to be communicated.

and the preparation of the Agenda-2030 sustainable-development goals (SDGs) adopted by the UN in 2015. The periodic procedure of monitoring and informing about Switzerland's performance in terms of sustainable development is intended to initiate debate about where action is required. The Federal Council furthermore decreed that for new and important projects in legislation, planning, or building that are relevant to sustainability it is mandatory to run a sustainability assessment in order "to enable political projects and undertakings to be assessed at an early stage, and if necessary optimized, from the point of view of sustainable development" (ARE 2012, 54). Thus, MONET is employed for ex-ante and ex-post analyses. In the latter case, for the period analyzed, each of the measured indicators informs about the targeted trend for the indicator, the observed trend for the indicator, and the resulting assessment. Results are communicated in the form of qualitative judgments about whether the comparison of the targeted and the observed trend is positive (i.e., more sustainable), negative, or unchanged. The 73 indicators relate to 12 subjects such as social cohesion, international cooperation, economy, or natural resources. In order to simplify the results, 17 of the 73 indicators act as key indicators (FSO et al. 2016).

Sustainable development is a way to conceive of the common good that has received strong democratic legitimacy as a complex societal ideal and policy goal from the global to the local level. Nevertheless, the concept as such is a fundamentally contested concept (Jacobs 1999), and political controversy about how to interpret this concept more specifically is certain to continue (for a more detailed exposition of the controversies see Brun and Hirsch Hadorn 2008). As a consequence, there are also considerable differences between the indicator systems, which result in different assessments. The UN indicator system, for instance, on which the MONET system builds, covers 14 themes (e.g., poverty, atmosphere), 44 subthemes, 50 core indicators, and 96 indicators in total (CSD 2007). As in the case of MCDA, multiple indicator systems may be an advantage. A general set of indicators may be inappropriate, for example, because indicators have to be valid in the specific context to which they are applied.

If a non-aggregate method is used for comparative ex-ante valuation of alternative policies that are up for choice, one may—in principle—calculate whether there exists a maximal option, that is, an option that is not dominated by one of its alternatives (Sen 2000a, 486). A dominated option performs worse on every evaluation criterion than the alternative options (Allenspach 2013, vi). It would therefore be rational to choose a maximal option. Applying a system with 73 indicators clearly faces an efficiency dilemma however, that consists in "a conflict between goal coverage and decisiveness […] since the more aspects we include, the larger can we expect the efficient alternatives to be" (Hansson 2016, 39). Thus, dominance rankings seem to make sense only in connection with a handful of indicators, the selection of which needs to be justified so as not to exclude relevant ones. This justification may well refer to previous uses of the indicator system for monitoring the state of affairs. Such monitoring provides necessary information for an explicit justification of the set-up of policy choices regarding the topic in need of action, the framing of the options, and the relevance of future outcomes to be included in the evaluation. A further important point to be explicitly considered in the set-up is the temporal structuring of the policy decision, that is, the question as to whether a decision should be taken in one attempt or stepwise by partitioning the decision and extending the process over time. The latter option may prove useful in real-world policy choices in two regards. First, it is possible to avoid the trading of improvements in some indicators for worsening in others by respecting thresholds and improving scores on the various indicators in sequential steps (Allenspach 2013). Second, taking decisions stepwise over time for repeated adaptation of a policy allows dealing with the uncertainty of future outcomes (Hirsch Hadorn 2016).

Since we cannot be sure of what will happen in the future, adapting the predictions and values that are adduced for policy design and justification may be indispensable.

In sum, non-aggregate methods perform well regarding relevance if the included multiple indicators are well determined. At the same time, they lack practicability because they are labor-intensive with respect to the generation of data. Furthermore, the application of non-aggregate methods only rarely results in a simple conclusion and straightforward policy advice, which is a consequence of the diversity and the complexity of information. To meet the expectations of those administrators, the public, and decision-makers that prefer a more straightforward type of results, a viable solution may consist in focusing on communicating the policy options' performances on selected indicators.

5 Summary and Conclusion

In this paper, I have argued for the claim that we should justify a policy by opting for the method that is useful for the policy issue in question both in terms of its relevance and in terms of its practicability. Relevance requires that the method can represent and account for what is at stake in the policy issue. Practicability refers to aspects such as easy versus difficult handling of the method. In order to substantiate my claim, I have evaluated three types of valuation methods with respect to how they represent values attributed to assumed future outcomes of a policy in an idealized, that is, intentionally distorted way and abstracted from aspects that are deemed irrelevant. (1) Cost–benefit analysis rests on unidimensional measurement and ranking, assuming that a single and simple covering value is capable of encompassing the values of all relevant possible outcomes. (2) Multi-criteria decision analysis applies multi-dimensional measurement but unidimensional ranking, which requires the construction of a new composite covering value out of the diverse values of the relevant possible outcomes as a basis for a betterness-ordering of the evaluated alternatives. (3) Non-aggregate indicator systems operate with multi-dimensional measurement and sometimes also multi-dimensional ranking, based on a list of contributory values that constitute a complex covering value. These three approaches instantiate the properties that result from abstraction and idealization such as conceptual simplicity versus complexity, or selectivity versus comprehensiveness of the values under consideration to different degrees. While relevance often requires a certain extent of complexity and comprehensiveness, practicability of a method is achieved through simplicity and selectivity. A method may be useless if it is not relevant to the problem at hand or if it is not practicable in a given context. Therefore, to be useful, considerations on the relevance of a method have to allow for practicability and vice versa: considerations relating to relevance have to respect a minimal degree of simplicity and selectivity required for practicability, while considerations relating to practicability have to respect a minimal degree of diversity and complexity that is required for relevance.

Providing justification for a policy by means of CBA is appropriate if the purpose is to determine which policies of a range of likely alternatives can be expected to perform best on a single criterion while best performance according to a set of relevant criteria requires some sort of MCDA. Non-aggregate indicator systems, by contrast, are appropriate if the aim is to find out whether and in what direction issues in the real world need to be changed by a new policy. In sum, none of these methods has proved to be best suited in general because each of them is appropriate for different purposes and contexts. It seems, however, that there are fewer cases in which CBA is useful than

one might expect. Usually, MCDA or non-aggregate indicator systems can be assumed to be better suited. One reason for this conclusion is that often, policies should explicitly account for multiple values that are neither hierarchically ordered nor capable of being freely compensated, for instance in the case of policies to promote sustainable development. Furthermore, justification may be required to determine how each of the values attributed to assumed future outcomes should be weighted (Hirsch Hadorn et al. 2011). Another reason is that policy evaluation should include a justification for the set-up of policy choices with respect to the topic in need of action, to the framing of the options, and to the relevant future outcomes. Thus, the justification for the choice of a valuation method is conditional not only on the democratic legitimacy of the included values but also on the causal assumptions on which the set-up of a policy choice and the design of the options rest. Scientific methods that are employed to justify the epistemic rationality of proposed policies, for example randomized controlled trials, require abstractions and idealizations too. Therefore, if findings from such trials are meant to serve as evidence for claims of effectivity, these claims may be epistemically misleading. Against this background, Cartwright and Stegenga (2011) argue that evaluation of the epistemic rationality of a policy necessitates a combination of several methods in order to substantiate a general empirical effectivity claim for a given context of application. Thus, according to their proposal, evidence from randomized controlled trials is not to be replaced altogether but rather needs to be complemented by adducing causal background knowledge as well as knowledge about the diverse causal interactions at work in the given context of application and about contextual interactions that have an impact on whether the implementation of the policy will be successful in the intended context of application. Cartwright and Stegenga (2011) assume this procedure to ensure the relevance of evidence claims concerning the effectivity of the proposed policy in question. Objecting to such kinds of proposals, Luján and Todt (2021) have argued for not regarding evidence from randomized controlled trials as a requirement or as a privileged type of evidence. They propose to regard evidence from randomized controlled trials as one out of several possible starting points for ensuring epistemic rationality which can be achieved by means of a variety of alternative methods. Since these methods usually result in different outcomes, the method that seems adequate in the context of a particular regulation needs to be chosen (Luján and Todt 2021, 3).

My analysis has focused on types of abstraction and idealization on which policy valuation methods rest. What I have not provided, however, is a discussion concerning the justification of the specific abstractions and idealizations that are required in a concrete application of a method, which, of course, is also very important with respect to the assessment of the usefulness of a method for policy justification. Policy evaluation methods that lack justification of the idealizations and abstractions on which they rest may cause concern, objections, and even irritation in political debates on policy decisions, which has been highlighted by Greenberg, the former chief-editor of the Journal "Risk Analysis":

> Having worked on risk analyses and environmental assessments for decades, I am accustomed to breathing a sigh of relief when we complete reports. But the relief is short-lived because, inevitably, my colleagues and I must respond to questions. Sometimes, the questions come from other technical people; other times, impatient elected officials, reporters, and the public. Inevitably, no matter how much time we spent on the study, some of the questions feel like personal attacks, and some are. For example, why did you answer these questions and not those questions? Why did you assume this instead of that? (Greenberg 2017, 843)

All in all, the provision of a justification for the various second-order decisions and their substantive assumptions in the context of choosing and specifying a method makes policy justification transparent to decision-makers. This allows them to judge the proposed results with respect to whether the assumptions on which these results build are appropriate for the policy issue at hand. At the same time, raising the awareness of the conditionality of the results from policy evaluation methods may prevent the public from holding false expectations as regards scientific policy justification. Science communication in the context of policy justification is not to be regarded as a marginal task, however. Rather, this is a process that is to be understood and planned as a demanding and quite challenging effort of its own (Pielke 2014). Such communication activities may include addressee-adapted rephrasing and contextualization of the justification of the methods and results. Furthermore, acknowledging and interpreting the findings of science requires at least an elementary level of scientific knowledge on the side of the administrators, the interest groups in the civil society, and the decision-makers because without this knowledge they are not able to grasp the information and related consideration. In this regard, participation of representatives from these groups of stakeholders in the development of the set-up of a CBA, a MCDA or non-aggregate indicator systems may prove to be a promising and worthwhile approach.

Declarations

Conflict of interest The author declares that there is no conflict of interest.

References

Ackerman, Frank, and Lisa Heinzerling. 2004. *Priceless. On knowing the price of everything and the value of nothing*. New York: The New Press.
Allenspach, Urs. 2013. *Sequences of choices with multiple criteria and thresholds. Implications for rational decisions in the context of sustainability*. Doctoral thesis ETH Zurich No. 20974. https://doi.org/10.3929/ethz-a-009773097
ARE [Federal Office for Spatial Development]. 2012. *Sustainable development in Switzerland. A guide.* Berne: Swiss Confederation. https://www.are.admin.ch/are/en/home/media/publications/sustainable-development/nachhaltige-entwicklung-in-der-schweiz-ein-wegweiser.html. Accessed 7 March 2020.
Aristotle. 2003. *The Nicomachean Ethics* (translation by H. Rackham). Cambridge, MA: Harvard University Press.
Briggs, Rachael A. 2017. Normative theories of rational choice: expected utility. In *The Stanford encyclopedia of philosophy* (spring 2017 edition), ed. Edward N. Zalta. https://plato.stanford.edu/archives/spr2017/entries/rationality-normative-utility/. Accessed 28 September 2020.

Broome, John. 2008. The ethics of climate change. Pay now or pay more later? *Scientific American* 298 (6): 69–73.

Brun, Georg, and Gertrude Hirsch Hadorn. 2008. Ranking policy options for sustainable development. *Poiesis & Praxis* 5 (1): 15–30. https://doi.org/10.1007/s10202-007-0034-y.

Cartwright, Nancy, and Jacob Stegenga. 2011. A theory of evidence for evidence-based policy. *Proceedings of the British Academy* 171: 291–322. https://doi.org/10.5871/bacad/9780197264843.003.0011.

Chall, Cristin, and Niels C. M. Martens. 2020. Simplicity in the sciences and humanities: Report on the Bonn "Simplicities and Complexities" Conference. *Journal for General Philosophy of Science* 51 (3): 491–497. https://doi.org/10.1007/s10838-020-09499-2.

Chan, Kai M. A., Patricia Balvanera, Karina Benessaiah et al. 2016. Why protect nature? Rethinking values and the environment. *PNAS* 113 (6): 1462–1465. https://doi.org/10.1073/pnas.1525002113.

Chang, Ruth. 1997. Introduction. In *Incommensurability, incomparability, and practical reason*, ed. Ruth Chang, 1–34. Cambridge, MA: Harvard University Press.

Costanza, Robert, Ralph d'Arge, Rudolf de Groot et al. 1997. The value of the world's ecosystem services and natural capital. *Nature* 387: 253–260. https://doi.org/10.1038/387253a0.

CSD [United Nations department of economic and social affairs. Commission on sustainable development]. 2007. *Indicators of sustainable development. Guidelines and methodologies* (3rd edition). New York: United Nations. https://sustainabledevelopment.un.org/content/documents/guidelines.pdf.. Accessed 22 March 2021.

Dupuit, Jules. 1844. De la mesure de l'utilité des travaux publics. *Annales des ponts et chaussées, Second series, 8*. Translated by R. H. Barback (1952) as "On the measurement of the utility of public works". *International Economic Papers, 2*, 83–110. Reprinted 1969 in *Readings in welfare economics*, ed. Kenneth J. Arrow and Tibor Scitovsky, pp. 255–283. London: Allen and Unwin.

Finnveden, Göran, Michael Z. Hauschild, Tomas Ekvall et al. 2009. Recent developments in Life Cycle Assessment. *Journal of Environmental Management* 91 (1): 1–21. https://doi.org/10.1016/j.jenvman.2009.06.018.

FSO [Federal Statistical Office], ARE [Federal Office for Spatial Development], SDC [Swiss Agency for Development and Cooperation], & FOEN [Federal Office for the Environment]. 2016. *Sustainable development pocket statistics 2016.* Neuchâtel: Swiss Confederation. https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/publikationen.assetdetail.1101247.html. Accessed 22 March 2021.

Greenberg, Michael. 2017. Book review "the argumentative turn in policy analysis: reasoning about uncertainty." *Risk Analysis* 37 (4): 843–844. https://doi.org/10.1111/risa.12805.

Grüne-Yanoff, Till. 2007. Bounded rationality. *Philosophy Compass* 2 (3): 534–563. https://doi.org/10.1111/j.1747-9991.2007.00074.x.

Grüne-Yanoff, Till, and Sven Ove Hansson. 2009. From belief revision to preference change. In *Preference change: approaches from philosophy, economics and psychology*, ed. Till Grüne-Yanoff and Sven Ove Hansson, 159–183. New York: Springer.

Hansson, Sven Ove. 2007. Philosophical problems in cost–benefit analysis. *Economics and Philosophy* 23 (2): 163–183. https://doi.org/10.1017/S0266267107001356.

Hansson, Sven Ove. 2016. The ethics of economic decision rules. In *The Oxford handbook of professional economic ethics*, ed. George DeMartino and Deirdre N. McCloskey, 29–54. New York: Oxford University Press.

Hansson, Sven Ove, and Gertrude Hirsch Hadorn, eds. 2016. *The argumentative turn in policy analysis. Reasoning about uncertainty*. Cham: Springer. https://doi.org/10.1007/978-3-319-30549-3.

Hansson, Sven Ove, and Gertrude Hirsch Hadorn. 2018. Argument-based decision support for risk analysis. *Journal of Risk Research* 21 (12): 1449–1464. https://doi.org/10.1080/13669877.2017.1313767.

Hawkins, Troy R., Bhawna Singh, Guillaume Majeau-Bettez, and Anders Hammer Strømman. 2012. Comparative environmental live cycle assessment of conventional and electric vehicles. *Journal of Industrial Ecology* 17 (1): 53–64. https://doi.org/10.1111/j.1530-9290.2012.00532.x.

Hirsch Hadorn, Gertrude. 2016. Temporal strategies for decision making. In *The argumentative turn in policy analysis. Reasoning about uncertainty*, ed. Sven Ove Hansson and Gertrude Hirsch Hadorn, 217–242. Cham: Springer. https://doi.org/10.1007/978-3-319-30549-3_9.

Hirsch Hadorn, Gertrude. 2018. On rationales for cognitive values in the assessment of scientific representations. *Journal for General Philosophy of Science* 49 (3): 319–331. https://doi.org/10.1007/s10838-018-9403-6.

Hirsch Hadorn, Gertrude, Urs Allenspach, and Georg Brun. 2011. Warum Nachhaltigkeitsbeurteilungen Alternativen zu Kosten-Nutzen-Analysen erfordern [Why alternatives to cost–benefit analysis

in sustainability assessments are necessary]. *Schweizerische Zeitschrift Für Forstwesen* 162 (11): 412–429. https://doi.org/10.3188/szf.2011.0412.

Hirsch Hadorn, Gertrude, Georg Brun, Carla Soliva et al. 2015. Decision strategies for policy decisions under uncertainties: the case of mitigation measures addressing methane emissions from ruminants. *Environmental Science & Policy* 52: 110–119. https://doi.org/10.1016/j.envsci.2015.05.011.

Hussain, Waheed. 2018. The common good. In *The Stanford encyclopedia of philosophy* (spring 2018 edition), ed. Edward N. Zalta. https://plato.stanford.edu/entries/common-good/. Accessed 10 July 2021.

IPCC [Intergovernmental Panel on Climate Change]. 1995. *IPCC second assessment climate change 1995. A report of the Intergovernmental Panel on Climate Change. Summary for policymakers: The economic and social dimensions of climate change*. IPCC Working Group IIII. https://archive.ipcc.ch/pdf/climate-changes-1995/ipcc-2nd-assessment/2nd-assessment-en.pdf. Accessed 6 March 2021.

Jacobs, Michael. 1999. Sustainable development as a contested concept. In *Fairness and futurity: essays on environmental sustainability and social justice*, ed. Andrew Dobson, 21–45. Oxford: Oxford University Press.

Lichtenstein, Sarah, and Paul Slovic, eds. 2006. *The construction of preference*. Cambridge: Cambridge University Press.

Luce, Robert D., and Howard Raiffa. 1957. Utility theory. In *Games and decisions*, ed. Robert D. Luce and Howard Raiffa, 12–38. New York: Wiley.

Luján, José L., and Oliver Todt. 2021. Evidence based methodology: a naturalistic analysis of epistemic policies in regulatory science. *European Journal for Philosophy of Science* 11 (26): 1–19. https://doi.org/10.1007/s13194-020-00340-7.

Marttunen, Mika, Fridolin Haag, Valerie Belton et al. 2019. Methods to inform the development of concise objective hierarchies in multi-criteria decision analysis. *European Journal of Operational Research* 277 (2): 604–620. https://doi.org/10.1016/j.ejor.2019.02.039.

Millennium Ecosystem Assessment. 2005. *Ecosystems and human well-being: A framework*. Washington, DC: Island Press. http://www.millenniumassessment.org/en/Framework.aspx. Accessed 22 Mar 2021.

Pearce, David W. 1983. The origins of cost–benefit analysis. In *Cost–benefit analysis studies in economics*, ed. David W. Pearce. London: Palgrave. https://doi.org/10.1007/978-1-349-17196-5_2.

Pielke, Roger. 2014. *The honest broker. Making sense of science in policy and politics (10th printing)*. Cambridge: Cambridge University Press.

Rauschmayer, Felix. 2001. Entscheidungshilfen im Umweltbereich. Von der mono-kriteriellen zur multikriteriellen Analyse [How to decide in the context of environmental issues: from mono-criterial to multi-criterial analysis]. *Jahrbuch Ökologische Ökonomik* 2: 221–241.

Rauschmayer, Felix, and Nathalie Risse. 2005. A framework for the selection of participatory approaches for SEA. *Environmental Impact Assessment Review* 25 (6): 650–666. https://doi.org/10.1016/j.eiar.2004.09.005.

Resnik, Michael D. 2011. *Choices. An introduction to decision theory (9th printing)*. Minnesota: University of Minnesota Press.

Roy, Bernard. 2005. Paradigms and challenges. In *Multiple criteria decision analysis*, ed. José Figueira, Salvatore Greco, and Matthias Ehrgott, 3–24. New York: Springer.

Sen, Amartya. 2000a. Consequential evaluation and practical reason. *The Journal of Philosophy* 97 (9): 477–502. https://doi.org/10.2307/2678488.

Sen, Amartya. 2000b. The discipline of cost–benefit analysis. *Journal of Legal Studies* 29 (52): 931–952. https://doi.org/10.1086/468100.

Singh, Rajesh K., H. R. Murty, S. K. Gupta, and A. K. Dikshit. 2012. An overview of sustainability assessment methodologies. *Ecological Indicators* 15 (1): 281–299. https://doi.org/10.1016/j.ecolind.2011.01.007.

Stern, Nicholas. 2007. *The economics of climate change: the Stern reviews*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511817434.

Stirling, Andy. 2006. Analysis, participation and power: justification and closure in participatory multi-criteria analysis. *Land Use Policy* 23 (1): 95–107. https://doi.org/10.1016/j.landusepol.2004.08.010.

Sunstein, Cass R. 2018. *The cost–benefit revolution*. Cambridge, MA: MIT Press.

UNFCCC [United Nations Framework Convention on Climate Change]. 2008. *Challenges and opportunities for mitigation in the agricultural sector*. Technical paper No. 8. http://unfccc.int/resource/docs/2008/tp/08.pdf. Accessed 2 July 2014.

Velleman, J. David. 2014. The story of rational action. In *The possibility of practical reason (2nd edition, online version)*, ed. J. David. Velleman. Ann Arbor: Michigan Publishing, University of Michigan Library. https://doi.org/10.3998/maize.13240734.0001.001.

Wegner, Giulia, and Unai Pascual. 2011. Cost–benefit analysis in the context of ecosystem services for human well-being: a multidisciplinary critique. *Global Environmental Change* 21 (2): 492–504. https://doi.org/10.1016/j.gloenvcha.2010.12.008.

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