

SPRU Report to the SPLiCE Project

A Review of ‘Social Appraisal’ Methodologies

Josie Coburn and Andy Stirling
SPRU, University of Sussex
September 2014

1: INTRODUCTION

1.1: Task Summary

The following document is a report from the SPRU team to Output 4 of the SPLICE Project. We begin by making some general remarks concerning the scope of this Output and on the way in which we have undertaken to address each aim and task.

The aims of Output 4 are set out in the Project Case for Support as follows:

To provide recommendations on whether the very different impacts (environmental, social and economic) of a very diverse range of energy supply and demand options could be assessed and valued in a way which allows them to be compared with each other in order to assist choices to be made between them.

In order to co-ordinate consistent inputs to meeting this overall aim, the SPRU team was asked to undertake a series of specific tasks. Each is indexed here (A) – (H) in order to quickly introduce the way in which each task has been delivered.

The first task (A) was to review an open array of specific ‘social appraisal’ methods broadly grouped as multicriteria analysis (MCA), multicriteria mapping (MCM) and participatory and deliberative approaches.

In each respect we were asked (B) to summarise the purpose, strengths, weaknesses opportunities and threats (SWOT), as well as potential linkages with other methodologies. It was clarified that the apparent redundancy between the ‘SW’ and ‘OT’ aspects of ‘SWOT’ means that potential linkages can be addressed in relation to opportunities. Likewise, the category of ‘potential threats’ might usefully be used to address impacts on wider debate.

In addition we were asked (C) to review overarching issues concerning the application of the focal methods across different scales (including time and space).

A particular focus was requested (D) concerning what was referred to as “*displaced local impacts of achieving national goals*”.

We were also asked to assess (E) where these methods have been used (successfully or otherwise) in relation to considering low carbon (or any other aspects of) energy supply and demand.

A further series of relevant factors was specified to concern (F) the particular metrics that are employed in each reviewed method and how these can be or have been compared with metrics from other types of impact assessment.

It was also requested (G) that we discuss the circumstances that affect whether a social impact assessment has been useful and the factors that influence its utility.

And a final query concerned (H) “*how, if at all, do/could social impact assessments use ecosystem service frameworks?*”

1.2: Task Delivery

Given the scope and diversity of methods encompassed by Task (A), this is addressed here by selecting a series of particular approaches that collectively span the main axes of difference in this field. What is meant by this will emerge in more detail in the course of the review itself. But the particular focal methods may be summarised as the following broad categories (together with some of their most salient contrasts):

- (i) conventional externalities assessment, CEA (otherwise variously known as cost-benefit analysis, benefit-cost evaluation; social costs assessment);
- (ii) deliberative monetary valuation, DMV (adding variously-structured forms of inclusive deliberation to conventional externalities assessment methods);
- (iii) staged multicriteria assessment, MCA (broadly based on principles of multiattribute utility theory, as operationalised in multicriteria decision analysis);
- (iv) social multicriteria evaluation, SME (based on a different approach to multicriteria modelling and involving various interactive practices to augment them);
- (vi) qualitative participatory deliberation, QPD (a wide diversity variously-designed approaches to inclusive public engagement, often based on citizens’ panels);
- (v) multicriteria mapping, MCM (qualitative / quantitative comparison of open-ended elicitations, that focuses not on aggregating, but mapping divergences);
- (vii) Q Method, QMA (a distinctive open-ended hybrid qualitative/quantitative approach to mapping out contrasting perspectives on any value-based issue);

The most important single factor to bear in mind in considering these methods – and the review that follows – is that when it comes to evaluating them one compared with another, ‘*the devil is in the detail*’^{1,2}. Each may be implemented in a wide variety of different ways, subject to a range of different general and context-specific evaluative imperatives^{3 4 5}. Although there are general tendencies in each respect, these are more often a reflection of the associated disciplinary cultures, than they are of technical necessity. As a result, any process of evaluation must contend with a ‘fractal’ structure of pros and cons⁶. There is no apparently positive characteristic in respect of any possible evaluation criterion that may not be reversed by some more detailed feature of the way in which a particular method is designed or implemented. The review that follows is therefore based around the general tendency in practice.

Task B concerning the strengths, weaknesses opportunities and threats is obviously subject to the above qualification. It is equally obviously a matter of perspective. In a highly charged policy arena such as that in which these techniques are applied, the values under which the methods themselves are appraised are just as subject to political controversy as the issues to which they are applied⁷. So what count as ‘strengths’ and ‘weaknesses’ will – to a large extent – be in the eye of the beholder.

This review addresses this challenge by means of the requested SWOT table, with entries given explicitly in relation to particular political aims. As specified, this table includes as opportunities the issue of 'linkages with other methods'. Under 'threats' it addresses as suggested, the question of potential impacts on wider public value based controversies. By way of further crucial context, before this table, the intended purpose of each method is elaborated further in the next section.

Task C concerning applicability to different scales is included – where relevant – in the next section concerning purpose. But a crucial point to bear in mind is that all the specific methods addressed here (as well as the general approaches they represent) have been variously used at virtually every spatial level and temporal scale ⁸. There is nothing intrinsic to the applicability of these different methods in this regard. Observable patterns of preference are more a reflection of cultural styles in the responsible disciplines, than any inherent features of the methods themselves.

Task D highlights in local/national trade-offs just one instance of what best be addressed in more general terms as the handling of issues of distribution and representation ⁹. These are similar to the topic of 'scale' in the questions they raise. In short, there is nothing intrinsic about any of the focal methods that makes any of them inherently or self-evidently more favourable on this count. Preferences in this regard will reflect the precise ways in which the methods are implemented, as well disciplinary affiliations and biases in how they are viewed. In principle, each approach can be designed in different ways, such as to address distributional issues like those concerning displaced local impacts ^{10 11 12}.

Task E concerning generally established patterns of usage across different policy areas, raises similar issues (and for the same reasons) to those discussed above in relation to Tasks C and D. In broad terms, every one of the selected methods has been used in some context and fashion to appraise strategic issues in energy policy in general, and low carbon transitions in particular. All those selected here are in principle highly applicable in this area. For instance, indicative examples of use in energy policy include CEA ^{13 14 15 16 17}; MCA ^{18 19 20 21}; SME ^{22 23} and QPD ^{24 25 26 27} ²⁸. The references given in each respect also reflect some of this diversity.

The question of chosen metrics addressed in task F is closely related to the purpose of the methods in each case, so is addressed as part of the summary for Task A in Section 2.

Task G concerning the 'utility' of the different methods raises challenges very similar to those discussed in relation especially to Task D above. Any unqualified expression of merit orders across these (or any other methods) would reflect the subjective values and assumptions on the part of the evaluator more than the inherent characteristics of the methods themselves. This is discussed especially in Section 2.5. Each mode of implementation reflects different fundamental notions of what might constitute the 'utility', usefulness, appropriateness or value of the technique itself and those with which it might be compared. So these central evaluative challenges must necessarily be addressed in a 'plural and conditional' fashion, rather than a matter of absolute or definitive objectivity ²⁹. This is how Task A has been undertaken.

Task H implies a very specific question over the extent to which each method is consistent with ecosystem service frameworks. Again, the inherent complexities and realities of this field of approaches and their political context mean (as already

discussed above), that any answer is likely to tell more about the perspective under which evaluation is conducted, than the material being evaluated. In short, all of the reviewed techniques are susceptible to being used in some way with ecosystem service frameworks. The main question that might be posed in this regard is 'why?'

It is arguable that many of the methods reviewed here actually constitute preferable means to appraise ecological values themselves (when compared with a narrow utilitarian 'service' framework)³⁰. Some of the reasons for this are explicated in the course of discussing the rest of this review. Whether or not it is agreed with, this does mean that an exercise like the present project should be extremely careful about uncritically accepting an instrumental service framework as self-evidently appropriate even in relation to ecosystems themselves, let alone in relation to wider societal dimensions of energy strategies.

2: BACKGROUND, PURPOSE AND METRICS

2.1: Conventional Externalities Assessment (CEA)

The family of techniques referred to here as 'conventional externalities assessment' are often referred to collectively as benefit-cost or cost-benefit analysis³¹. These share a central common feature, in that the issues in relation to which all options are appraised, are typically addressed by means of the single (apparently simple) **metric** of monetary value. The reference to 'externalities' reflects the fact that many of the most important monetary values involved, are 'external' to existing markets³². They must therefore be elicited, inferred or modelled by means of various technical procedures³³. In cases where the issues under scrutiny are those related to provision of ecological services, then the resulting 'ecosystem service frameworks' constitute a member of this broad family of methods³⁴.

The **purpose** of doing this is to deal with the massive practical inconvenience for the purposes of justifying particular policies, that the various relevant issues are typically 'incommensurable'³⁵. In other words, they are 'apples and oranges' – intrinsically not subject to aggregation under a single metric. So, what CEA apparently offers in this regard, is a way to reconcile this fundamental impossibility³⁶. By assigning a single distribution of monetary values across all such issues, it renders alternative policy options not only roughly comparable in broad qualitative terms, nor just as subject to a neat approximate ordinal (i.e.: relative) sequence; but as apparently amenable to an unambiguous ordering on a precise cardinal (i.e.: quantitative) scale – expressing the exact ratios and intervals separating the overall merits of each policy option.

The trouble is, that it has been demonstrated in Nobel Prize winning analysis in the field of welfare economics and rational choice theory underlying these methods, that even the aspiration (let alone the claim) of a single definitive ordering of incommensurable issues is, in a plural society, not only impossible in practice to guarantee – but inherently meaningless even to contemplate^{37 38 39}. So, the

apparent policy utility comes at the price of serious inaccuracy in relation to real-world complexities, uncertainties and subjectivities.

Nonetheless, in terms of pros and cons, CEA can still hold strong attractions under an instrumental perspective – where it is simply assumed that the aim is automatically to comply with the aims of clients (or wider incumbent interests in any given controversy), whether this be government, business or an NGO ⁷. In such cases, few policy rhetorics are more potent than one expressing unequivocal confidence over the aggregation of incommensurable issues and expressing these in the familiar and highly operational terms of monetary value, without acknowledging any uncertainty or ambiguity ⁴⁰. But it follows from this same apparent strength under one view, that there exists a corresponding serious weakness under other views. These are, that CEA in all its forms serves effectively to suppress uncertainty, deny ambiguity and force one particular evaluative perspective at the expense of others – thus undermining both science (which it misrepresents) as well as democracy itself ⁴¹.

2.2: Deliberative Monetary Valuation (DMV)

In this family of methodologies, various deliberative processes are used alongside conventional optimisation analysis, to assign a **metric** of monetary value to reflect the performance of a range of policy options across a set of relevant issues ⁴². The main purpose is to address some of the acknowledged challenges summarised above in the case of more purely analytical forms of CEA (Section 1.2) ⁴³.

Depending on how DMV is reported, the main general difference with CEA is that the process of assigning these monetary values can be more transparent in relation to diverse extant social and political perspectives ⁴⁴. And the particular values assigned are much more subject to the agency of those participants who are able to engage in the process. Processes of deliberation can offer a learning experience for those involved. And they open the possibility of subjecting the final results to some kind of sensitivity analysis to reflect the divergent views expressed during the process of deliberation. This is not usually undertaken, but might in principle be reconstructed by a third party in order to reveal some of the concealed ambiguities.

However, in the end, DMV produces as an output, the same kind of discrete arrays of monetary values that are typically produced in CEA. So it is in principle subject to the same kinds of concern addressed in that case (Section 2.1). In this regard, a review by Stagl for DEFRA concludes “*deliberative monetary valuation is most suitable for the appraisal of projects whose impacts are relatively well understood, where the impacts do not reach far into the future, and which do not affect complex ecosystem services such as biodiversity*” ⁴⁵.

2.3: Staged Multicriteria Analysis (MCA)

The label used above for this family of methods is the one chosen by Stagl in a useful review for DEFRA ⁴⁵. This term is applied here, as it was by her, to address a diverse array of multicriteria techniques ^{46 47} – also reviewed elsewhere in detail for DEFRA ⁴⁸. There exist many divisions within this field, some of which lead to methods that may derive contrasting findings when applied to the same kinds of policy challenge. But a common feature of all these methods, is a move away from the apparently unambiguous (but as we have seen, potentially highly misleading) metric of monetary value. The **metrics** used instead are more abstract measures of relative value, as variously produced by each method.

This said, all these methods share a basic overall **purpose**, which is to further address the challenges of incommensurability described in relation to CEA in Section 2.1. This is done by affording variously more sophisticated ways to explore the implications of divergent priorities and values – and sometimes uncertainties – across contrasting social perspectives. The ‘three stages’ sometimes referred to in the labelling of this methodology, is simply one means by which this can be achieved – as set out in a particularly relevant approach developed and widely applied in Germany by Renn and Webler ^{49 50}, including to energy issues.

In this form, MCA uses a ‘co-operative discourse’ approach, in which key uncertainties and ambiguities in appraisal are addressed in distinct ways at different stages of analysis. In short, the evaluation criteria are selected in advance by major stakeholders. The scoring of benefits and impacts under these criteria is undertaken entirely by experts. Here, there is some attention to uncertainties. But these are generally treated as if they were relatively tractable ‘risks’ (i.e.: amenable to the assigning of probabilities) ⁵¹. So – as is typically the case in CEA and DMV – uncertainties of a more challenging kind are correspondingly excluded from analysis ³⁵.

Crucial to this kind of approach, is that the input of citizens is restricted to the exploration of alternative values. So, there is no opportunity for citizens to question the scope or structuring of issues as determined by major stakeholders. And the scoring of benefits and impacts by experts remains similarly inaccessible to interrogation. Since expression of uncertainties is also somewhat reduced (as discussed above), this leaves MCA to be rather circumscribed in its ability to explore a full range of alternative perspectives and possibilities ⁵².

In her earlier study for DEFRA, Stagl found that three-stage MCA is most appropriately used when the impacts of a policy, programme or project are reasonably well understood by experts but where there is a significant technical component ⁴⁵. But this kind of application leaves unaddressed the issues of ambiguity and uncertainty mentioned above.

2.4: Social Multicriteria Evaluation (SME)

Social multicriteria evaluation differs from MCA in a number of ways that are important in methodological terms, but often less so in respect of the practical implications for policy and wider political debate. It was mentioned in discussing MCA (Section 2.3), that the field of multicriteria analysis is divided between many divergent approaches. Arguably the single most important such division is between a 'Francophone' tradition involving a procedure for pairwise comparison between options as compared under each criterion and an 'Anglophone' tradition more directly based on conventional utility theory and neoclassical ideas of rational choice⁵³.

Like MCA – and unlike CEA and DMV – SME makes use of an abstract value *metric*⁵⁴. But the purpose in this case is to address more specifically than do any of these techniques, the challenges of complexity, ambiguity and uncertainty. SME does this by combining participatory techniques with a pairwise comparison approach to multicriteria analysis. This affords greater agency to participants of all kinds (including citizens) to frame the ways in which different policy, programme or project options are taken into account – and how conflicting interests are handled⁵⁵.

A particular focus of SME lies in the provision of transparency – both to participants and third parties. It is intended that this help foster 'social learning' – so that the appraisal exercise itself is not simply about the outputs that are produced, but also about the process in which the different parties are engaged⁵⁴.

In her earlier analysis for DEFRA, Stagl finds that "[t]his method is most suitable for the appraisal of policies, programmes or projects whose impacts are not well understood yet and therefore benefit from a multidisciplinary modelling of impacts"⁴⁵. In these terms, Stagl is referring more to a 'transdisciplinary' than a 'multidisciplinary' value, since the latter is more closely shared with the other techniques reviewed earlier here⁵⁶. After all, BCA and DMV are equally typically multidisciplinary (although the typical dominance of economics in the framing of the method means they are less interdisciplinary than MCA or SME). If it is used to involve citizens in more transparent, respectful and less circumscribed ways as aimed, then SME may also by this means claim some degree of 'transdisciplinarity'^{57,58}.

However, it remains the case that under principles of rigour shared by all rational choice approaches (including CEA, DMV and MCA), SME can be argued to display serious methodological deficiencies. In some circumstances, these can lead to artefacts in the ranking process, such as rank reversals that may confuse or undermine confidence⁵⁹. Despite the positive efforts that distinguish this approach over the others mentioned here, SME may also be challenged concerning the extent and depth in which it permits participants to explore the full range of ambiguities and uncertainties. So it may correspondingly prove limited in the degree to which it can

deliver on the aims of social learning. But SME remains favourable in comparison with all techniques reviewed thus far, in relation to these particular aims.

2.5: Qualitative Participatory Deliberation (QPD)

Under this category of approach, are included an enormous diversity of different methods ranging variously through focus groups, citizen's panels, stakeholder negotiation, interactive modelling, community visioning, do-it-yourself juries, open space, consensus conferences^{60 61 62 63 64 65}. Each particular method typically displays a variety of quite radical contrasting alternative ways in which it can be designed, commissioned, recruited, framed, bounded, overseen, focused, facilitated, staged, structured, reported, evaluated and articulated with other methods and with policy debates. Each of these attributes in the constituting of any particular instance of qualitative participatory deliberation is spelled out explicitly here, because each presents a dimension on which (as explained in Section 1.1) '*the devil is in the detail*' in any attempt to draw general analytic or evaluative conclusions⁶⁶.

Compounding this complexity, is the fact that virtually any particular method of participatory deliberation (and any detail of the above kind in the implementation of each), can also be combined with any of the other methods reviewed in this survey. For instance, one particular approach to QPD reviewed by Stagl for DEFRA is 'stakeholder decision analysis' (SDA)^{45 67}. This initially employed a qualitative form of multicriteria analysis. The method was later articulated with the quantitative procedure at the heart of MCM to form the synthetic approach called 'deliberative mapping' (DM)^{68 69}. And DMV and SME inherently involve the use of some kind of participatory deliberation in association with their own quantitative procedures. Perhaps most flexible of all, 'multicriteria mapping' (MCM) can be used as an adjunct to some variant of virtually any of the broad participatory approaches identified above (Section 2.6)⁷⁰.

The task of generalising mentioned repeatedly here, is therefore especially difficult in respect of this category of approach. However, the bottom line response in relation to the key queries of interest here are as follows. The fundamentally qualitative nature of these processes means that the issue of metrics remains secondary. In short, it is possible to make use of any **metric** that might be considered relevant, but it is recognised that any comparison across different metrics will be subject to qualitative considerations – and that these properly form the central focus of appraisal. For those for whom adherence to a particular single metric is a matter of principle, then, all forms of QPD will tend to be seen as correspondingly deficient.

With regard to **purpose**, there arises another important point. In common with the real-world implementation of all other appraisal methods considered here, the underlying purpose of appraisal will typically differ as between different participants. Powerful incumbent interests will typically wish to use the exercise to justify some

policy decision, such as to enhance the degree to which they are trusted, increase acceptance of their interests and reduce the risks attached to dissent and protest^{40,71}. This may, or may not, involve a desire to assert a particular pre-conceived referred decision⁷.

Likewise, various stakeholder interests will wish to use appraisal as part of wider strategies to assert particular favoured policy outcomes – or to give a voice to perspectives that they have reason to believe are otherwise excluded. Practitioners of appraisal will typically experience great pressure to align with one or other of these powerful interests. But they may also hold strong interests of their own – for instance in broadening out the scope of appraisal and ‘opening up’ the picture given to policy concerning the implications of contrasting perspectives⁷². For their part, ostensibly ‘disinterested’ participants like ordinary citizens will typically always have their own biases and enthusiasms – and will (like policy makers) often wish to gain a sense of satisfaction in contributing to a tangible policy outcome, sometimes even if this is at the expense of reducing complexity.

In general then, there is with participatory deliberation as with other approaches to appraisal, a need to be cautious about attributing any single clear-cut ‘purpose’^{73,74}. Even individual perspectives may oscillate in complex ways between an instrumental purpose, aiming at some particular pre-conceived ‘right decision’, or a substantive purpose of finding in an open, balanced way what looks like the ‘the best decision’ under different views; and/or a normative purpose in ensuring that whatever method is used (and irrespective of the outcomes), the process itself is conducted appropriately⁷. As with other methods, it is impossible to evaluate participatory deliberation in abstract terms, without being explicit as to the particular purpose.

2.6: Multicriteria Mapping (MCM)

Multicriteria mapping constitutes one attempt to address all the issues raised thus far in this review – spanning qualitative and quantitative approaches; giving balanced attention to an unconstrained array of issues and options; allowing participants ‘to be in the driving seat’ (without steering or constraining them in particular directions); but at the same time imposing clear principles of rigour in the ways that options and issues are appraised and the transparency with which this is conveyed to third parties for peer review^{70 75}. It is recommended in a DEFRA manual as an especially effective means to address these kinds of challenges⁴⁸.

At the same time MCM is distinct from all other methods reviewed here, in taking the fullest account of uncertainties and ambiguities and clearly expressing these in the final result – serving to help ‘open up’ the practical policy implications of divergent values, assumptions and contexts⁷⁶. In short, MCM aims at the same time rigorously and accountably to deliver on all three kinds of purpose discussed in Section 2.5: instrumental (in allowing expression of particular interests); substantive (in

illuminating the diverse concrete implications for decision making) and normative (in doing this in ways that are compliant equally with quantitative and qualitative principles of rigour) ⁵².

Like other methods, MCM can be implemented in various ways in order to meet different instantiations of these aims in contrasting contexts. It can be used purely as an interview technique, with deliberation carried out later on the basis of presentation of the qualitative and quantitative results. And – since the basic tool is an accessible form of web-based software – this may (depending on the aims) be undertaken either in person or remotely. With due caution, either approach can be combined and integrated with variously-staged group based deliberative processes. And as part of this, MCM can use as inputs (among others), the outputs of any of the other methods reviewed here – or scientific environmental assessment techniques. Likewise, it may itself be taken as an input to exercises in participatory deliberation reviewed in the previous section (Section 2.5) ⁷⁷.

On the positive side, MCM is relatively broad and flexible in scope, avoiding the imposition of constraints on the type of issue that can be taken into account or the way they can be defined. This contrasts with other multicriteria techniques where appraisal is virtually always based exclusively on utilitarian trade-offs, where options and even criteria are sometimes prescribed in advance, where participants' criteria are often aggregated on a single 'value tree', where scoring is usually performed by a narrow specialist group, leaving citizen or stakeholder input restricted to criteria definition and weighting. These features allow MCM to faithfully reflect perspectives from a wide range of different participants without imposing undue constraints or engendering counterproductive tensions ⁷⁵.

On the negative side, MCM in itself and as it stands is still largely an individual interview-based tool. The interview process is quite demanding. Unless special additional arrangements are made, provision for effective group deliberation and citizen (rather than specialist) engagement can be limited. These deficiencies are readily addressed by incorporating MCM into a broader process providing both for citizen participation and intensive in-depth group deliberation involving both citizens and specialists. This more elaborate approach is termed 'Deliberative Mapping' ⁶⁸.

It is important to recall, though, that under instrumental objectives prioritising the securing of justification, the distinctive degree of 'broadening out' and 'opening up' offered by MCM ⁷² can be viewed as a disadvantage. Although MCM can be used to illuminate a single 'average' picture of rankings across all perspectives, this is qualified by transparent acknowledgement of the degree to which this picture varies. Correspondingly, the lack of such transparency in other methods (like CEA, DMV, MCA and even consensus-oriented participatory approaches), can be seen as an advantage if the aim is simply to justify decisions ⁷⁸.

2.7: Q Method (QME)

Originating in social psychology ⁷⁹, Q Methodology is a powerful, mature and well-established approach, which unusually (like MCM) combines hybrid quantitative and qualitative dimensions ^{80 75}. It is particularly easy and cost-effective to implement. However, the style tends to be less well oriented to addressing the comparative performance of concrete policy options. Instead, the strengths of Q method lie in illuminating key distinct perspectives concerning the divergent reasons why different possible policies might be considered positive or negative. It is especially powerful as a way of identifying associations between contrasting ostensibly entirely separate enthusiasms and concerns. This can be useful, where the purpose is to understand better how different perspectives relate to each other ⁸¹.

Q method is based on the compilation by the analyst of a large set of short clear statements on an issue in question, drawn from a rich diversity of sources and perspectives and covering a full envelop of the different evaluative dimensions associated with contrasting policy options. Engaged in relatively short individual interviews, representative individuals are recruited from a wide range of different perspectives to order these statements according to how much they agree or disagree with each. The results of these 'Q sorts' are processed statistically to reveal the degree to which positions on different dimensions associated and diverge from each other. As a result, Q can be very effective at identifying "similarities among individual attitudes, *which may not have been known a priori*" ⁸² (p. 35, emphasis added).

The **metric** used in Q, such as it is, then, is an abstract measure of proximity and distance between perspectives. Like MCM, the **purpose** is as much to illuminate diversity and distinguish between contending reasons for different possible actions, rather than to focus single-mindedly on aggregated 'best practice' policies.

The main disadvantage of Q, in these terms, is that it is not primarily designed for application directly to alternative policy interventions. It is also not so much a directly interactive deliberative method – the kinds of learning that it can contribute to are more individual or collective on the basis of the results. There is typically no group interaction as part of the process itself. It is more effectively used to cast light on the divergent conditions and pros and cons associated with a set of policy options taken as a whole. However, by using Q to differentiate between contrasting perspectives and identify their principal priorities and concerns, it can be used as a powerful means to inform the recruitment of participants for other methods involving public engagement – like DMV, MCA, SME, QPD and MCM.

SWOT Table

This table summarises in the recommended form, the substantive material dealt with in more detail in the above narrative account.

Impact Evaluation Frameworks	Type of outputs and indicative costs (k£) (on basis of comparable modest sized policy scoping project)	<u>S</u> trength	<u>W</u> eakness	<u>O</u> pportunity	<u>T</u> hreat
Conventional Externalities Assessment (CEA)	Ostensibly precise rankings across policy options based on familiar-sounding monetary values. 20-40	The major strengths (where these aims apply) lie in the justificatory power for policy making, and associated capacity for narrowing down the scope of assessment and closing down associated political debates. Also where this aim applies, a further major strength is the credibility associated with use of the familiar and ostensibly objective metric of monetary value.	Major weaknesses are the lack of scope for public and stakeholder engagement; inadequate treatment of uncertainty and ambiguity; lack of transparency in reflecting key sensitivities; lack of robustness in relation to aggregation and – where these aims apply – a tendency to narrow in appraisal and close down associated political debates.	Because this claims directly to derive a metric that is generally comparable, the associated body language and culture can (however details are expressed) seriously undermine scope for appreciating the value of other techniques. So - other than a means to address specific economic issues in MCA, SME and MCM, CEA does not fit well with other methods.	The fact that uncertainties and ambiguities are concealed in ostensibly precise and definitive monetary values has the effect of seriously suppressing the quality of policy debates and wider democratic discourse in the field.
Deliberative Monetary Valuation (DMV)	Ostensibly precise rankings across policy options based on familiar-sounding monetary Values. Some learning as part of process. 30-60	The major strengths (where these aims apply) lie in the justificatory power for policy making, and associated capacity for narrowing down the scope of assessment and closing down associated political debates. The scope for deliberation, illumination of divergences and process learning are also additional strengths. And there is also the benefit of using, but somewhat diminished by the reduced credibility fostered by deliberation revealing the	Although labelled as “deliberative”, the constraints of the monetary focus mean this is significantly less so in comparison with other methods (like SME, QPD and MCM) This can be seen as combining the worst of all worlds, in compromising on the justificatory power of CEA, whilst adding further complexity but failing to deliver the corresponding benefits of flexibility, transparency and robustness associated with multicriteria or participatory	There is greater openness than in CEA for articulation with other techniques, because the deliberative aspect allows greater latitude than is the case in rigid calculative externalities assessment, for taking account of the qualitative implications of the pictures yielded by different appraisal methods. But, in the end, the monetary idiom raises similar difficulties.	If sensitivity analysis is used to convey the implications of divergent perspectives in the reporting of results, then this can be less threatening than conventional externalities assessment in respect of the issues raised for CEA.

		shortcomings of this metric.	deliberative methods.		
Staged Multicriteria Assessment (MCA)	Complete ranking of policy options as derived from a highly structured and stylised process and subject to weighting of different views. 20-50	Major strengths (where these apply) are shared with CEA and DMV – in the capacity to close down and so justify particular decisions. That this is associated with a greater degree of flexibility and transparency can under some views enhance these benefits. The more fully the method is combined with participatory processes, the more this applies.	A bit like DMV, the compromise on different evaluative imperatives embodied in MCA mean that it can ‘fall between stools’. With respect to the aim of ‘opening up’, it lacks the flexibility breadth, transparency of MCM and the robustness and social learning associated with participatory approaches. But it also lacks the power to justify decisions, often associated with a monetary metric.	Due to the focus on closing down, there may under some views be a strong responsibility to link such approaches with other methods that open up wider perspectives to ensure more, democratically accountable policy making.	If sensitivity analysis is used to convey the implications of divergent perspectives in the reporting of results, then this can be less threatening than conventional externalities assessment in respect of the issues raised for CEA.
Social Multicriteria Evaluation (SME)	Complete or partial ranking of a particular aggregated picture of policy options and some learning as part of process. 30-60	This can display essentially the same strengths as DMV, but with additional scope for stakeholder engagement, transparency and social learning. But it is less strong (where this aim applies) with regard to the use of the familiar and ostensibly objective metric of monetary value.	In many ways, this displays a similar set of compromises to MCA, but weighted somewhat more favourably on the side of opening up and learning and somewhat less so on the side of closing down and justification decision.	The relative neutrality of these methods allow greater potential for articulation with other methods than above. But – to the extent that most remain utilitarian – these methods tend to exclude a full account of in-principle issues that are covered in QPD or MCM approaches.	The problem identified with respect to opportunities to the left presents a threat of a lesser degree but similar kind to that discussed above.
Qualitative Participatory Deliberation (QPD)	Deliberated consensus with strong learning as part of process and some possibility for selected illumination of dissenting views and reasons for key divergences. 30-60	A diverse array of methods with different detailed strengths. If conducted in an open fashion, all tend to display high flexibility and robustness in relation to participants’ interests. Depending on how undertaken, can also promote strong learning. But if closed down in order to deliver	Under a viewpoint prioritising quantitative procedure as an end in itself, the purely qualitative form of these approaches is a serious disadvantage. Conversely, where value is attached to rigorous exploration of uncertainties and ambiguities and the opening up policy debates, then the frequent focus of these methods on closure	These techniques can be used as an overarching way to take account of the outputs of any other approach discussed here.	Where the design, framing or implementation QPD are used to close down policy debates and justify particular decisions, then they can present essentially the same threats (albeit in different form) that are noted for CEA. However, there can be mitigating factors, in that the widespread practice of QPD can help foster the (albeit

		legitimation, justification or consensus, then these 'strengths' (where seen as such), can bring reduced legitimacy, transparency and learning.	and consensus can also be a serious disadvantage – in the extreme broadly comparable to CEA.		unintended) effect of raising wider expectations concerning responsibilities to be transparent and responsive in policy making.
Multicriteria Mapping (MCM)	Map of contrasting rankings of policy options under different perspectives (including an overall average), as well as illumination of related uncertainties, ambiguities and relevant values plus discourse analysis and learning as part of the process. 20-60	Major strengths lie in the flexibility and broadening out of the scope of appraisal and the opening up of policy debates. Quantitative data is substantiated by rich qualitative material and a rigorous and transparent picture of uncertainties, ambiguities and divergent values. The process also fosters significant learning.	The main weakness of MCM (where this quality is seen as such), is the fact that a technique aimed at opening up policy debates, can have the effect of destabilising closure and the justification of particular decisions. The rigorous exploration of different aspects of appraisal required in this process can also be demanding both for analysts and participants.	The flexibility of MCM in addressing in a balanced way such a diversity of input, issues and perspectives, means it offers an especially strong tool for integrating the outputs of a range of other methods.	There is no guarantee that MCM will be implemented in accordance with its own driving principles. Like other kinds of MCA and SME, it can be used in a narrow fashion to close down decisions. Where this occurs, some of the design features offer greater opportunities for interrogation, but the explicit aims do lend a particular vulnerability to legitimation.
Q Method (QME)	More aligned to illuminating underlying issues rather than policy options. Clear differentiation of principal divergent perspectives on a problem. 10-20	A very effective way to scope out underlying issues and illuminate how different perspectives and aspects relate to each other – possibly revealing associations and distinctions that are entirely unexpected. In this sense, Q method can be a powerful aid to opening up decisions. Q is also a relatively quick and easy to implement.	Q is not primarily geared towards a focus on concrete policy actions, but rather at understanding the issues and perspectives that determine how these are viewed.	Q can offer a powerful tool to inform the recruitment of participants in participatory processes and the framing of issues and options for deliberation and analysis alike..	The complex and ambiguous configurations of perspectives that can be identified in Q, can be difficult to interpret or to relate in concrete ways to practical policy choices.

References

1. Mohr, A., Raman, S. & Gibbs, B. *Which publics? When? Exploring the policy potential of involving different publics in dialogue around scienc and technology.* (2014).
2. Bussu, S., Davis, H. & Pollard, A. *The best of Sciencewise reflections on public dialogue.* (2014).
3. Dryzek, J. S. *Social Choice Theory and Deliberative Democracy : A Reconciliation.* (2002).
4. Rowe, G. & Frewer, L. J. Public Participation Methods: A Framework for Evaluation. *Sci. Technol. Human Values* **25**, 3–29 (2000).
5. Bohmann, J. *“Public Deliberation: pluralism, complexity and democracy.”*(MIT Press, 1996).
6. Stirling, A. Pluralising progress: From integrative transitions to transformative diversity. *Environ. Innov. Soc. Transitions* **1**, 82–88 (2011).
7. Stirling, A. “Opening Up” and “Closing Down”: Power, Participation, and Pluralism in the Social Appraisal of Technology. *Sci. Technol. Hum. Values* **23**, 262–294 (2008).
8. Giampietro, M. *Multi-Scale Integrated Analysis of Agroecosystems.*
9. O’Neill, J. Representing people, representing nature, representing the world. *Environ. Plan. C Gov. Policy* **19**, 483–500 (2001).
10. *Equality , Participation and Inclusion 1: diverse perspectives.*
11. *Justice and Democracy.*
12. *Inclusion, Participation and Democracy: what is the purpose.* (Kluwer, 2003).
13. ExternE. *Externalities of Energy - Methodology 2005 Update.* (2005).
14. Sundqvist, T. & Soderholm, P. Valuing the Environmental Impacts of Electricity: a critical survey.
15. Jensen, D. & Al., E. *Studies of the Environmental Costs of Electricity September 1994.* (1994).
16. Stirling, A. Limits to the Value of External Costs. *Energy Policy* **25**, 517–540 (1997).
17. St, G. Energy Policy and Externalities : An Overview David PEARCE. 1–19 (2001).
18. Trutnevyte, E. The allure of energy visions: Are some visions better than others? *Energy Strateg. Rev.* 1–9 (2013). doi:10.1016/j.esr.2013.10.001
19. Robinson, J. B. Of Maps and Territories The Use and Abuse of Socioeconomic Modeling in Support of Decision Making. *Technol. Forecast. Soc. Change* **42**, 147–164 (1992).
20. Keeney, R. L., Renn, O. & Winterfeldt, D. von. Structuring West Germany’s Energy Objectives. (1987).
21. Mirasgedis, S. & Diakoulaki, D. Multicriteria analysis vs. externalities assessment for the comparative evaluation of electricity generation systems. *Eur. J. Oper. Res.* **102**, 364–379 (1997).
22. Kurka, T. Application of the analytic hierarchy process to evaluate the regional sustainability of bioenergy developments. *Energy* 1–10 (2013). doi:10.1016/j.energy.2013.09.053
23. Saaty, T. L. & Vargas, L. G. *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks.*
24. Prikken, I. Future national energy mix scenarios : public engagement processes in the EU and elsewhere. (2012).
25. Chilvers, J. & Longhurst, N. *Participation, politics and actor dynamics in low carbon energy transitions: report of a Transition Pathways Project workshop.* 21–22 (2012).
26. Giurco, D., Cohen, B., Langham, E. & Warnken, M. Backcasting energy futures using industrial ecology. *Technol. Forecast. Soc. Change* **78**, 797–818 (2011).
27. Einsiedel, E. F., Boyd, A. D., Medlock, J. & Ashworth, P. Assessing socio-technical mindsets: Public deliberations on carbon capture and storage in the context of energy sources and climate change. *Energy Policy* **53**, 149–158 (2013).
28. Raven, R. P. J. M., Mourik, R. M., Feenstra, C. F. J. & Heiskanen, E. Modulating societal acceptance in new energy projects: Towards a toolkit methodology for project managers. *Energy* **34**, 564–574 (2009).

29. Stirling, A. Opening Up the Politics of Knowledge and Power in Bioscience. *PLoS Biol* **10**, e1001233 (2012).
30. Sagoff, M. *The Economy of the Earth: philosophy, law and the environment*. (Cambridge Univ Press, 2008).
31. Byrd, D. M. & Cothorn, C. R. *Introduction to Risk Analysis: a systematic approach to science based decision making*. (Government Institutes Press, 2000).
32. Ottinger, R. L. Incorporating Externalities: the wave of the future. (1993).
33. Hanley, N. & Spash, C. *Cost-Benefit Analysis and the Environment*. (Edward Elgar, 1993).
34. Common, M. S. & Stagl, S. *Ecological economics: an introduction*. (Cambridge University Press, 2005).
35. Stirling, A. in *Negot. Chang. new Perspect. from Soc. Sci.* (Berkhout, F., Leach, M. & Scoones, I.) (Edward Elgar, 2003).
36. O'Neill, J. *Ecology, Policy and Politics*. (Taylor & Francis, 1993). doi:10.4324/9780203416570
37. Arrow, K. J. *Social Choice and Individual Values*. (Yale University Press, 1963).
38. Kelly, J. S. *Arrow Impossibility Theorems*. (Academic Press, 1978).
39. MacKay, A. F. *Arrow's Theorem: the paradox of social choice - a case study in the philosophy of economics*. (Yale University Press, 1980).
40. Collingridge, D. *Critical Decision Making: a new theory of social choice*. (Frances Pinter, 1982).
41. Getzner, M., Spash, C. & Stagl, S. *Alternatives for environmental valuation*. (Routledge, 2005).
42. Spash, C. L. Deliberative Monetary Valuation and the Evidence for a New Value Theory. (1999).
43. Aldred, J. Are the Alternatives to the Contingent Valuation Method Any Improvement? (1996).
44. O'Neill, J. & Spash, C. L. *Conceptions of value in environmental decision-making*. *Environ. Valuat. Eur.* (2000).
45. Stagl, S. *SDRN Rapid Research and Evidence Review on Emerging Methods for Sustainability Valuation and Appraisal*. (2007).
46. Treasury, H. M., ILGRA & Spackman, M. *The Setting of Safety Standards: a report by an interdepartmental group and external advisers*. (HM Treasury, 1996).
47. Clemen, R. T. *Making Hard Decisions: an introduction to decision analysis*. (PWS Kent Publishing, 1991).
48. Dodgson, J., Spackman, M., Pearman, A. & Phillips, L. *DTLR multi-criteria analysis manual*. (2003).
49. Renn, O., Webler, T., Rakel, H., Dienel, P. & Johnson, B. Public participation in decision making: A three-stage procedure. *Policy Sci.* **26**, 189–214 (1993).
50. Renn, O. & Webler, T. in *Abfallpolitik im Koop. Diskurs. Bürgerbeteiligung bei der Standortsuche für eine Deponie im Kant. Aargau* (Renn, O., Kastenholz, H., Schild, P. & Wilhelm, U.) (Hochschulverlag AG an der ETH Zürich, 1998).
51. Stirling, A. Risk at a turning point? *J. Environ. Med.* **1**, 119–126 (1999).
52. Stirling, A. Analysis, participation and power: justification and closure in participatory multi-criteria analysis. *Land use policy* **23**, 95–107 (2006).
53. Salo, A., Gustafsson, T. & Ramanathan, R. Multicriteria methods for technology foresight. *J. Forecast.* **22**, 235–255 (2003).
54. Munda, G. Social multi-criteria evaluation: Methodological foundations and operational consequences. *Eur. J. Oper. Res.* **158**, 662–677 (2004).
55. Munda, G. & Russi, D. Social multicriteria evaluation of conflict over rural electrification and solar energy in Spain. *Environ. Plann. C. Gov. Policy* **26**, 712 (2008).
56. *Reflexive Governance for Sustainable Development*. (Edward Elgar, 2006).
57. Nowotny, H., Scott, P. & Gibbons, M. "Mode 2" Revisited: The New Production of Knowledge. *Minerva* 179–194 (2003).

58. Gibbons, M. *et al.* *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. (Sage, 1994).
59. Salo, A. & Punkka, A. Rank inclusion in criteria hierarchies. *Eur. J. Oper. Res.* **163**, 338–356 (2005).
60. Smith, G. *Democratic Innovations: designing institutions for citizen participation*.
61. Fischer, F. Democracy and Expertise: reorienting policy inquiry.
62. OECD. *Citizens as Partners: information, consultation and public participation in policy making*.
63. Grove-White, R., Macnaghten, P., Mayer, S. & Wynne, B. "Uncertain World: genetically modified organisms, food and public attitudes in Britain."(1997).
64. Blamey, R. K., James, R. F., Smith, R. & Niemeyer, S. J. Citizens' juries and environmental value assessment. *Canberra, Aust. Natl. Univ.* (2000).
65. Paper, S. W. Empowering Designs: towards more progressive appraisal of sustainability.
66. Leach, M., Scoones, I. & Stirling, A. *Dynamic Sustainabilities: technology, environment, social justice*. (Routledge, 2010).
67. Burgess, J. in *Cult. Turns/Geographical Turns* (Naylor, S., Ryan, J., Cook, I. & D.Crouch) (Prentice Hall, 2000).
68. Davies, G. *et al.* *Deliberative Mapping : Appraising Options for Addressing " the Kidney Gap ."*1–10 (2003).
69. Burgess, J. *et al.* Deliberative mapping: a novel analytic-deliberative methodology to support contested science-policy decisions. *Public Underst. Sci.* **16**, 299–322 (2007).
70. Stirling, A. Multi-criteria Mapping: mitigating the problems of environmental valuation. (1997).
71. Collingridge, D. *The Social Control of Technology*. (Open University Pres, 1980).
72. Ely, A., Van Zwanenberg, P. & Stirling, A. Broadening out and opening up technology assessment: Approaches to enhance international development, co-ordination and democratisation. *Res. Policy in press*, (2013).
73. Fiorino, D. J. Citizen Participation and Environmental Risk: a survey of institutional mechanisms. *Sci. Technol. Human Values* **15**, 226–243 (1990).
74. Fiorino, D. J. Environmental Policy As Learning : A New View of an Old Landscape. (1999).
75. Stirling, A., Simmons, P. & Spash, C. APPROACHES TO THE MAPPING OF VALUES A Review of Q-Methodology , Multi-Criteria Mapping and Attitudinal Scales by. **44**, (2003).
76. Stirling, A. & Mayer, S. A novel approach to the appraisal of technological risk: a multicriteria mapping study of a genetically modified crop. *Environ. Plan. C-Government Policy* **19**, 529–555 (2001).
77. Stirling, A. & Coburn, J. *Multicriteria Mapping Manual Version 1.0*. (2014).
78. Collingridge, D. *Technology in the Policy Process: controlling nuclear power*. (Frances Pinter, 1983).
79. Stephenson, W. *The Study of Behavior: Q technique and its methodology*. (University of Chicago Press, 1953).
80. McKeown, B. & Dan Thomas. *Q Methodology*. (Sage, 1988).
81. Cairns, R. & Stirling, A. "Maintaining Planetary Systems" or "Concentrating Global Power?" High Stakes in Contending Framings of Climate Geoengineering. *Glob. Environ. Chang.* (2014).
82. Addams, H. in *Soc. Discourse Environ. Policy An Appl. Q Methodol.* (Addams, H. & Proops, J.) (Edward Elgar, 2000).