

This is the accepted manuscript version of the contribution published as:

Schaafsma, M., Bartkowski, B., Lienhoop, N. (2018):

Guidance for deliberative monetary valuation studies

International Review of Environmental and Resource Economics **12** (2-3), 267 – 323

The publisher's version is available at:

<http://dx.doi.org/10.1561/101.00000103>

Guidance for Deliberative Monetary Valuation studies

Authors:

Marije Schaafsma^{a,*}, Bartosz Bartkowski^b, Nele Lienhoop^{b,c}

Affiliations:

^a University of Southampton, School of Geography and Environmental Science, University Road, Southampton, SO17 1BJ, United Kingdom

^b Helmholtz Centre for Environmental Research - UFZ, Department of Economics, Permoserstraße 15, 04318 Leipzig, Germany

^c Bochum University of Applied Sciences, Faculty of Economics, Lennershofstraße 140, 44801 Bochum, Germany

* Corresponding author: M.Schaafsma@soton.ac.uk

Abstract

There is growing demand for more pluralistic valuation approaches, for which Deliberative Monetary Valuation (DMV) is suitable. Guidance is needed for valid and reliable DMV application, as exist for conventional willingness to pay studies using stated preference methods. The purpose of this paper is to develop a set of minimal requirements for study design and reporting aimed at DMV practitioners, based upon the existing DMV literature as well as related social science literatures. The core contribution of our paper are the practical recommendations for DMV study design focusing on the deliberation process and elicitation format, the analysis of the deliberation and willingness to pay results, and validity. We summarise reporting requirements for reliability, before offering conclusions and suggestions for promising future research directions.

Keywords: environmental valuation; non-market valuation; stated preferences; survey; workshops

JEL codes: C93, H41, Q51, Q57

Acknowledgements

MS would like to thank the financial support from an ESPA Early Career Fellowship Grant (FELL-2014e104) with support from the Ecosystem Services for Poverty Alleviation (ESPA) programme. The ESPA programme is funded by the Department for International Development (DFID), the Economic and Social Research Council (ESRC) and the Natural Environment Research Council (NERC).

1 Introduction

In the past decades the importance of considering environmental goods and services in decision-making and policy assessment has become widely recognised. This has led to an increased interest in the economic value of such goods and services. Particularly with regard to ecosystem services, monetary values have gained attention both in the scientific and policy realm (MEA 2005, TEEB, 2010, European Commission 2013). At the same time, conventional valuation methods, particularly stated preference (SP) methods, have met with substantial criticism (Sagoff 1988; Sen 1995; Gómez-Baggethun and Ruiz-Pérez 2011; Hausman 2012; Parks and Gowdy 2013), which has led to the development of a new method type – Deliberative Monetary Valuation (DMV). For conventional valuation methods, guidelines and recommendations were recently published (Bishop and Boyle 2018; Johnston et al. 2017), aiming at standardisation and thus quality assurance and relevance of results. Such recommendations or guidelines for DMV are still missing.

In most general terms, DMV is a combination of SP valuation with deliberative (group discussion) elements. The term was coined by Spash (2007), but several earlier publications have come to be identified with DMV, both theoretical papers (Brown et al. 1995; Sagoff 1998; Niemeyer and Spash 2001; Kenyon et al. 2001; Wilson and Howarth 2002; Howarth and Wilson 2006) and applications (MacMillan et al. 2002; Christie et al. 2006; Álvarez-Farizo and Hanley 2006). In the closely related literatures on environmental decision making and environmental planning, deliberative approaches had been proposed and applied even earlier (e.g. Burgess et al. 1988a,b; Gregory 2000; Pelletier et al. 1999; Renn and Webler 1992). Recent years have witnessed a surge of new literature on DMV, including overview papers (Lo and Spash 2013; Bunse et al. 2015), theoretical and conceptual contributions (Bartkowski and Lienhoop 2018; Frame and O'Connor 2011; Howarth and Wilson 2006; Raymond et al. 2014) and many applications (e.g. Lienhoop and Völker 2016; Szabó 2011; Kenter et al. 2011; Vargas et al. 2016).

The term deliberation originates from *deliberare*, meaning to carefully consider or weigh well. Most DMV approaches draw to some extent upon the theory and practice of deliberative democracy and particularly Habermas' notion of the ideal speech situation (Habermas 1984, 1987; Rawls 1971; Dryzek 2000; Smith 2005). The ideal speech situation describes a safe and egalitarian setting, in which people seek common understanding and make joint decisions by reasoned argument, consensus and cooperation, without power asymmetries (Cuppen 2018; Mendelberg 2002). The reasons for combining SP valuation with deliberative elements generally fall into three categories: substantive, normative and instrumental (Chilvers 2009). In the substantive approach, to which most applications to date belong, the focus is on the formation of informed and stable preferences, learning and information and knowledge sharing to improve the rigour of science and decisions. As opposed to conventional SP surveys, deliberative approaches place emphasis on learning, education, knowledge and information exchange, as well as time for reflection to form preferences before the SP valuation exercise. The normative approach focuses on democratising economic valuation and on value pluralism (see also Bunse et al. 2015). Many conceptual studies and some more recent applications belong to this category. Here, increasing the democratic quality of the process itself is the aim, as the right thing to do. This is in line with Bulkeley and Mol's (2003) assessment that "increasingly, non-participatory forms of policy making are defined as illegitimate, ineffective and undemocratic, both by politicians and by stakeholders" (p. 144). The instrumental "moralising" motivation for DMV is to achieve particular governance goals through redesigning the value articulating institution, such as building trust, legitimacy, accountability, acceptance or conflict reduction

(Fish et al. 2011; Chilvers 2009; Degeling et al. 2015; Vatn 2009; Edwards et al. 2016). As normative and instrumental motivations for deliberation often go hand in hand and result in similar methodological approaches, we employ in the following the stylised distinction between “Valuation Workshops” (substantive approach) and “Value Juries” (normative and/or instrumental approach; see Section 2). This differentiation corresponds to a large extent to the distinction between deliberated preferences and deliberative democratic monetary valuation by Orchard-Webb et al. (2016), and to Lo and Spash (2013) who distinguish between “economising” and “moralising” approaches to DMV. We refer to the sub-distinction between normative and instrumental approaches where appropriate.

Support in political circles for more pluralism in environmental management by using deliberative valuation seems to increase (Albon et al. 2014; Pascual et al. 2017), in tandem with academic emphasis on social values and value pluralism (Special Issues in the journal *Ecosystem Services*, Kenter 2016; Jacobs et al. 2016). DMV has a number of claimed advantages over the (very narrow interpretation of) the homo oeconomicus model. Arguably, it allows for consideration of unfamiliar, complex environmental goods, makes possible the inclusion of non-utilitarian values, takes into account the social, process nature of preference formation and overcomes motivational crowding-out, contributes to increasing the democratic quality of environmental policy and is conducive of revealing social values (Niemeyer and Spash 2001; Lo and Spash 2013; Raymond et al. 2014).

DMV proponents claim that DMV studies lead to “better” (Kenter et al. 2016b), “more meaningful” (Gómez-Baggethun and Martín-López 2015), “better considered” (Lienhoop and MacMillan 2007a) and “potentially more robust” (Kenter 2016) outcomes, as well as lower non-response rates to the willingness to pay (WTP) question and respondents who are “more motivated” to bear the costs of participating in the complex valuation task (Lienhoop and MacMillan 2007a; Szabó 2011). However, the acclaimed substantive, normative and instrumental advantages of deliberation, and participative methods more generally, are not strongly supported by empirical evidence (Chilvers 2009; Chilvers and Kearnes 2017; Mendelberg 2002; Arias-Maldonado 2007; Abels 2007). Existing applied DMV studies report surprisingly little about the motivation for deliberation, the form and duration of the deliberative activities, or the type and content of the information that is exchanged. Therefore, it is difficult to empirically verify the claims made in favour of DMV in the theoretical literature, or indeed to replicate the research.

The motivation for this paper is that to respond to the growing demand for more pluralistic valuation approaches, and DMV in particular, guidance is needed in the form of recommendations for valid and reliable DMV application, like those for conventional WTP studies using standard SP methods (Johnston et al. 2017; Arrow et al. 1993). The purpose of this paper is therefore to develop a set of minimal requirements for study design and reporting for DMV practitioners.

Our recommendations are based on the existing DMV literature, including a wide range of deliberative methods, from “pragmatic” (substantive) to normative and instrumental deliberative approaches (Raymond et al. 2014), and both deliberated preferences and deliberative democratic monetary valuation approaches (Kenter et al. 2016b). But the small body of literature on DMV applications means that we also draw upon related social science literature on participation, deliberative democracy, political psychology, collective decision making, group psychology, qualitative social research, as well as existing validity and reliability guidelines for “standard” SP literature and more general economic literature.

In the next section, we present the different types of DMV and their theoretical foundations. This provides the background for our first set of recommendations on the first step of DMV studies: deciding on the purpose and theoretical underpinning for deliberation (Section 3.1). The core contribution of our paper is the practical recommendations that follow. We address DMV study design focusing on the deliberation process and elicitation format (Section 3.2), the analysis of both the deliberative and WTP results (Section 3.3), and relevant (external) validity criteria (Section 3.4). We summarise reporting requirements for reliability (Section 3.5), before offering conclusions and suggestions for promising future research directions (Section 4).

2 DMV types and theoretical foundations

The term DMV is used to label approaches that combine monetary valuation (usually SP) from economics with deliberative methods that were inspired by deliberative democratic theory and originally developed to enhance participation in political decisions and improve their democratic quality (Kuyper 2017). Since DMV is of interest to scholars from a range of disciplines, including economists (Brown et al. 1995; Spash 2007; Wilson and Howarth 2002), decision theorists (Gregory et al. 2005), and philosophers (Sagoff 1998; O'Neill 2001), there are several variations of DMV. These differ in terms of practical implementation, conceptualization and underlying theoretical assumptions (Bunse et al. 2015). In this section, we disentangle the different types of DMV as well as their respective theoretical foundations.

Overall, DMV comprises two main categories: Value Juries and Valuation Workshops (also called Market Stall). Note, however, that these are not standardised terms and many alternative names can be found in the literature. The main differences between these two categories are value type elicited (social versus individual choices and WTP) and their main theoretical underpinning (deliberative democratic theory versus neoclassical economic theory). Value Juries fall into the democratising line of work (normative and instrumental purpose), whereas Valuation Workshops lean towards economic rationality assumptions (substantive purpose).

Value Juries have their roots in Citizens' Juries. Invented by Ned Crosby and Peter Dienel in the 1970s, a Citizens' Jury in its classic form consists of 12–25 citizens, chosen in a stratified random selection process to represent a diversity of socio-demographic criteria (e.g. gender, age, ethnicity). These representatives deliberate over a community issue over a couple of days. To do so, they are provided with information on the topic, hear and question witnesses, discuss different aspects of the issue and commonly arrive on a preferred course of action (Coote and Lenhaglan 1997; Bunse et al. 2015). When the jury's aim is to generate a common (often labeled social, group-based or collective) WTP to obtain or avoid a certain environmental change, the approach is called Value Jury (Brown et al. 1995; Robinson et al. 2008).

The communicative rationality assumptions underlying Value Juries are drawn from the theory of deliberative democracy (Smith and Wales 1999). Communicative rationality refers to Habermas' theory of communicative action (Habermas 1984), in which an action is rational if it is the result of an open-minded exchange of arguments and the 'unforced force of the better argument' (Habermas 1996). Deliberative democratic theory views the individual as a reflexive citizen who holds socially constructed and non-myopic preferences (Ward et al. 2003; Vatn 2009). Deliberative democracy stands for a procedural approach and takes public

involvement as the core of legitimate political decision making. Essential elements are freedom and equality of participation, the importance of reasoned arguments, and consensus (Cohen 1997). In Value Juries participants are regarded as reflexive citizens who represent and are deeply embedded in society, so that in the ideal speech situation statistical representation is not required.

Deliberative democracy is closely interlinked with Habermas' theory of communicative action (Habermas 1984; 1987). It suggests that a deliberative process enables participants to construct their preferences. In this process participants first consider their individual interests and, through deliberation, transcend these and adopt other-regarding perspectives. Eventually participants achieve a mutual understanding by means of exchange of argument and reach consensus about the decision task. Important prerequisites for deliberation include that citizens (i) are educated and informed about the decision issue, (ii) have the opportunity to extensively reflect on their preferences, (iii) are encouraged to ask questions, and (iv) are spurred to express arguments for one outcome over another (Fishkin 1991).

The Valuation Workshop approach was developed by MacMillan et al. (2002). The approach involves a group meeting, usually lasting approximately two hours with up to 12 participants. During the meeting a facilitator conveys relevant information on the environmental change to be valued, the hypothetical market and the payment vehicle, using different communication forms (verbal, information folders, posters). Participants are then encouraged to ask questions and discuss the issue with each other. At the end of the meeting participants express their preferences individually and anonymously, either through a Contingent Valuation (CV) or Choice Experiment (CE) type of WTP question. To give respondents time to reflect and re-adjust their preferences, Market Stalls often involve a second meeting or a follow-up phone call after about one week (MacMillan et al. 2002; Lienhoop et al. 2015).

The rationality assumptions of Valuation Workshops draw from both economic theory and deliberative democratic theory. Standard SP methods start from the model of homo oeconomicus (fully informed, self-interested, with predefined preferences), consequentialism/utilitarianism, social welfare as additive aggregate of individual welfare estimates and, thus, require statistical representation (Spash 2007). While mostly leaning towards economics, Valuation Workshops strive to assimilate the insights from deliberative democracy (i.e. the prerequisites for good deliberation mentioned above), to support the preference formation processes of participants. Thus, they include elements such as discussion with other members of the general public as well as time to think. The main argument for DMV in this line of work is that individuals need to be reminded of salient issues and that group interaction enables this. Nonetheless, Valuation Workshops distance themselves from deliberative democratic theory in that they stick to eliciting and aggregating individual preferences. Meanwhile, Value Juries and related approaches target non-individual values (shared, social or transcendental values) more directly and explicitly (Kenter, 2016).

Most DMV literature invokes Habermasian theory of communicative action (e.g. Orchard-Webb et al. 2016; Raymond et al. 2014; Lo 2013), but a few theoretical contributions have explicitly invoked Rawlsian theory of justice (e.g. Brown et al. 1995; Wilson and Howarth 2002). It is not clear, however, what the consequences are for DMV depending on which of these theoretical foundations (Habermas vs. Rawls) is chosen; this question has not been explicitly addressed in the DMV literature (see also next section). Recently, Amartya Sen's approach to rationality (Sen 2009), which is not unlike Habermas', but less demanding than the "ideal speech situation" and more sceptical towards consensus, has been proposed as a

possible theoretical reconciliation of deliberative principles and economic theory (Bartkowski and Lienhoop 2018).

3 Recommendations for conducting DMV studies

In this section, we derive recommendations to guide practitioners who would like to conduct DMV studies. We point towards relevant sources and evidence from diverse research fields.

3.1 Purpose and chosen theoretical foundation

The most important question for any DMV study is: why is there a need for deliberation? In other words: what is deliberation supposed to achieve? A clear answer to this question also implies the choice of a theoretical foundation for the study (Lienhoop et al. 2015) and determines the suitable characteristics of the process and the expected outcomes (Gastil et al. 2017). It is a specifically challenging issue to combine deliberative democratic theory with economic valuation without sacrificing theoretical consistency. The purpose of the study and, specifically, the reason for including deliberation determine how ambitious and theoretically consistent the deliberative elements must be.

Deliberation serves different purposes in the two different DMV types we described in Section 2. Valuation Workshops, common in applied studies, aim to improve the informational basis of preferences by facilitating preference formation in a group setting. These applications motivate the choice for deliberation by the complexity of the valuation scenario or good, and the scientific uncertainty of provision of the good. The purpose of Value Juries is the “democratisation” of the overall valuation process and its results. Here, motivations can relate to substantive (participation improves quality of science or decisions, through inclusion of multiple knowledges and meanings), instrumental (deliberation as a means to an end, e.g. legitimacy) or normative (deliberation as an end in itself, the right process in democracy) goals (Chilvers 2009). In the context of environmental valuation, moral uncertainty and value conflicts are prominent substantive reasons (Lo and Spash 2013). Often, deliberative and participatory studies do not state explicitly whether the purpose is to educate people, inform policy or work with stakeholders (Degeling et al. 2015, Fish et al. 2011).

Both choices, i.e. on the purpose and theoretical foundation, translate into minimum quality requirements for deliberation. If the purpose of deliberation is mainly preference formation, the demands are relatively low and more technical. Introduction of the issue by a facilitator and moderated group discussion are central. A DMV workshop can be considered deliberative if participants can speak and listen to others’ opinions, can ask questions to clarify difficult issues and can (get time to) reflect upon the issue for themselves. Whether this has taken place in practice can be checked by follow-up questions regarding self-reported certainty of WTP or motives underlying WTP, by econometric testing procedures to identify impacts of learning on parameter estimates or variance across choice tasks or by comparing WTP before and after deliberation (Schaefer et al. 2013; Lienhoop and Völker 2016; Brouwer et al. 2010).

When democratisation is the goal, the demands are higher. It is relevant here which specific theoretical background is invoked (e.g. Habermasian, Rawlsian, Senian). For instance, if the chosen theoretical foundation is Habermasian discourse ethics, the central question is: how to approximate an ideal speech situation to achieve communicative rationality and discourse (cf.

Himmelroos 2017)? In other words, how do you ensure the democratic quality of the outcome and process (Edwards et al. 2016), including non-coercion and connection of specific interests to generic principles (Dryzek and Hendriks 2012)? This is particularly challenging as Habermasian deliberation is never finished or final (Habermas 1996), while a DMV study is inherently limited in time. Furthermore, in the ideal speech situation validity claims are evaluated in the process of deliberation. To be considered “ethical”, the discourse process must be open, respectful, and accept and assimilate different voices and views (O’Hara 1996). Participants must demonstrate constructive group behaviour for a communicatively rational debate. In practice, however, individuals in groups may display counterproductive behaviour or strategic rationality and exploit their power or identity (Mendelberg 2002). Issues such as power imbalances, implicit hierarchies, knowledge asymmetries, (un)willingness to engage and accept others’ views are even more relevant when the aim is to seek consensus, compared to economising DMV studies where individuals maintain more autonomy and independence. This translates into specific requirements for moderation, duration and frequency of deliberation activities as well as sampling (Section 3.2.1).

Other types of deliberative rationality include Senian and Rawlsian reasonableness. The former recognises rational pluralism and has agency as a core concept. Sen’s reasoned argument aims for selecting one or more better choices, but not one optimal choice. Rawls’ concept of public reason is based on the concept of reasonable pluralism and involves a balance or negotiation between individual pursuits of ends and willingness to propose fair terms (justify their claims in a way acceptable to others), based on shared values that are acceptable to all free and equal members of society, despite alternative views (Kuyper 2017). It does not require agreement on best reasons, only on a shared understanding of what is reasonable. Neither of these approaches has been operationalised in applied DMV studies, but Kuyper (2017) argues that Rawlsian approaches may be more idealised than Habermas. Differences between these different interpretations of deliberative democratic theory have not been much discussed or explored empirically in the DMV literature.

A further requirement for DMV studies, related to the purpose of the study, is to report on the positionality of the researcher(s), especially when normative or substantive purposes are at stake. Some existing DMV studies on public goods seem to express a clear instrumental political objective: to strengthen arguments for sustainable development, in particular by calling upon (otherwise latent) transcendental, shared values (Kenter et al. 2016a). Value pluralism requires that indeed all values or worldviews have a place in the debate; the chosen theoretical foundation will then determine whether these values should somehow be reconciled in a consensual choice or if agreement to disagree is allowed. This is particularly evident for international, cross-cultural studies (Christie et al. 2012; Kenter et al. 2011), where researchers and participants may hold very different worldviews and perceptions about human-nature relationships.

Our key recommendation is to clearly state the purpose (substantive, normative, instrumental) of deliberation and the underlying motivations. The choice of theoretical foundation, type of rationality aimed for and all subsequent design choices should be consistent with the purpose of deliberation.

3.2 Design

3.2.1 Deliberation process

The main difference between DMV studies and conventional SP studies is the deliberative activity. The design of the deliberative process includes choices about sampling, logistics, information provision, moderation, topics of deliberation, and democratisation. These design factors influence the outcome of the deliberative process (Kenter et al. 2016a).

Sampling

The first question is sampling: who is going to participate in the workshops and how should participants be chosen? The purpose and theoretical foundation (see 3.1) often determine the type of representation required. In Valuation Workshops, some level of statistical representation should be strived for, but may be difficult because of the high costs of conducting a DMV study.

Practitioners should carefully consider whether strong statistical results (i.e. small confidence intervals, models with high explanatory power, hypothesis testing through split-sample designs) are needed for the academic or policy context of the study. Such objectives require larger samples and potentially individual choices. The WTP elicitation format should also be designed with the potential sample in mind. For example, if sample sizes and the number of choices that can be made are small, a Choice Experiment can only include a small number of attributes. When using Choice Experiments in Valuation Workshops, researchers can pilot the study and use preliminary outcomes as priors for experimental design to reduce the sample size requirements of the main study.

Instead, stratified sampling which is suited for small sample sizes may be chosen and reduce costs (Mitchell and Carson 1989). Stratified sampling involves the division of the population along characteristics of importance to the researcher (e.g. gender, social class, education level, religion). Following this, the population is randomly sampled within each category (Groves et al. 2009).

In Value Juries, the deliberative democratic theoretical basis translates into a requirement for political or discursive representation (Dryzek and Niemeyer 2008; Raymond et al. 2014). But full political representativeness may be hard to achieve: invited participation often follows existing institutional structures but may fail to involve parties that tend to instigate social conflict (Cuppen 2018). Moreover, remote stakeholders, e.g. of global environmental benefits are often excluded in sampling strategies (Sen 2009). Thereby standard approaches to politically representative sampling can be insufficient for inclusive deliberative normative appraisals.

Stratified random sampling is usually applied for its advantages of: (i) inclusion of minorities (because of stratification) and (ii) prevention of co-option by vested interests (because of randomness) (Lafont 2017). A risk of stratification is that participants may view themselves as representatives of their respective societal strata, which may be undesirable for some purposes, e.g. when it is preferable that participants view themselves simply as citizens of a given polity (Smith and Wales 2000).

The next question related to sampling is: who should participate in the study and why? In the deliberative concept of inclusive governance, all potentially affected parties should be represented, but DMV proponents also refer to other concepts of inclusive governance as laid

out by Renn and Schweizer (2009), such as the potential to include representatives of different knowledges and rationalities, non-interested parties and marginalised groups. The identification of groups or strata can be based on power (over resources, decisions, contacts), competence (knowledge, expertise) or those who are impacted (Fish et al. 2011). Stakeholder groups can be identified through stakeholder mapping or interviewing experts, who can also advise on the inclusion of marginalised groups. The role of the publics involved in the study must be clear (Degeling et al. 2015): is it to provide expertise, experience and knowledge (substantial), to provide citizen legitimacy to decisions (instrumental), or to generate engagement (normative)?

Self-selection bias and non-participation are problematic for DMV studies, both for political and statistical representation. If people choose not to participate because they lack interest or do not perceive themselves to be affected, the opportunity of including an external perspective of a person with common sense but no immediate stake (“anthropological inclusive governance” in Renn and Schweizer 2009, also Sen 2009) is lost, although deliberative inclusive governance criteria still hold. If, however, certain groups decide not to participate upon invitation, for example because the outcome appears to be a loss to them, this can de-legitimise the process (French and Laver 2009). Some deliberative studies have included uninterested participants (who acknowledge an issue but have no desire to engage with it) (e.g. Lorenzoni and Hulme 2009), but there is no DMV study on the effect of including uninterested participants on the preferences of the uninterested, other participants or overall outcomes.

Group composition, representation within a group and group size can all have impacts on group behaviour (Mendelberg 2002). In general, groups of 8–12 participants work well for deliberation, both for Valuation Workshops and Value Juries, as this group size is manageable and can incorporate multiple viewpoints (see *Group effects*). Over-recruitment is often necessary to address no-shows. If equal standing of all participants cannot be achieved, it may be useful to organise separate deliberative activities for marginalised or seldom-heard groups. Remuneration of participants should be considered especially if the costs of attendance are prohibitively high for (some) groups or to motivate people that would otherwise not choose to attend a group exercise (i.e. to counter self-selection). It should be noted that separating groups can also have negative consequences later; cooperation within small groups towards a common choice may be irreconcilable with between-group cooperation (Mendelberg 2002).

Logistics

A seemingly trivial issue concerns logistics – where is the workshop to take place, how should the room(s) be prepared beforehand etc. Yet, different room layouts facilitate different group dynamics (Bellamy et al. 2017). Psychological research suggests that the physical setting of discussions, e.g. the position of the facilitator in the room or the seating arrangement, influence participants’ behaviour (Greenberg 1976). Locations and layout elements considered by some participants as non-neutral can have negative effects on participation, for instance, when the workshop takes place at the location of an involved stakeholder or in the presence of religious symbols. As DMV studies usually strive for equality and inclusivity, practitioners should bear the potential effect of the workshop location in mind. Settings should be well suited not only for facilitating inclusive discussions and answering questionnaires (i.e. chairs and tables arranged in a circle or u-shape), but also for the effective and efficient use of envisioned moderation techniques and information

materials (e.g. on-screen presentations, posters etc.). The required access to (external) information should also be considered in location choice.

Information provision

The content and mode of information provision seems a priori an even more crucial element of DMV design than in standard SP studies (e.g. Richards et al. 2017; Tisdell et al. 2008), and includes aspects of quality, access and medium. Ideally, the information provided is transparent, with clear references to sources and with input from experts, stakeholders or citizens, providing a range of different – and where appropriate opposing – perspectives. It should avoid scientific language and be translated into laypeople's terms to ensure comprehensibility. The materials must provide information about both pros and cons associated with the environmental change under investigation and remain as neutral on further decisions as possible. Information provision should be informed by previous interactions with stakeholders, e.g. via exploratory focus groups and in pre-tests.

Practitioners should ensure equal access to this information, especially where there are differences across participants in literacy, language, or hearing and visual abilities. Access to additional information resources (experts, witnesses, specialists, etc.), when possible in the chosen DMV approach, may be crucial towards achieving the purpose of deliberation (Fish et al. 2011).

Good quality information can have significant positive influence on deliberative quality (Gastil et al. 2017). Individual SP studies try to fully control the information provided and delivered, and many experimental economics studies use digital platforms in which all information is controlled and all interactions are recorded. However, in DMV that control may be much lower or absent. In principle, the richness and variety of information that arises from different deliberations or information sources prior to WTP elicitation is the key substantive purpose of DMV, and so this control may not be warranted.

But beyond issues of comparison across groups, there are potential pitfalls that affect the quality of information shared through deliberation. Firstly, the knowledge shared or gathered by participants may not be true. If quality control of information is implicit or assumed and experts or participants can deliberately provide unreliable information, there is a risk that power asymmetries (information asymmetries) are exploited. In addition, in Valuation Workshops, there is little control over which folders or information leaflets the participants pick up and read, so which information is used is unclear. When the purpose of deliberation is to improve the informational basis of policy decisions through reasoned and open-minded argument (substantive), the key question is whether the deliberative process leads to consideration of that information by participants when forming their preferences (Peter 2013). One option to approximate learning would be to test knowledge before and after participation (e.g. LaRiviere et al. 2014); other DMV studies have asked respondents after participation whether they learned anything that changed their opinion (Lienhoop and MacMillan 2007b).

Secondly, in Value Juries, external experts or witnesses may be called upon by participants, but their testimonies are beyond control of the researchers. Expertise, both held by “lay” participants and “expert witnesses” (including, in DMV contexts, the researchers), which is to a large extent co-produced during the deliberative process, can have two problematic consequences of “expertise discourse”: false obviousness of solutions suggested by “experts” and the creation of hierarchies between those (seemingly) in possession of expertise and those without (Sprain and Rainig 2018). Information provided by an outside specialist may

not be perceived accurate, legitimate or reliable (Davies and Burgess 2004; Mansbridge et al. 2012). The accountability of the expert must be clear to the DMV participants.

Thirdly, dominant experts may frame a public-choice problem in their favour. Empirical research has shown that deliberation can only reduce such framing effects to some extent, especially when groups are heterogeneous, but it cannot eliminate them (Druckman 2004).

Finally, there is the question about which medium is chosen to convey the information. In DMV studies, an additional layer of complexity and potential pitfalls is present, since the information is usually not provided as a standardised text at the beginning of a questionnaire, but often presented by a facilitator so that verbal and especially non-verbal factors come into play. For instance, the well-known halo effect may lead to participants “trusting” or “believing” the facilitator simply because they are rhetorically gifted or nice-looking (Nisbett and Wilson 1977). Whilst interviewer-bias plays a role in individual face-to-face surveys too, the halo effect may be much stronger during prolonged interaction of the deliberative process. Sandorf et al. (2016) tried to avoid this bias by using video to convey the information in the hypothetical scenario.

Moderation

Moderation and facilitation of group discussions, in particular on controversial topics, requires a large amount of skill and experience. Sandfort and Quick’s (2017) concept of deliberative technology emphasises the importance of “methodological techniques, material objects, and conceptual frameworks” in deliberative contexts. What the specific requirements of “good facilitation” are, however, is not clear:

“While it is broadly acknowledged that ‘good’ facilitation empowers stakeholders, manages power imbalances and prevents the dominance of professional voices, there is limited understanding of what constitutes good facilitation in practical terms.”
(Crompton et al. 2017, pp. 16-17)

Guidelines for focus groups and other participatory methods of qualitative social science provide quality standards for moderation and facilitation (Fern 2001; Gibbs 1997; Krueger and Casey 2014). A body of literature in sociology and education sciences has looked at effective group work and cooperation, which practitioners may want to consult, also for its practical guidance (Johnson and Johnson 1994). Although it seems to be common practice that DMV workshops are moderated by the researcher rather than by an external facilitator, hiring an experienced facilitator whenever feasible and specifically if researchers do not have the necessary skills is important for the quality of the results.

In general, the moderator should aim to create a friendly, rather informal atmosphere where participants feel safe and free to express their views. Where preference formation through information transmission is the main purpose of deliberation, the moderator will need to focus on understanding of the good under valuation amongst all participants. Again, deliberative activities with the purpose of democratisation have farther reaching requirements. Achieving unbiased, free and equal participation of different groups of participants requires careful attention to how some of those groups participate. If a group is systematically “silent”, this requires action by the moderator. Carter et al. (2018) suggest that sufficient time for discussions can help alleviate the problem of systematic underrepresentation; restraining particularly dominant participants and encouraging “shy” ones to express their opinions may also help.

Where the (normative) objective of deliberation is to empower marginalised groups, inequality in communicative competence and other resources need to be considered (Clifford 2012). Moderation in mixed groups could focus on empowering those groups to actively engage in discussion and assist in articulation of contributions. Alternatively, other forms of deliberation that rely less on rational argumentation may be needed to improve inclusion (Vargas et al. 2017).

Consensus seeking will require a different approach, where guidance towards a shared understanding and vision on solutions may be a key component of moderation to achieve the substantive purpose. Brown et al. (1995) point out that achieving consensus is a process that the facilitator must understand and guide, whilst remaining impartial throughout the discussion. Tools and models to guide the process are available, such as Consensus Voting, Blocking, Quaker-based model, and consensus-based decision-making.¹ In cases of conflict or lack of consensus, moderators may have to allow agree-to-disagree outcomes, even if this does not fulfil the ideal of consensus (e.g. Murphy et al. 2017). At the same time, in cultures that value consensus seeking highly, deliberation may not reveal contradictory thoughts, values and preferences. If consensus is rushed, the moderator may intervene to slow the conversation. Most important for informing subsequent policy or decision making is to ensure that all sides have had the opportunity to express their views and see those recorded and communicated with final decision makers.

To address the problems involved in expertise discourse discussed above, facilitators can use different techniques, which resemble methods to deal with dominant participants. Sprain and Reinig (2018: 367) propose that moderators should (i) reframe the discussion when presumably “obvious” solutions are offered, so as to prevent “narrowly focusing on a single possible solution”; (ii) engage in “passionate impartiality”, trying to restore power balance and equity within the group, e.g. by “reframing an expert comment as contestable by asking for reactions or counter arguments.” Here, the researcher can play a role in the deliberation as an impartial expert, who can intervene when a participant, whether intentionally or not, makes distorted statements or claims that may sway the deliberation and are not corrected by other participants. Such interventions must be done in a polite, respectful manner, be restricted to factual information, and be based on detailed and referenced knowledge of the case. An example from our experience where an intervention was suitable was when a participant quoted a study that did not exist – the statement sounded reliable and the participants were unable to correct it.

Group effects

Of concern in group-based discussions are potential negative effects of group dynamics, including communication and power effects. Group systems theory points at the interactions between all members of a group, the structure of groups and formation of coalitions or sub-groups who share the same values, interests and needs, and the patterns of communication (Kottler and Englar-Carlson 2014). Participants can take up different roles during deliberation, some of which are constructive, while others are destructive and steer away from deeper deliberation (Kottler and Englar-Carlson 2014). Participants may be aggressive and try to control others, talk so much that no one else can contribute, or frame or steer the discussion to their advantage. Others may suppress any form of emotion or conflict between others, agree with everyone, or just disengage from the discussion. Functional roles would

¹ See www.consensusdecisionmaking.org and www.seedsforchange.org.uk/consensus#proc (retrieved 22 May 2018).

include members who mediate conflicts, build trust and consensus, motivate others into action, seek information or keep people on task and ensures progress.

Deliberation is not naïve to power issues or group effects, and there is debate to what extent deliberation can deal with polarisation and other group effects. Group polarisation occurs when deliberation moves the group opinion to a more extreme point (in the direction of the original inclination) than indicated by individual non-deliberated opinions. This tends to happen more in homogenous groups and is driven by social comparison and persuasion of arguments delivered with confidence (Sunstein 2002, 2005; Sunstein and Hastie 2015). Proponents of deliberation argue, however, that polarisation effects do not occur in heterogeneous groups which follow a structured deliberative process (Curato et al. 2017). In mixed groups that hold various viewpoint rather than a uniform one, participants do not as easily conform to the dominant or majority position in the group (Brodbeck et al. 2007). Völker and Lienhoop (2016) hence suggest mitigating such group dynamics *ex ante* by composing each group so that it contains participants with heterogeneous preferences (i.e. include participants that advocate and oppose the policy issue to be discussed as opposed to homogeneous group with participants who are in favour of the policy issue).

Other problems identified in deliberating groups include the amplification, rather than correction, of individual errors in judgement; cascade effects, where group members follow others; and emphasis on common, rather than critical but not widely known, information (Sunstein and Hastie 2015). The presence of social norms may reward conformity and reduce the need for reason and reflection in preferences expressed in groups, thereby masking contradictory opinions. Individuals may hold on to their initial position and remain entrenched (Watkins et al. 2013). These group behaviour problems undermine the preconditions necessary for successful deliberation, i.e. equality of participants and willingness to reason, reflect and adjust preferences.

Group dynamics are often subtle and therefore difficult to detect and prevent. However, researchers can alleviate group dynamics by allocating respondents to groups so that socio-demography and views and attitudes towards the topic of deliberation are heterogeneous. Such information about participants can be gathered with a short questionnaire during the recruitment process (Völker and Lienhoop 2016). Excellent facilitators can counter dynamics stemming from dominant versus shy respondents by appropriately steering the discussion (see *Moderation*). Finally, it may be possible to mediate group polarisation by choosing a specific voting mechanism. The results of Vargas and Díaz (2017) suggest that non-consensual rules, where respondents express their WTP individually and anonymously, reduce group-polarisation effects while maintaining the benefits of deliberation.

Topics of deliberative activities

As discussed in Section 3.1, typical motivations for deliberation relate to complexity, uncertainty and value conflicts. Decisions for which deliberation is pertinent involve multidimensional impacts, fairness of outcomes and processes, including the rights and well-being of future generations, uncertainty and risk, mistrust and urgency (Stern, 2005). When discussing the most desirable development trajectory, people may differ widely in their preferred outcome, means to that end, pathway, as well as motivation, and it may be important to explore these differences. But for consensus, it can be unnecessary for participants to agree on all aspects of a choice (Ben-Arieh and Easton 2007; Sen 2009).

An important design choice involves the determination of question and discussion topics. Existing studies tend to discuss, at the very least, the interpretation of the valuation scenario

including the good under valuation. The phrasing of these questions must be as neutral as possible.

Specific topics can either be set *a priori* by the researcher or emerge during the discussion. *A priori* framing of information and specification of debate topics or questions dictate the discussion agenda and may introduce issues that the participants themselves had not considered and may not have deemed relevant. Confronting participants with “outside” perspectives and arguments is, in fact, a frequently voiced goal of deliberation (Bartkowski and Lienhoop 2018). On the one hand it is deemed important to discuss the underlying values of decision information (e.g. Stern 2005), on the other such questions to stimulate discussions frame the debate. Recent DMV studies have included sets of questions about moral values, which participants complete at individual level, before further deliberation and WTP are expressed (Orchard-Webb et al. 2016). Like leading questions, this may manipulate outcomes: value frames focus participants discussions on one value dimension but may reduce the depth and quality of the deliberation (Brewer and Gross 2005).

Open or hybrid formats are also possible, which are not as structured and focused and give participants the chance to discuss what they want, which may reveal topics (e.g. well-being aspects of the environmental good) that researchers or policy makers were unaware of (Wright and Street 2007). In hybrid formats, participants (in)form the agenda prior to deliberation. The framing of issues emerging from the group itself during deliberation can also be problematic, if these frames become dominant, polarising or group-based (Calvert and Warren 2014). Such framings reduce communicative freedom, undermine mutual respect and rely on prejudice, but they can be remedied by introducing alternative frames by the facilitator (Sunstein 2002).

Duration

Ideally, a Habermasian process is not restricted by time. DMV studies, however, usually are. This raises the question whether a workshop lasting one and a half or two hours, as common e.g. in Valuation Workshops, will suffice to ensure Habermas’ communicative rationality (or Rawlsian or Senian reasonableness). It can be expected that the elicited decisions or values in such a context are not “final”, but rather contingent and may change in the future. Guidelines for focus groups will not be sufficient; rather, the extensive literature on the practice of deliberative democracy such as citizens’ juries, deliberative polls etc. should be consulted (e.g. Smith and Wales 2000; Coote and Lenaghan 1997, Stagl 2007). In general, the length of a “democratising” workshop should depend on when its purpose (e.g. reaching consensus on a collective WTP or achieving a democratic process) is achieved. Not only duration, but also the quality or intensity of the social interaction and deliberation determine quality (Kenter et al. 2016a, see also 3.3.1).

Democracy and transparency of the design

In all topics above, it is assumed that the researcher decides on the design of the deliberative process, or perhaps co-designs it with the stakeholders. The power to make decisions on who is involved, where deliberation takes place, what information is provided or shared, who moderates and how, and which topics are deliberated, can be decisive in the outcome of the DMV study. In general, for the purposes of research, the researcher may want to fix certain design characteristics to test hypotheses. However, for “democratising” reasons, a more flexible, adaptive study design may be more fit for purpose, for example, in the flexible inclusion of stakeholders (Metzger et al. 2017). Respect and inclusion are important ethical and democratic requirements.

Lancsar and Swait (2014) argue that the decision-making context should reflect the influences and exogenous constraints that decision makers would face in reality. This realistic choice context is what they refer to as main question framing in external validity, and what Kenter et al. (2016a) call the institutional context. In the same way that SP studies aim for incentive-compatible designs (e.g. Czajkowski et al. 2017, Zawojka and Czajkowski 2017), for deliberative processes to be demand-revealing participants should believe that their engagement will make a difference, i.e. is consequential, whether that is by informing policy makers or decisions or by their own learning of the subject and its policy context. Mentioning the institution that may use the results can moderate the incentive compatibility (e.g. Vargas and Diaz 2017, Oehlmann and Meyerhoff 2017). Both for participants and non-participants, the deliberative process must be transparent, so that it is clear how and when the results will be used, and what the aim and objectives, as well as the policy and participatory context are.

Moderators should clarify what participation involves and which risks or distress participants may experience; this follows standard ethical guidelines for data collection of prior and informed consent. Most of the ethical criteria for DMV are the same as for individual SP studies. In addition, it is important to point out to participants that the information they share within the group discussion will thus be known to all participants. This allows participants to make an informed decision about what they will share with the group when their safety or privacy is at stake. Consent must also be obtained for video or audio recording. It is good practice to inform participants about the results of the analysis once ready.

The deliberative democracy ideals of inclusivity and value pluralism rely on participants' ability to engage in rational debate. In case those abilities are limited, practitioners should consider not (only) using group-based discussions or Value Juries, but (also) assessing individually expressed preferences (votes, rankings, etc), or other types of expression including through arts-led interpretive activities (Edwards et al. 2016).

3.2.2 Elicitation of WTP: aggregation rule and social welfare function

The elicitation format that defines preference expressions (WTP) is a distinguishing criterion for DMV methods and therefore involves a crucial decision that affects the interpretation of DMV results. A first important question in this respect concerns the aggregation rule. In practice, the most common choice in Valuation Workshop approaches is additive aggregation of individual WTP, in line with conventional SP methods, which is heavily debated.

Value Juries have based the elicitation of WTP on group or collective choices, usually to be arrived at consensually (e.g. Kenter et al. 2011; 2016c; Kenter 2016; Álvarez-Farizo et al. 2007; 2009; Álvarez-Farizo and Hanley 2006; Ito et al. 2009). Forced consensus has been criticised repeatedly, in different contexts and for different reasons (Arias-Maldonado 2007; Bartkowski and Lienhoop 2018). Dryzek (2000) refers to problems of power and coercion, and calls forced consensus unnecessary, unattainable and undesirable; furthermore, he refers to “deep moral disagreement” and “deep division” as possible obstacles to consensus that may necessitate other decision mechanisms, such as negotiations or voting (Dryzek 2013; see also Habermas 1996). Where individual choices or votes can be expressed, the effect of deliberation on the level of agreement can be analysed (List et al. 2012).

A second important design issue especially for collective WTP elicitation is whether the aggregation rule (how individual preferences are combined into social preferences, e.g. majoritarian vs. consensus) is prescribed or left open. The former option has the advantage of consistency across groups; the latter ensures that agreement has not been “forced” upon participants or that attempts to achieve one do not end up in stalemate. If set *a priori*, the

decision regarding the preference aggregation rule should be informed by the goals and theoretical foundation chosen for the study (see Sections 2 and 3.1). It would also be possible for participants to decide on the preferred aggregation rule, and because of the normative content this question would lend itself well to a deliberative process. In their consensus-oriented deliberative valuation study, Murphy et al. (2017) allowed in one group for an agreement to disagree outcome. Economic valuation is based on the principle of consumer sovereignty, which assumes unquestioning elicitation of individual preferences. This principle is in stark contrast with the fact that the *aggregation rule* applied to these preferences is imposed by the analyst - even though people also have preferences over aggregation rules and unweighted additive aggregation of individual preferences may not be widely preferred in practice (Stefan Baumgärtner, personal communication). Recognition of such preferences implies, however, the problem of infinite regress.

For both individual and group elicitation formats, different aggregation rules are available. For instance, Schläpfer (2016) proposed a valuation methodology that involves voting, i.e. a majority-based aggregation of individual preferences. Murphy et al. (2017) showed that aggregation rules different from simple additive aggregation (i.e. plurality, Borda, Kemeny, Hare, Copeland rule) help to approximate consensual collective preferences better. For collective WTP, Ito et al. (2009) found differences between consensus-based and majority-based rules, showing that consensus leads to less pronounced differences between individual preferences and collective choice.

A closely related issue is the choice of a social welfare function that combines individual welfare estimates into a social welfare estimate using aggregation, with variations such as utilitarian or Rawlsian functions. In practice, most Valuation Workshop studies use utilitarian functions in which individuals are assumed to have the same marginal utility of income. A first critique on aggregating individual preferences in DMV studies is that the individual preferences elicited during or after group interaction are not independent, violating aggregation assumptions (Howarth and Wilson 2006). In practice it has been (sometimes implicitly) assumed that individuals state their WTP post-deliberation independently.

One might argue that respondents who have participated in a deliberative process are no longer representative of the general population, and therefore aggregating their WTP would not be appropriate for social welfare estimation. However, the assumption of DMV studies is that with appropriate deliberative processes, any non-participant would come to the same conclusion as participants after deliberation. Then, if the choice outcome of the deliberation is based on a Habermas-type rational communicative debate, and the public supports the democratic rule that such rationality is required for policy choices, the DMV outcome would apply to all citizens.

A second critique concerns equity. Individual WTP depends on ability to pay (Jacobsen and Hanley 2009). If the WTP estimates are used in cost-benefit analysis, this may lead to decisions for options preferred by a minority of more affluent individuals even if an alternative was preferred by a majority of poorer people. The Kaldor-Hicks compensation rule, which forms the rationale in cost-benefit analysis, requires potential compensation at the point of weakly Pareto efficient solutions to ensure no one can be made better off; it hence artificially separates efficiency and equity criteria and strategies. However, the deliberative democratic theoretical basis of DMV makes the issue of material (and other types of) inequality and actual, rather than potential, pay-offs highly relevant. Empirical findings show that people are inequality averse (Groom and Maddison 2013), thereby providing legitimacy to correct additive aggregated WTP values for policy purposes in certain cases (Johansson-

Stenman 2005; Baumgärtner et al. 2017; Drupp et al. 2018). Studies on the social costs of carbon have used utilitarian social welfare functions, but then imposed equity-weighting (e.g. Anthoff et al. 2009), but in practice such corrections are rare. Alternative social welfare functions, such as prioritarian functions or Rawlsian maximin or leximin (Adler 2012), put greater marginal weight on lower welfare levels, but are also rarely used in applied studies.

DMV studies either avoid statistical aggregation by assuming any equity weighting to be embedded in the group decision, or by explicitly asking the group to choose a social welfare function. Kenter et al. (2016c) asked groups to state a fair price for the good. This WTP format merges issues of equality and efficiency, as it aims to reveal the maximum WTP that most if not all citizens would be willing and able to pay.

3.2.3 Summary of main design recommendations

Sampling: Where budgets are sufficient, we recommend samples as large as possible based on stratified random sampling along demographic variables, modified if necessary for Value Juries (e.g. to include marginalized groups).

Logistics: Room layouts and locations should be as neutral as possible; seating arrangements and access to information materials should not imply hierarchies or create sub-groups.

Information provision: Information (content and mode) should be understandable to all participants, transparent and neutral (diverse and representative). We advise to pre-test to adapt information provision to the needs of participants and test framing effects. Researchers should reflect on the framing of information and accountability of participants and external experts or witnesses.

Moderation: We recommend hiring a professional, experienced moderator whenever possible, and principally for Value Juries where the process of reaching group consensus must be expertly facilitated. Moderators should be familiar with the topic under investigation and related debates, be able to identify frames (e.g. according to a list of predefined frames), be aware of potential power asymmetries, and be able to deal with them through reframing or impartiality. Inclusive, free, unbiased and equal participation should be strived for, e.g. by means of moderation guides or deliberative formats.

Group effects: We recommend addressing potential negative group dynamics through planning means of detection of polarisation and social norms, forming heterogeneous groups and using professional moderation. Non-consensual or individual WTP formats should be considered if polarisation or other strong group effects are very likely to occur.

Topics: We recommend including and facilitating discussion of a broad range of relevant topics. Whether setting topics *a priori*, letting topics emerge from the group during deliberation (to avoid researcher-driven framing effects), or hybrid formats are most suitable depends on the purpose of the study.

Duration: The duration of deliberative sessions should be based on their purpose. Participants should be offered sufficient time for discussion and the opportunity to express their views, especially in Value Juries. We recommend a flexible deliberative process design (e.g. time buffer, multiple meetings) and inclusion of multiple forms of expression to address variation in the capacity of participants to engage in communicative rational debate.

Democracy/transparency: The deliberative process, its goals and the use of the results should be made transparent to both participants and non-participants *a priori*.

Elicitation format and aggregation: WTP elicitation format should be consistent with the purpose and theoretical foundation of the study. The aggregation method and the social welfare function (equity considerations) can be discussed with participants.

Pre-test: We strongly recommend practitioners to pre-test both the deliberative and the WTP elicitation elements (including hypothetical scenario/market).

3.3 Analysis

3.3.1 Deliberation process and qualitative content

Many existing DMV studies report very little detail about the content of the discussions and debates, which we see as both a missed opportunity and a lack of transparency and reliability. To tap into the rich qualitative dataset of the discussions, recording the interactions during the deliberation is a first step. Data collection can be done by taking notes (transcripts), recording audio(-visuals) (MacLean and Burgess 2009), or – when using software-supported platforms – by recording the digital interactions. Participant observation, either by researchers or video, can be useful to analyse non-verbal communication, respondents' engagement and group dynamics expressed through body language (e.g. Lienhoop and Fischer, 2009; see also *Group effects*). Body language conveys information, for example about embodied emotions, agreement, or engagement - all relevant to the quality of engagement in deliberation, but it does not lend itself well to comparison of value judgements or perceptions between people (Tillman 2016).

The main purpose of recording the discussion is to evaluate the content and quality of the deliberative process. Content analysis, where a coding scheme is used to convert the qualitative data partially to numerical data, can either be inductive or deductive. The latter aims to analyse the content of discussions around themes defined by the researcher. Q-methodology is another option to study different viewpoints. Various software packages support this type of analysis. One potential theme would be to look for demonstration of (shifts to) transcendental values (Kenter et al. 2016a,c) and other-regarding behaviour or expression of second-order preferences (see *Validity of valuation outcomes*). Deliberation should not be assumed to lead to value pluralism or acceptance of multiple viewpoints; this assertion should be tested explicitly (Lo and Spash 2013).

The analysis of the content of discussions is particularly important when the objective of deliberation is to increase the information basis of individual choices, or to improve the substantive quality of choice through sharing knowledge and information. The analysis should aim to see whether participants were willing to learn and adapt their values, beliefs and perceptions, and to reduce their confidence in their original beliefs (Peter 2013) – this would demonstrate the value of the procedure. Such changes can be assessed not only by comparing CE or CV responses before and after deliberation, but also with additional qualitative methods to explore the impact of deliberation on beliefs according to participants (e.g. Robinson and Bryan 2013).

Evaluations of the effects of deliberation on outcomes have looked at cognitive effects (information), attitudinal and opinion shifts, political sophistication and knowledge, argument quality, socio-affective effects (satisfaction), and behavioural effects (political participation intent) (see Gastil et al. 2017, Zhang 2015). One way to assess the quality of deliberative processes is the *Discourse Quality Index* developed by Steenbergen et al. (2003). This approach involves the coding of voice recordings for their discourse quality with five coding

categories that follow the principles of Habermas' discourse ethics. These categories include (i) participation, which measures the participant's ability to freely engage in the discussion, (ii) the level of justification, which measures the extent to which a speaker gives justifications for demands, (iii) content of justifications, which depicts whether appeals are made in terms of group interests, the common good, or both, (iv) respect, including three types of respect – for the group(s) affected by a certain policy intervention, towards demands of others, for voiced counterarguments, (v) constructive politics, assessing to what extent is constructive consensus building possible. Similar, partly complementary measures of various aspects of discourse quality have been proposed e.g. by Dutwin (2003), Holzinger (2004) and Stromer-Galley (2007). In a deliberative valuation study, Murphy et al. (2017) used a short list of debriefing questions to check the quality of deliberations as assessed by the participants themselves (Table 1). These questions cover the outputs of deliberation related to the environmental policy issue; additional questions could be asked about increases in deliberation capacity of participants (Kenter et al. 2016a) – a potential normative objective of the deliberation.

[Table 1 here]

Discussion recordings can also be used to illuminate the motives underlying participants' preferences, and how these vary across groups, as these provide information on the arguments for or against an environmental change and are therefore of interest to policy makers. This wealth of information has high political value in addition to WTP results.

Validity of valuation outcomes

Validity criteria for SP studies can be translated into criteria for DMV studies. Firstly, like protest bids in SP studies where participants refuse to reveal their true WTP, participants in DMV may protest against the process or aims of deliberation or refuse to engage in discussion or consensus-seeking group WTP formulation (Saam 2017). Lack of engagement in deliberation reduces the content validity of the DMV study. Reasons for protest and refusal can usually be traced back to dissatisfaction with the framing, mode or structure of the deliberation (Saam 2017). This underscores the importance of careful design, pretesting and skilled facilitation for avoiding protest behaviour.

The theoretical validity of DMV studies can be tested in different ways and depends on the assumed rationality. DMV studies aiming to improve conforming to the standard axioms of rational choice theory could build in tests of lexicographic or discontinuous preferences, transitivity, monotonicity, and stability (including order effects), and scope and embedding tests (Rakontonarivo et al. 2016), ideally pre- and post-deliberation.

Lancsar and Swait (2014) propose to use content analysis of qualitative information to assess several validity issues. This includes choice set formation, where participants eliminate alternatives during elicitation, for example because of non-negotiable constraints or incommensurable values. Next is the evaluation process, which supports the specification of the systematic utility or WTP function: should a linear or non-linear function be used, are attributes ignored (in case of a CE format), where lies the largest preference heterogeneity among participants, what is the influence of other agents on choices? The evaluation of this decision process may also reveal which decision rule participants use (Schkade and Payne 1994): are they rational utility maximisers (consequentialist, compensatory), or random among acceptable alternatives (satisficing)? Do they rule out alternatives (elimination-by-aspect) as they consider thresholds, social or ethical norms, time or legislative constraints (deontologic, non-compensatory)? If risk is an element of the choice, what type of choice

behaviour under risk do individuals and groups display, and how does deliberation affect this? Do other participants influence the chosen decision rule of the individual? In studies where group choices are elicited, the discussion during the choice processes may provide information that reveals the decision-making strategy that the group employs and, akin to “think-aloud” designs in individual SP studies, may help to understand how decisions were made and what information was used. Building upon Sen’s (1977) distinction between first-order and second-order preferences (i.e. “conventional” preference orderings vs. meta-orderings over possible preference orderings, e.g. given external constraints), Bartkowski and Lienhoop (2018) suggest that content analysis of audio- or video-recorded deliberations may allow for disentangling the two and thus “filter out” the influence of perceived constraints on first-order preferences.

Process-criteria may be sufficient for studies that employ deliberation for normative reasons. The evaluation of the process of the workshop requires reflection on the realisation of the design: representativeness in sampling, fairness in deliberation and ability to deliberate, access to information, the process of facilitation and reporting, transparency and accountability (see Fish et al. 2011).

For substantive purposes, some studies ask respondents in follow-up questions whether their satisfaction with WTP results increased with deliberation, e.g. how certain they are that WTP reflects their preferences. Such questions provide some information about the legitimacy of using the results for policy (e.g. MacMillan et al. 2006; Völker and Lienhoop 2016), although respondent-reported satisfaction measures may not be reliable (Hess and Beharry-Borg 2012).

Group effects

Reflection on the functioning and dynamics of a group, during or after the DMV activity, can be guided with questions such as: what is the role that each member plays? Which members have formed coalitions, who sits together? Do boundaries between groups prevent new information from being adopted? What information is (not) shared, accepted, neglected, synthesised? How do people communicate, who do they address? Which social norms regulate what can be said? How are conflicts resolved and how is (dis)agreement shown? Who creates distractions? These questions help to understand the group process and roles that various members adopt in group processes (Kottler and Englar-Carlson 2014). In turn, the revealed group patterns may further guide the analysis of the qualitative content of the deliberations.

We have not come across any study that explicitly focuses on assessing and addressing power asymmetries and inequality aspects that may arise in a deliberative process. Besides creating heterogeneous groups, we suggest addressing power and inequality through a quality assessment of the discourse. In their Discourse Quality Index (DQI, see also above), Steenbergen et al. (2003) provide codes to measure a speaker’s respect towards social groups represented by other participants as well as respect towards demands and counterarguments voiced during the discussion. This measurement involves the coding of transcripts from voice recordings and provides an *ex post* assessment.

3.3.2 WTP analysis

Most DMV studies have used standard random utility models (RUM). In general, DMV studies tend to have smaller samples, especially for studies using political representativeness, which can lead to insignificant parameter estimates with small t-statistics, or large confidence

intervals around WTP estimates. There are several possible adjustments for DMV in the estimation of WTP statistics compared to standard SP studies. Valuation Workshop studies typically use standard logit models for individual choices. Not many studies account for group influences, or even the correlation between choices made by members who formed part of the same group. Völker and Lienhoop (2016) assess the relation between the initial individual distribution of preferences on the group outcome, comparing homogenous (all proponents) and heterogeneous (proponents and opponents) groups. They estimate generalised multinomial logit (G-MNL) models, which allow the scale parameter (rather than the taste parameters pertaining to the attributes) to vary with group type. Although they observe changes in choices and WTP after deliberation, the marginal WTP estimates are not significantly different for either group type.

The number of groups in DMV studies is often small, so model results lack the statistical power necessary for WTP estimates with small confidence intervals. Therefore, even when participants in groups are to make a consensual choice based on their citizen preferences, most studies have analysed the choices as if they were independent, individual choices, and do not correct for intra-group correlation (e.g. Álvarez-Farizo & Hanley 2006; Álvarez-Farizo et al. 2007, 2009; Ito et al. 2009; Kenter et al. 2016c). One exception is the study by Kenter et al. (2011) in which 46 groups provided collective choices and these collective choices were taken as the dependent variable in the choice models. Although we recognise potential infeasibility of organising multiple deliberative sessions (which may come at the expense of discursive quality), assuming independence of individual choices and utility maximisation in consensual choices for ‘social WTP’ only seems valid in deliberations of the highest quality where individual and group choices are equalised (i.e. where deliberation has structurally transformed preferences if consumer and citizen choices were different before deliberation).

Further options for analysis may be provided by a limited number of SP studies that have looked at interdependent choices of households (Adamowicz et al. 2005), where households act as small, non-productive groups that are constrained by a common budget. Scarpa et al. (2012) model the difference in taste parameters between members of couples within a standard WTP-space framework, through estimating the deviance of women from the couple’s preferences. Rungie et al. (2014) control for the choice dependence in their RUM by decomposing the error term of the utility functions to identify the effect of each individual on household choice. More specifically, this model accounts for the influence of individuals’ taste coefficients as latent determinants of the joint choice of their household in a structural choice modelling framework. Dosman and Adamowicz (2006) propose a model where bargaining is through their individual utilities: the parameters corresponding to the joint decision are defined as a weighted mean of the individual coefficients. The “bargaining coefficient” can identify polarisation or denial by individuals. Mariel et al. (2018) assess the individually stated choices of parent couples together with the couples’ revealed joint choice in a bargaining model. They account for the (fixed) difference in scale parameter from their SP and revealed preference (RP) data. Other ways of establishing these weights can be explored if observational data that reveal dominant members are available.

None of these studies found particularly strong differences between husbands and wives, suggesting that intra-household differences are not sufficiently large to warrant separate studies. However, further research is necessary to assess whether this conclusion transfers to DMV studies with unrelated group members who have different budgets, although the difficulty of obtaining large enough sample sizes (number of DMV groups) may prevent this analysis due to lack of statistical power.

If the assessment of the decision-making process revealed divergent (non-maximising) behaviour, alternative models could be employed. One option is the random regret model proposed by Chorus (2012), which takes the minimization of regret of not choosing the best option, rather than maximization of utility, as its objective. Such a specification may be useful in the context of consensus.

3.3.3 Summary of main recommendations for analysis

Deliberation and qualitative content: We recommend recording of deliberations as precondition for qualitative analysis, for which standard methods for content analysis can be used. If possible, researchers should pay attention to non-verbal communication. Deliberation quality should be assessed, for example, by means of standard indices or debriefing questions.

Validity: We recommend using content analysis to evaluate the validity of valuation in terms of *a priori* rationality assumptions, incentive compatibility and “protest non-deliberation”. Assessment of the decision-making process or choice structure of the participants can be used to inform the WTP analysis. The process (for studies with a normative purpose) or the perceived quality of the outcomes (for substantive purposes) should be evaluated.

Group effects: We recommend reflecting upon the functioning and dynamics of the group deliberation and linking this to individual (Valuation Workshops) or group (Value Juries) choices.

WTP analysis: We recommend choosing econometric approaches to WTP analysis depending on the results of content analysis of deliberations, such as alternatives to models for utility maximisation or the analysis of heterogeneity across groups with shared views.

3.4 External validity

Depending on the motivation for and purpose of deliberation and the chosen theoretical foundation, there may be options to compare deliberated and non-deliberated studies. A robust test of external validity, both criterion and convergent validity, would require that the epistemological backgrounds of the data-generating processes are the same; if these paradigms are deemed incompatible, then comparison is not meaningful (Raymond et al. 2014).

In the few existing comparisons of non-deliberated and deliberated, individually elicited WTP values, some studies find no significant difference (e.g. Álvarez-Farizo et al. 2007, 2009; Christie et al. 2006; Dietz et al. 2009; Lienhoop and Fischer 2009; Lienhoop and Völker 2016; Philip and MacMillan 2005). Others find non-deliberated WTP estimates to be higher (Álvarez-Farizo and Hanley 2006) or lower than deliberated WTP values (e.g., Balderas-Torres et al. 2013; MacMillan et al. 2006; Urama and Hodge 2013). This can relate to different theoretical foundations, non-additive values, or biases in either individual or group WTP values. Even if the epistemological differences were deemed irrelevant, then the comparisons cannot be deemed robust because the deliberative studies confound effects of learning (gaining information to construct and update preferences) and other interpersonal effects (dominance, pressure, conformity to social norms). Moreover, some of these studies are based on between-sample comparison while others use within-sample tests. Within-sample tests of the effect of deliberation on choices and WTP values (e.g. Dietz et al. 2009) are not clean because they are affected by order effects. Between-sample comparisons, where

individual WTP without deliberation is compared to Valuation Workshop WTP estimates, may suffer from differences in sample characteristics, especially if the sampling strategies differ.

Robust convergent validity tests of hypothetical DMV studies are also difficult because it is highly unlikely to find an existing DMV-type process combined with an RP or market outcome related to the same environmental good. Similar problems exist for individual SP studies. Potentially, one could compare a hypothetical DMV study with a real or simulated deliberative format, such as a deliberated referendum or poll, where some form of price, tax or other WTP expression related to the same environmental change is discussed, but with a method other than an SP technique. Ensuring that valuation studies are incentive-compatible and avoiding hypothetical bias seems to play a key role in external validity of SP studies (Zawojka and Czajkowski 2017), a result which may also hold for DMV.

Beyond WTP comparison, one could test the validity of DMV studies using triangulation by comparing the policy recommendations of WTP studies to those emerging from other methods. For example, Hattam et al. (2015) compare Citizen's Juries with individual WTP estimates. Arguably, the real value, but also difficulty, arises when findings from different methods are contradictory; Hattam et al. (2015) argue that such mixed messages should be understood as challenges and focal points for environmental management (for a similar argument in the context of deliberative mini-publics, see Lafont 2017). The pragmatic approaches proposed by Raymond et al. (2014) combine individual and deliberative methods for assessing social values to enhance legitimacy, address potential value incommensurability, and achieve large scale representativeness. By comparing valuation workshops with a control group consisting of in-person interviews, Lienhoop and MacMillan (2007a) found that the deliberative process was superior to interviews: the R^2 -values show that WTA is better explained by the valuation workshop data, and the valuation workshop produced fewer protest responses and item-non-responses.

Overall, we recommend caution with interpreting comparisons between deliberated and non-deliberated WTP estimates, both within-sample and between-samples. Triangulation in a mixed-method approach may improve political relevance of results. Reference studies may be used that are compatible with the purpose and theoretical background of the research; if not available, differences in policy recommendations from two studies should be translated into management and further stakeholder engagement activities.

3.5 Reporting

Since no uniform standard for conducting DMV can be prescribed - which would be contrary to the very idea of this valuation approach - it is essential for validity and relevance of DMV studies that the choices made during both design and analysis are reported as transparently as possible.

First, we recommend that studies clearly state the purpose of deliberation and the theoretical foundation (Section 3.1), and report on the extent to which the deliberation fulfilled this purpose (Section 3.3). Secondly, to complement standard reporting requirements for individual SP studies, we provide a list of issues that researchers should report on for reliability purposes (see Text Box 1).

[Text Box 1 here]

In addition to presenting results, practitioners are advised to reflect on whether the results are fit for use in their wider political context, whether that is for democratisation or legal or accounting purposes. Most of the quality criteria for individual SP studies apply, including the sound scientific basis related to the environmental change and its objective description, relevance and realism. The analyst should also reflect on whether the DMV study measured preferences for the environmental good it intended to measure and follow statistical guidelines for WTP analysis.

4 Open questions and future research directions

Strictly speaking, DMV is not one method, but rather a diverse method type. The unifying characteristic is that it combines elements of conventional SP valuation methods with insights from deliberative democratic theory. Depending on the specific purpose, from preference formation to value pluralism and democratisation, elements of both underlying theories are mixed to different extents. Moreover, DMV research is still rather young, so there are plenty directions for future research that will contribute to increasing DMV's validity and relevance. In this final section, we suggest those issues that we see as most important (see also Kenter et al. 2016b).

We have highlighted several design issues that can introduce biases in DMV-generated WTP values resulting from information provision, group effects and time to think to name a few. These require systematic and careful testing to develop validity and reliability criteria for DMV. Currently, there is still too little research available to derive such recommendations. For instance, there is the question of temporal stability of deliberation-induced preference change. Market Stall applications have tried to tackle this by introducing an additional elicitation round a couple of weeks after the DMV workshop (e.g. Lienhoop and Völker 2016). However, similar to temporal stability tests in standard SP studies, it is nearly impossible to control for confounding factors and disentangle the effects of deliberation from other influences in the meantime. In the language of Raymond and Kenter (2016), the issue may be interpreted as the difficulty to distinguish between (supposedly stable) transcendental values and the more dynamic contextual values.

Another important issue that can and should be tested in DMV settings is framing, especially how making certain topics explicit, whether in individual or group-exercises prior to the WTP question, influences the content of deliberation and the WTP results or policy recommendations, including well-known framing effects such as gain-loss asymmetry (Tversky and Kahneman 1981).² Validity tests of consequentiality and protest bids, related to (trust in) institutions, and their interaction with design choices to foster outcome and procedural fairness are important for deepening our understanding of the legitimacy of DMV studies (Bianchi et al. 2015).

With the scepticism about DMV due to group dynamics, systematic analysis of group behaviour and comparisons across groups of different composition could help to understand what triggers different types of group dynamics affect the valuation environmental changes. Interdisciplinary collaboration with psychologists and sociologists on group behaviour is advised. Group dynamics analysis could start with highly controlled settings, such as online DMV where participants do not see each other and factors such as tone of voice or body language are excluded (Gastil et al. 2017).

² The possibility of testing framing effects is an advantage of DMV that would fall into the substantive category. We thank an anonymous referee for this point.

Another open issue is the role of emotions in deliberation and preference elicitation, which would violate assumptions of neoclassical rationality, and to some extent rationalist communication, but supposedly expand the inclusiveness of deliberation if emotions are more important for certain groups. Recently, Hanley et al. (2017) found that emotions did not have a significant effect on preferences, WTP or model estimates elicited by means of a choice experiment in a laboratory setting. The role and normative interpretation of emotions in deliberative settings has been intensively debated; Saam (2017) argues, for example, that emotions can have detrimental effects in terms of equality of participation and democratic quality of deliberations. Emotions can influence group dynamics, which are one of the most difficult and complex challenges faced by DMV research. Emotional effects have not yet been studied in DMV context, and as their role is potentially larger in deliberation than in individual SP studies, there is a need for research into their detection, control (if deemed desirable) and analysis.

The effects of social choice rules and welfare and preference aggregation are another key contested issue in the DMV debate and beyond (Arrow 2012). One avenue of research would be to identify preferred aggregation rules in deliberated and non-deliberated settings. A second option would be to engage with recent developments in the economics literature on aggregation, where small group aggregation looks at production within groups rather than taking a group as a single decision maker (Chiappori and Ekeland 2011). A third approach would follow Murphy et al.'s (2017) analysis of different (in their case, ordinal) aggregation rules for individual preferences and their relationship with deliberated preferences. The relevance of this question seems to go well beyond DMV research, as conventional SP approaches also uniformly use simple additive aggregation approaches.

Given that DMV is still developing and searching for a coherent theoretical foundation (Bartkowski and Lienhoop 2018; Bunse et al. 2015), its relevance for decision making must be assessed on a case-by-case basis. In terms of structured decision-making support, only Valuation Workshops seem to currently have some limited applicability – for instance, for large enough samples, DMV results can be used in cost–benefit analysis (CBA) if the scope is local or regional. Conversely, DMV results are not suitable for standard environmental-economic accounts, as even the consistency of conventional SP methods with them has been debated (Obst et al. 2016; Droste and Bartkowski 2017). DMV-derived values seem consistent with the pluralist approach of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) (Pascual et al. 2017). Others may deem DMV to be only suitable for qualitative support of decision-making processes. More research into its relevance for and uptake in decision making is needed.

References

- Abels, G., 2007. Citizen involvement in public policy-making: Does it improve democratic legitimacy and accountability? The case of pTA. *Interdisciplinary Information Sciences*, 13(1), pp.103–116.
- Adamowicz, W., Hanemann, M., Swait, J., Johnson, R., Layton, D., Regenwetter, M., Reimer, T. and Sorkin, R., 2005. Decision strategy and structure in households: a “groups” perspective. *Marketing Letters*, 16(3-4), pp.387–399.
- Adler, M., 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford University Press, New York.
- Albon, S., Turner, K., Watson, B., Anger, A., Baker, J., Bateman, I., Bentley, S., Blyth, N., Bowles-Newark, N., Brown, C. and Brown, I., 2014. *UK National Ecosystem Assessment follow-on: Synthesis of key findings*. UNEP World Conservation Monitoring Centre, Cambridge, UK.
- Álvarez-Farizo, B. and Hanley, N., 2006. Improving the process of valuing non-market benefits: combining citizens’ juries with choice modelling. *Land Economics*, 82(3), pp.465–478.
- Álvarez-Farizo, B., Hanley, N., Barberan, R. and Lazaro, A., 2007. Choice modeling at the “market stall”: Individual versus collective interest in environmental valuation. *Ecological Economics*, 60(4), pp.743–751.
- Álvarez-Farizo, B., Gil, J.M. and Howard, B.J., 2009. Impacts from restoration strategies: Assessment through valuation workshops. *Ecological Economics*, 68(3), pp.787–797.
- Anthoff, D., Hepburn, C. and Tol, R.S., 2009. Equity weighting and the marginal damage costs of climate change. *Ecological Economics*, 68(3), pp.836–849.
- Arias-Maldonado, M., 2007. An imaginary solution? The green defence of deliberative democracy. *Environmental Values*, pp.233–252.
- Arrow, K.J., 2012. Social choice and individual values (Vol. 12). Yale university press.
- Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R. and Schuman, H., 1993. Report of the NOAA panel on contingent valuation. *Federal register*, 58(10), pp.4601–4614.
- Balderas-Torres, A., MacMillan, D.C., Skutsch, M. and Lovett, J.C., 2013. The valuation of forest carbon services by Mexican citizens: The case of Guadalajara city and La Primavera biosphere reserve. *Regional Environmental Change*, 13(3), pp.661–680.
- Bartkowski, B. and Lienhoop, N., 2018. Beyond rationality, towards reasonableness: Enriching the theoretical foundation of Deliberative Monetary Valuation. *Ecological Economics*, 143, pp.97–104.
- Baumgärtner, S., Drupp, M.A., Meya, J.N., Munz, J.M. and Quaas, M.F., 2017. Income inequality and willingness to pay for public environmental goods. *Journal of Environmental Economics and Management*, 85, pp.35–61.
- Bellamy, R., Lezaun, J. and Palmer, J., 2017. Public perceptions of geoengineering research governance: An experimental deliberative approach. *Global Environmental Change*, 45, pp.194–202.
- Ben-Arieh, D. and Easton, T., 2007. Multi-criteria group consensus under linear cost opinion elasticity. *Decision Support Systems*, 43(3), pp.713–721.
- Bianchi, E.C., Brockner, J., Van den Bos, K., Seifert, M., Moon, H., van Dijke, M. and De Cremer, D., 2015. Trust in decision-making authorities dictates the form of the interactive relationship between outcome fairness and procedural fairness. *Personality and Social Psychology Bulletin*, 41(1), pp.19–34.
- Bishop, R.C. and Boyle, K.J., 2018. Reliability and validity in nonmarket valuation. *Environmental and Resource Economics*, in press.

- Brewer, P.R. and Gross, K., 2005. Values, framing, and citizens' thoughts about policy issues: Effects on content and quantity. *Political Psychology*, 26(6), pp.929–948.
- Brodbeck, F.C., Kerschreiter, R., Mojzisch, A. and Schulz-Hardt, S., 2007. Group decision making under conditions of distributed knowledge: the information asymmetries model. *Academy of Management Review* 32(2), pp.459–479.
- Brouwer, R., Dekker, T., Rolfe, J. and Windle, J., 2010. Choice certainty and consistence in repeated choice experiments. *Environmental and Resource Economics*, 46, 93–109.
- Brown, T.C., Peterson, G.L. and Tonn, B.E., 1995. The values jury to aid natural resource decisions. *Land Economics*, 71(2), pp.250–260.
- Bulkeley, H. and Mol, A.P., 2003. Participation and environmental governance: consensus, ambivalence and debate. *Environmental Values*, 12(2), pp.143–154.
- Bunse, L., Rendon, O. and Luque, S., 2015. What can deliberative approaches bring to the monetary valuation of ecosystem services? A literature review. *Ecosystem Services*, 14, pp.88–97.
- Burgess, J., Lamb, M. and Harrison, C.M., 1988a. Exploring environmental values through the medium of small groups: 1. Theory and practice. *Environment and Planning A*, 20(3), pp.309–326.
- Burgess, J., Lamb, M. and Harrison, C.M., 1988b. Exploring environmental values through the medium of small groups: 2. Illustrations of a group at work. *Environment and Planning A*, 20(3), pp.457–476.
- Calvert, A. and Warren, M.E., 2014. Deliberative democracy and framing effects: Why frames are a problem and how deliberative mini-publics might overcome them. In: Grönlund, K., Bächtiger, A. and Setälä, M. (Ed), *Deliberative Mini-Publics: Involving Citizens in the Democratic Process*, ECPR Press, Colchester, UK. pp.203–24.
- Carter, A., Croft, A., Lukas, D. and Sandstrom, G., 2017. Women's visibility in academic seminars: women ask fewer questions than men. *arXiv:1711.10985*.
- Chiappori, P.A. and Ekeland, I., 2011. New developments in aggregation economics. *Annual Review of Economics*, 3(1), pp.631–668.
- Chilvers, J., 2009. Deliberative and participatory approaches in environmental geography. In: Noel, C., David, D., Diana, L. and Bruce, R. (Eds), *A Companion to Environmental Geography*, Wiley Blackwell, Chichester, UK, pp.400–417.
- Chilvers, J. and Kearnes, M., 2015. *Remaking Participation: Science, Environment and Emergent Publics*. Taylor & Francis, London.
- Chorus, C., 2012. Random regret minimization: an overview of model properties and empirical evidence. *Transport Reviews*, 32(1), pp.75–92.
- Christie, M., Hanley, N., Warren, J., Murphy, K., Wright, R. and Hyde, T., 2006. Valuing the diversity of biodiversity. *Ecological Economics*, 58, pp.304–317.
- Christie, M., Fazey, I., Cooper, R., Hyde, T. and Kenter, J.O., 2012. An evaluation of monetary and non-monetary techniques for assessing the importance of biodiversity and ecosystem services to people in countries with developing economies. *Ecological Economics*, 83, pp.67–78.
- Clifford, S. 2012. Making disability public in deliberative democracy. *Contemporary Political Theory* 11(2), pp.211–228.
- Cohen, J., 1997. Deliberation and democratic legitimacy. In: Bohman, J. and Rehg, W. (Eds.), *Deliberative Democracy: Essays on Reason and Politics*. MIT Press, Cambridge, MA, pp. 67–91.
- Coote, A. and Lenaghan, J., 1997. *Citizens' Juries: Theory into Practice*. Institute for Public Policy Research, London, UK.

- Crompton, A., Waring, J., Roe, B. and O'Connor, R., 2017. Are we all on the same page? A qualitative study of the facilitation challenges associated with the implementation of deliberative priority-setting. *Public Management Review*, pp.1–20.
- Cuppen, E., 2018. The value of social conflicts. Critiquing invited participation in energy projects. *Energy Research & Social Science*, 38, pp.28–32.
- Curato, N., Dryzek, J.S., Ercan, S.A., Hendriks, C.M. and Niemeyer, S., 2017. Twelve Key Findings in Deliberative Democracy Research. *Daedalus*, 146(3), pp.28–38.
- Czajkowski, M., Vossler, C.A., Budziński, W., Wiśniewska, A. and Zawojka, E., 2017. Addressing empirical challenges related to the incentive compatibility of stated preferences methods. *Journal of Economic Behavior & Organization*, 142, pp.47–63.
- Davies, G. and Burgess, J., 2004. Challenging the 'view from nowhere': citizen reflections on specialist expertise in a deliberative process. *Health & Place*, 10(4), pp.349–361.
- Degeling, C., Carter, S.M. and Rychetnik, L., 2015. Which public and why deliberate? – A scoping review of public deliberation in public health and health policy research. *Social Science & Medicine*, 131, pp.114–121.
- Dietz, T., Stern, P.C. and Dan, A., 2009. How deliberation affects stated willingness to pay for mitigation of carbon dioxide emissions: An experiment. *Land Economics* 85, pp.329–347.
- Dosman, D. and Adamowicz, W., 2006. Combining stated and revealed preference data to construct an empirical examination of intrahousehold bargaining. *Review of Economics of the Household*, 4(1), pp.15–34.
- Droste, N. and Bartkowski, B., 2017. Ecosystem service valuation for national accounting: A reply to Obst, Hein and Edens (2016). *Environmental and Resource Economics*, in press.
- Druckman, J.N., 2004. Political preference formation: Competition, deliberation, and the (ir) relevance of framing effects. *American Political Science Review*, 98(4), pp.671–686.
- Drupp, M.A., Meya, J.N., Baumgärtner, S. and Quaas, M.F., 2018. Economic inequality and the value of nature. *Ecological Economics*, in press.
- Dryzek, J.S., 2000. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford University Press, New York.
- Dryzek, J.S., 2013. The deliberative democrat's *Idea of Justice*. *European Journal of Political Theory*, 12(4), pp.329–346.
- Dryzek, J.S. and Hendriks, C., 2012. Fostering deliberation in the forum and beyond. In: Fischer, F. and Gottweis, H. (Eds), *The Argumentative Turn Revisited: Public Policy as Communicative Practice*. Duke University Press, Durham, pp.31–57.
- Dryzek, J.S. and Niemeyer, S., 2008. Discursive representation. *American Political Science Review*, 102(4), pp.481–493.
- Dutwin, D., 2003. The character of deliberation: Equality, argument, and the formation of public opinion. *International Journal of Public Opinion Research* 15(3), pp.239–264.
- Edwards, D.M., Collins, T.M. and Goto, R., 2016. An arts-led dialogue to elicit shared, plural and cultural values of ecosystems. *Ecosystem Services*, 21, pp.319–328.
- European Commission, 2013. Mapping and Assessment of Ecosystems and Their Services. An Analytical Framework for Ecosystem Assessment under Action 5 of the EU Biodiversity Strategy to 2020. Technical Report No. 2013-067. Brussels: European Commission.
- Fern, E.F., 2001. *Advanced Focus Group Research*. SAGE Publications, Thousand Oaks.
- Fish, R., Burgess, J., Chilvers, J., Footitt, A., Haines-Young, R., Russel, D., Turner, K. and Winter, D.M., 2011. *Participatory and deliberative techniques to embed an ecosystems approach into decision making: an introductory guide*. DEFRA, London.

- Fishkin, J. S., 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. London, Yale University Press.
- Frame, B. and O'Connor, M., 2011. Integrating valuation and deliberation: the purposes of sustainability assessment. *Environmental Science & Policy*, 14(1), pp.1–10.
- French, D. and Laver, M., 2009. Participation bias, durable opinion shifts and sabotage through withdrawal in citizens' juries. *Political Studies*, 57(2), pp.422–450.
- Gastil, J., Richards, R., Ryan, M. and Smith, G., 2017. Testing Assumptions in Deliberative Democratic Design: A Preliminary Assessment of the Efficacy of the Participedia Data Archive as an Analytic Tool. *Journal of Public Deliberation*, 13(2).
- Gibbs, A., 1997. Focus groups. *Social Research Update*, 19(8), pp.1–8.
- Greenberg, J., 1976. The role of seating position in group interaction: A review, with applications for group trainers. *Group & Organization Studies*, 1(3), pp.310–327.
- Gregory, R.S., 2000. Valuing environmental policy options: a case study comparison of multiattribute and contingent valuation survey methods. *Land Economics*, 76(2), pp.151–173.
- Gregory, R., Fischhoff, B. and McDaniels, T., 2005. Acceptable input: using decision analysis to guide public policy deliberations. *Decision Analysis*, 2(1), pp.4–16.
- Gómez-Baggethun, E. and Martín-López, B., 2015. Ecological economics perspectives on ecosystem services valuation. In: Martinez-Alier, J. and Muradian, R. (Eds.), *Handbook of Ecological Economics*. Edward Elgar Publishing, Cheltenham, pp.260–282.
- Gómez-Baggethun, E. and Ruiz-Pérez, M., 2011. Economic valuation and the commodification of ecosystem services. *Progress in Physical Geography*, 35(5), pp.613–628.
- Groom, B. and Maddison, D., 2013. Non-identical quadruplets: Four new estimates of the elasticity of marginal utility for the UK. *Centre for Climate Change Economics and Policy Working Paper*, 141.
- Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R., 2009. *Survey Methodology*. Wiley and Sons, New York.
- Habermas, J., 1984. *The Theory of Communicative Action* (Vol. 1). Beacon Press, Boston.
- Habermas, J., 1987. *The Theory of Communicative Action* (Vol. 2). Beacon Press, Boston.
- Habermas, J., 1996. *Between Facts and Norms*. MIT Press, Cambridge, MA.
- Hanley, N., Boyce, C., Czajkowski, M., Tucker, S., Noussair, C. and Townsend, M., 2017. Sad or Happy? The Effects of Emotions on Stated Preferences for Environmental Goods. *Environmental and Resource Economics*, 68(4), pp.821–846.
- Hattam, C., Böhnke-Henrichs, A., Börger, T., Burdon, D., Hadjimichael, M., Delaney, A., Atkins, J.P., Garrard, S. and Austen, M.C., 2015. Integrating methods for ecosystem service assessment and valuation: Mixed methods or mixed messages? *Ecological Economics*, 120, pp.126–138.
- Hausman, J., 2012. Contingent valuation: from dubious to hopeless. *Journal of Economic Perspectives*, 26(4), pp.43–56.
- Hess, S. and Beharry-Borg, N., 2012. Accounting for latent attitudes in willingness-to-pay studies: the case of coastal water quality improvements in Tobago. *Environmental and Resource Economics* 52(1), pp.109–131.
- Himmelroos, S., 2017. Discourse quality in deliberative citizen forums – A comparison of four deliberative mini-publics. *Journal of Public Deliberation*, 13(1).
- Holzinger, K., 2004. Bargaining through arguing: An empirical analysis based on speech act theory. *Political Communication* 21(2), pp.195–222.
- Howarth, R.B. and Wilson, M.A., 2006. A theoretical approach to deliberative valuation: Aggregation by mutual consent. *Land Economics*, 82(1), pp.1–16.

- Ito, N., Takeuchi, K., Kuriyama, K., Shoji, Y., Tsuge, T. and Mitani, Y., 2009. The influence of decision-making rules on individual preferences for ecological restoration: Evidence from an experimental survey. *Ecological Economics*, 68, pp.2426–2431.
- Jacobs, S., Dendoncker, N., Martín-López, B., Barton, D.N., Gomez-Baggethun, E., Boeraeve, F., McGrath, F.L., Vierikko, K., Geneletti, D., Sevecke, K.J. and Pipart, N., 2016. A new valuation school: Integrating diverse values of nature in resource and land use decisions. *Ecosystem Services*, 22, pp.213–220.
- Jacobsen, J.B. and Hanley, N., 2009. Are there income effects on global willingness to pay for biodiversity conservation? *Environmental and Resource Economics*, 43(2), pp.137–160.
- Johansson-Stenman, O., 2005. Distributional weights in cost-benefit analysis—should we forget about them? *Land Economics*, 81(3), pp.337–352.
- Johnson, D. W., and Johnson, F. P., 1994. *Joining Together: Group Theory and Group Skills* (5th ed.). Allyn & Bacon, Needham Heights, MA, US.
- Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R. and Tourangeau, R., 2017. Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), pp.319–405.
- Kenter, J.O., Hyde, T., Christie, M. and Fazey, I., 2011. The importance of deliberation in valuing ecosystem services in developing countries - Evidence from the Solomon Islands. *Global Environmental Change*, 21, pp.505–521.
- Kenter, J.O., 2016. Shared, plural and cultural values. *Ecosystem Services*, 21, pp.175–183.
- Kenter, J.O., Reed, M.S. and Fazey, I., 2016a. The Deliberative Value Formation model. *Ecosystem Services*, 21, pp.194–207.
- Kenter, J.O., Bryce, R., Christie, M., Cooper, N., Hockley, N., Irvine, K.N., Fazey, I., O'Brien, L., Orchard-Webb, J., Ravenscroft, N., Raymond, C.M., Reed, M.S., Tett, P. and Watson, V., 2016b. Shared values and deliberative valuation: Future directions. *Ecosystem Services* 21, pp.358–371.
- Kenter, J.O., Jobstovogt, N., Watson, V., Irvine, K.N., Christie, M. and Bryce, R., 2016c. The impact of information, value-deliberation and group-based decision-making on values for ecosystem services: integrating deliberative monetary valuation and storytelling. *Ecosystem Services*, 21, pp.270–290.
- Kenyon, W., Hanley, N. and Nevin, C., 2001. Citizens' juries: an aid to environmental valuation?. *Environment and Planning C: Government and Policy*, 19(4), pp.557–566.
- Kottler, J.A. and Englar-Carlson, M., 2014. *Learning Group Leadership: An Experiential Approach*. SAGE Publications, Thousand Oaks.
- Krueger, R.A. and Casey, M.A., 2014. *Focus Groups: A Practical Guide for Applied Research*. SAGE Publications, Thousand Oaks.
- Kuyper, J., 2017. The means and ends of deliberative democracy: Rejoinder to Gunn. *Critical Review*, 29(3), pp.1–23.
- Lafont, C., 2017. Can democracy be deliberative & participatory? The democratic case for political uses of mini-publics. *Daedalus*, 146(3), pp.85–105.
- Lancsar, E. and Swait, J., 2014. Reconceptualising the external validity of discrete choice experiments. *Pharmacoeconomics*, 32(10), pp.951–965.
- LaRiviere, J., Czajkowski, M., Hanley, N., Aanesen, M., Falk-Petersen, J. and Tinch, D., 2014. The value of familiarity: effects of knowledge and objective signals on willingness to pay for a public good. *Journal of Environmental Economics and Management*, 68(2), pp.376–389.

- Lienhoop, N. and MacMillan, D.C., 2007a. Contingent Valuation: Comparing participant performance in group-based approaches and personal interviews. *Environmental Values*, 16, pp.209–232.
- Lienhoop, N. und MacMillan, D., 2007b. Valuing Wilderness: Investigating Property Rights and WTA/WTP Using an Innovative Approach to Contingent Valuation. *Land Use Policy*, 24, pp.289–295.
- Lienhoop, N. and Fischer, A., 2009. Can you be bothered? The role of participant motivation in the valuation of species conservation measures. *Journal of Environmental Planning and Management*, 52(4), pp.519–534.
- Lienhoop, N., Bartkowski, B. and Hansjürgens, B., 2015. Informing biodiversity policy: The role of economic valuation, deliberative institutions and deliberative monetary valuation. *Environmental Science & Policy*, 54, pp.522–532.
- Lienhoop, N. and Völker, M., 2016. Preference refinement in deliberative choice experiments for ecosystem service valuation. *Land Economics*, 92(3), pp.555–577.
- List, C., Luskin, R.C., Fishkin, J.S. and McLean, I., 2012. Deliberation, single-peakedness, and the possibility of meaningful democracy: evidence from deliberative polls. *the Journal of Politics*, 75(1), pp.80–95.
- Lo, A.Y., 2013. Agreeing to pay under value disagreement: Reconceptualizing preference transformation in terms of pluralism with evidence from small-group deliberations on climate change. *Ecological Economics*, 87, pp.84–94.
- Lo, A.Y. and Spash, C.L., 2013. Deliberative Monetary Valuation: In search of a democratic and value plural approach to environmental policy. *Journal of Economic Surveys*, 27, pp.768–789.
- Lorenzoni, I. and Hulme, M., 2009. Believing is seeing: laypeople's views of future socio-economic and climate change in England and in Italy. *Public Understanding of Science*, 18(4), pp.383–400.
- MacLean, S. and Burgess, M.M., 2009. In the public interest: assessing expert and stakeholder influence in public deliberation about biobanks. *Public Understanding of Science* 19(4), pp.486–496.
- MacMillan, D.C., Philip, L., Hanley, N. and Álvarez-Farizo, B., 2002. Valuing the non-market benefits of wild goose conservation: a comparison of interview and group based approaches. *Ecological Economics*, 43(1), pp.49–59.
- MacMillan, D., Hanley, N. and Lienhoop, N., 2006. Contingent valuation: Environmental polling or preference engine? *Ecological Economics*, 60(1), pp.299–307.
- Mansbridge, J., Bohman, J., Chambers, S., Christiano, T., Fung, A., Parkinson, J., Thompson, D.F. and Warren, M.E., 2012. A systemic approach to deliberative democracy. In: Parkinson, J. and Mansbridge, J. (eds.). *Deliberative Systems: Deliberative Democracy at the Large Scale*, Cambridge University Press, Cambridge, pp.1–26.
- Mariel, P., Scarpa, R. and Vega-Bayo, A., 2018. Joint parental school choice: Exploring the influence of individual preferences of husbands and wives. *Regional Science and Urban Economics*, 68, pp.23–35.
- Millennium Ecosystem Assessment (MEA), 2005. Ecosystems and Human Well-being: General Synthesis. Washington, DC: World Resources Institute.
- Mendelberg, T., 2002. The deliberative citizen: Theory and evidence. *Political decision making, deliberation and participation*, 6(1), pp.151–193.
- Metzger, J., Soneryd, L. and Linke, S., 2017. The legitimization of concern: A flexible framework for investigating the enactment of stakeholders in environmental planning and governance processes. *Environment and Planning A: Economy and Space* 49(11), pp.2517–2535.

- Mitchell, R.C. and Carson, R.T., 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Resources for the Future, Washington.
- Murphy, M.B., Mavrommati, G., Mallampalli, V.R., Howarth, R.B. and Borsuk, M.E., 2017. Comparing group deliberation to other forms of preference aggregation in valuing ecosystem services. *Ecology & Society*, 22.
- Niemeyer, S. and Spash, C.L., 2001. Environmental valuation analysis, public deliberation, and their pragmatic syntheses: a critical appraisal. *Environment and Planning C: Government and Policy*, 19(4), pp.567–585.
- Nisbett, R.E. and Wilson, T.D., 1977. The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4), pp.250–256.
- Obst, C., Hein, L. and Edens, B., 2016. National accounting and the valuation of ecosystem assets and their services. *Environmental and Resource Economics*, 64(1), pp.1–23.
- Oehlmann, M. and Meyerhoff, J., 2017. Stated preferences towards renewable energy alternatives in Germany—do the consequentiality of the survey and trust in institutions matter?. *Journal of Environmental Economics and Policy*, 6(1), pp.1–16.
- O’Hara, S.U., 1996. Discursive ethics in ecosystems valuation and environmental policy. *Ecological Economics*, 16(2), pp.95–107.
- O’Neill, J., 2001. Representing people, representing nature, representing the world. *Environmental Planning C Government & Policy*, 19, pp.483–500.
- Orchard-Webb, J., Kenter, J.O., Bryce, R. and Church, A., 2016. Deliberative Democratic Monetary Valuation to implement the Ecosystem Approach. *Ecosystem Services*, 21, pp.308–318.
- Parks, S. and Gowdy, J., 2013. What have economists learned about valuing nature? A review essay. *Ecosystem Services*, 3, pp.1–10.
- Pascual, U., Balvanera, P., Díaz, S., Pataki, G., Roth, E., Stenseke, M., Watson, R.T., Başak Dessane, E., Islar, M., Kelemen, E., Maris, V., Quaas, M., Subramanian, S.M., Wittmer, H., Adlan, A., Ahn, S., Al-Hafedh, Y.S., Amankwah, E., Asah, S.T., Berry, P., Bilgin, A., Breslow, S.J., Bullock, C., Cáceres, D., Daly-Hassen, H., Figueroa, E., Golden, C.D., Gómez-Baggethun, E., González-Jiménez, D., Houdet, J., Keune, H., Kumar, R., Ma, K., May, P.H., Mead, A., O’Farrell, P., Pandit, R., Pengue, W., Pichis-Madruga, R., Popa, F., Preston, S., Pacheco-Balanza, D., Saarikoski, H., Strassburg, B.B., van den Belt, M., Verma, M., Wickson, F. and Yagi, N., 2017. Valuing nature’s contributions to people: the IPBES approach. *Current Opinion in Environmental Sustainability*, 26, pp.7–16.
- Pelletier, D., Kraak, V., McCullum, C., Uusitalo, U. and Rich, R., 1999. The shaping of collective values through deliberative democracy: An empirical study from New York’s North Country. *Policy Sciences*, 32(2), pp.103–131.
- Peter, F., 2013. The procedural epistemic value of deliberation. *Synthese*, 190(7), pp.1253–1266.
- Philip, L.J. and MacMillan, D.C., 2005. Exploring Values, Context and Perceptions in Contingent Valuation Studies: The CV Market Stall Technique and Willingness to Pay for Wildlife Conservation. *Journal of Environmental Planning and Management*, 48(2), pp.257–274.
- Rakotonarivo, O.S., Schaafsma, M. and Hockley, N., 2016. A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. *Journal of Environmental Management*, 183, pp.98–109.
- Rawls, J., 1971. *A Theory of Justice*. Harvard University Press, Cambridge, MA.
- Raymond, C.M., Kenter, J.O., Plieninger, T., Turner, N.J. and Alexander, K.A., 2014. Comparing instrumental and deliberative paradigms underpinning the assessment of

- social values for cultural ecosystem services. *Ecological Economics*, 107, pp.145–156.
- Raymond, C.M. and Kenter, J.O., 2016. Transcendental values and the valuation and management of ecosystem services. *Ecosystem Services*, 21, pp.241–257.
- Renn, O. and Schweizer, P.J., 2009. Inclusive risk governance: concepts and application to environmental policy making. *Environmental policy and governance*, 19(3), pp.174–185.
- Renn, O. and Webler, T., 1992. Anticipating conflicts: Public participation in managing the solid waste crisis. *GAIA-Ecological Perspectives for Science and Society*, 1(2), pp.84–94.
- Richards, D.R., Warren, P.H., Maltby, L. and Moggridge, H.L., 2017. Awareness of greater numbers of ecosystem services affects preferences for floodplain management. *Ecosystem Services*, 24, pp.138–146.
- Robinson, S. and Bryan, S., 2013. Does the process of deliberation change individuals' health state valuations? An exploratory study using the person trade-off technique. *Value in Health*, 16(5), pp.806–813.
- Robinson, J., Clouston, B., Suh, J. and Chaloupka, M., 2008. Are citizens' juries a useful tool for assessing environmental value?. *Environmental Conservation*, 35(4), pp.351–360.
- Rungie, C., Scarpa, R. and Thiene, M., 2014. The influence of individuals in forming collective household preferences for water quality. *Journal of Environmental Economics and Management*, 68(1), pp.161–174.
- Saam, N.J., 2017. Recognizing the emotion work in deliberation. Why emotions do not make deliberative democracy more democratic. *Political Psychology*, in press.
- Sagoff, M., 1988. *The Economy of the Earth: Philosophy, Law, and the Environment*. Cambridge University Press, Cambridge, New York.
- Sagoff, M., 1998. Aggregation and deliberation in valuing environmental public goods: A look beyond contingent pricing. *Ecological Economics*, 24(2-3), pp.213–230.
- Sandfort, J. and Quick, K., 2017. Deliberative technology: a holistic lens for interpreting resources and dynamics in deliberation. *Journal of Public Deliberation*, 13(1).
- Sandorf, E.D., Aanesen, M. and Navrud, S., 2016. Valuing unfamiliar and complex environmental goods: A comparison of valuation workshops and internet panel surveys with videos. *Ecological Economics*, 129, pp.50–61.
- Scarpa, R., Thiene, M. and Hensher, D.A., 2012. Preferences for tap water attributes within couples: an exploration of alternative mixed logit parameterizations. *Water Resources Research*, 48(1).
- Schaefer, M., Heinze, H.J., Rotte, M. and Denke, C., 2013. Communicative versus strategic rationality: Habermas Theory of Communicative Action and the social brain. *PloS one*, 8(5), pp.e65111.
- Schkade, D.A. and Payne, J.W., 1994. How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics & Management*, 26, pp.88–109.
- Schläpfer, F., 2016. Democratic valuation (DV): Using majority voting principles to value public services. *Ecological Economics*, 122, pp.36–42.
- Sen, A., 1977. Rational fools: A critique of the behavioural foundations of economic theory. *Philosophy and Public Affairs*, 6(4), pp.317–344.
- Sen, A., 1995. Environmental evaluation and social choice: Contingent valuation and the market analogy. *Japanese Economic Review*, 46(1), pp.23–37.
- Sen, A., 2009. *The Idea of Justice*. Harvard University Press, Cambridge, MA.
- Smith, G., 2005. *Beyond the Ballot: 57 Democratic Innovations from Around the World*. Power Inquiry, London.

- Smith G. and Wales, C., 1999. The theory and practice of citizens' juries. *Policy and Politics*, e27, pp.295–308.
- Smith, G. and Wales, C., 2000. Citizens' juries and deliberative democracy. *Political Studies*, 48(1), pp.51–65.
- Spash, C.L., 2007. Deliberative monetary valuation (DMV): Issues in combining economic and political processes to value environmental change. *Ecological Economics*, 63(4), pp.690–699.
- Sprain, L. and Reinig, L., 2018. Citizens speaking as experts: Expertise discourse in deliberative forums. *Environmental Communication*, 12(3), pp.357–369.
- Stagl, S., 2007. *SDRN Rapid Research and Evidence Review on Emerging Methods for Sustainability Valuation and Appraisal*. University of Sussex, Brighton.
- Steenbergen, M.R., Bächtiger, A., Spörndli, M. and Steiner, J., 2003. Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1, pp.21–48.
- Stern, P.C., 2005. Deliberative methods for understanding environmental systems. *AIBS Bulletin*, 55(11), pp.976–982.
- Stromer-Galley, J., 2007. Measuring deliberation's content: A coding scheme. *Journal of Public Deliberation* 3(1).
- Sunstein, C.R., 2002. The law of group polarization. *Journal of Political Philosophy*, 10(2), pp.175–195.
- Sunstein, C.R., 2005. Group judgments: Statistical means, deliberation, and information markets. *NYU Law Review*, 80, pp.962–1049.
- Sunstein, C.R. and Hastie, R., 2015. *Wiser: Getting Beyond Groupthink to Make Groups Smarter*. Harvard Business Press: Boston, MA.
- Szabó, Z., 2011. Reducing protest responses by deliberative monetary valuation: Improving the validity of biodiversity valuation. *Ecological Economics*, 72, pp.37–44.
- TEEB. 2010. *The Economics of Ecosystems and Biodiversity: Mainstreaming the Economics of Nature: A Synthesis of the Approach, Conclusions and Recommendations of TEEB*. Geneva: TEEB.
- Tillman, J.J., 2016. *An Integrative Model of Moral Deliberation*. Palgrave Macmillan, London.
- Tisdell, C., Wilson, C. and Nantha, H.S., 2008. Contingent valuation as a dynamic process. *The Journal of Socio-Economics*, 37(4), pp.1443–1458.
- Tversky, A. and Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*, 211(4481), pp.453–458.
- Urama, K.C. and Hodge, I., 2006. Participatory environmental education and willingness to pay for river basin management: empirical evidence from Nigeria. *Land Economics*, 82(4), pp.542–561.
- Vargas, A. and Díaz, D., 2017. Going along with the crowd? The importance of group effects for environmental deliberative monetary valuation. *Cuadernos de Economía*, 36(70), pp.75–94.
- Vargas, A., Lo, A., Howes, M., Rohde, N., 2017. The problem of inclusion in deliberative environmental valuation. *Environmental Values*, 26, pp.157–176.
- Vargas, A., Lo, A.Y., Rohde, N., Howes and M., 2016. Background inequality and differential participation in deliberative valuation: Lessons from small-group discussions on forest conservation in Colombia. *Ecological Economics*, 129, pp.104–111.
- Vatn, A., 2009. An institutional analysis of methods for environmental appraisal. *Ecological Economics*, 68, pp.2207–2215.
- Völker, M. and Lienhoop, N., 2016. Exploring group dynamics in deliberative choice experiments. *Ecological Economics*, 123, pp.57–67.

- Ward, H., Norval, A., Landman, T. and Pretty, J., 2003. Open citizens' juries and the politics of sustainability. *Political Studies*, 51(2), pp.282–299.
- Watkins, C., Massey, D., Brooks, J., Ross, K. and Zellner, M., 2013. Understanding the mechanisms of collective decision making in ecological restoration: an agent-based model of actors and organizations. *Ecology and Society*, 18(2): 32.
- Wilson, M.A. and Howarth, R.B., 2002. Discourse-based valuation of ecosystem services: establishing fair outcomes through group deliberation. *Ecological Economics*, 41(3), pp.431–443.
- Wright, S. and Street, J., 2007. Democracy, deliberation and design: the case of online discussion forums. *New media & society*, 9(5), pp.849-869.
- Zawojka, E. and Czajkowski, M., 2017. Re-examining empirical evidence on stated preferences: importance of incentive compatibility. *Journal of Environmental Economics and Policy*, 6(4), pp.374-403.
- Zhang, W., 2015. Perceived procedural fairness in deliberation: Predictors and effects. *Communication Research*, 42(3), pp.345–364.

Table 1. Debriefing questions for deliberative quality check (Murphy et al. 2017)

Question	Factor studied	Relevance for Value Juries (VJ) / Valuation Workshops (VW)
How much did your opinions about the relative importance of ecosystem services change over the course of the workshop?	self-perceived preference change	both
How well do you feel that your opinion was heard during the group deliberation?	being heard	both
How influential were you on the outcome of the group deliberation?	own influence within group	VJ
How influential were the scientists on the outcome of the group deliberation?	moderator or scientist influence (bias)	both
How satisfied were you with the outcome of the group deliberation?	general satisfaction with result	VJ

Text box 1. List of reporting requirements

Design – logistics & deliberative format:

- Sampling: number of participants, background, type of representation,
- Purpose of participation in the deliberative process
- Group composition, group size
- Participants' ability to participate in communicative rational debate
- Length of process, duration of each activity, number of activities
- Type of deliberative activities
- Location, setting of deliberative activities, access
- Remuneration (if applicable)
- Moderation: experience, impartiality, instructions
- Method of data collection: audio, transcripts, observation, questionnaires, etc.
- Process model used to reach consensus (only for Value Juries)

Design – information:

- Information provided and materials used in deliberative activities
- Media used for information provision, access to information
- Experts consulted or invited to provide input to deliberation
- Frames of environmental issue provided
- Discussion guide: a priori formulated questions, topics
- Positionality of the researcher(s) – in relation to the participants and the topic, and the design of the study (questions, information, analysis)

Design – pre-test and pilot:

- Methods, time, location, repetitions, sample
- Adjustments to deliberative process design following pre-test(s)

Analysis:

- Method of data analysis of deliberation
- Themes in content analysis
- Aggregation rule and welfare function
- Explanatory power of WTP models in relation to sample size

Results

- Response and participation rate, achieved representation
- Protest behaviour, quality or lack of engagement, fatigue
- Level of understanding of information and process
- Emerging topics in deliberation, e.g. identification of additional well-being effects
- Deliberation quality
 - group effects, dominance, polarisation
 - incentive compatibility, strategic behaviour
- Changes in preferences, motivations, attitudes
- Level of agreement or consensus, especially in Value Juries' stated WTP or choices
- Changes in choices or WTP statements that are attributed to deliberation