European Journal of Operational Research 299 (2022) 589–599

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Cognitive change and consensus forming in facilitated modelling: A comparison of experienced and observed outcomes

Vincent de Gooyert[a,*], Etiënne Rouwette[a], Hans van Kranenburg[a], Edward Freeman[b], Harry van Breen[c]

[a] Institute for Management Research, Radboud University, P.O. Box 9108, 6500, HK, Nijmegen, The Netherlands
[b] Darden School of Business, University of Virginia, P.O. Box 6550, Charlottesville, VA 22906, USA
[c] Alliander N.V., P.O. Box 50, 6920 AB, Duiven, The Netherlands

ARTICLE INFO

Article history:
Received 23 March 2020
Accepted 7 September 2021
Available online 15 September 2021

Keywords:
Behavioural OR
Facilitated modelling
Problem structuring
Group decisions and negotiations

ABSTRACT

Facilitated modelling is an approach where operational researchers act as facilitators to model an issue collaboratively with stakeholders. The potential benefits of facilitated modelling, including cognitive change and consensus forming, have been extensively described. However, the evidence that these benefits are achieved often remains anecdotical, and mainly relies on self-reported assessments. This is problematic, as self-reported measures can prove unreliable. In this study we adopted a pre-test post-test design to analyse whether the intended outcomes of eight facilitated modelling workshops were achieved. We compared the outcomes as experienced by participants of the workshops with observed outcomes. Our results show that cognitive change and consensus forming were achieved. However, experienced and observed cognitive change were not related. For consensus forming, we did find a modest correlation between experienced and observed outcomes. We discuss implications for facilitated modelling research and practice.

1. Introduction

Facilitated modelling is an OR tradition that revolves around operational researchers working together with stakeholders (Franco & Montibeller, 2010; Franco & Rouwette, 2011). In facilitated modelling, operational researchers intervene in a situation to help stakeholders address a problem by acting as modellers and facilitators of the group process. The potential benefits of facilitated modelling are widely described, and include cognitive change and consensus forming. Cognitive change is the idea that participants of a facilitated modelling workshops come in with a certain worldview, but that the intervention leads them to learn about the issue and accordingly "change their minds" (Ackermann & Eden, 2011). The intention of facilitated modelling is that the participants change their minds in such a manner that after the workshop their view on the problem is more similar to those of the other participants than before the workshop. In other words: facilitated modelling helps to form consensus (Rouwette, 2011). However, whether these intended outcomes of facilitated modelling are also achieved in practice is often not studied, and if evidence is presented this often remains anecdotical and relies on the assessment of outcomes by the participants of the workshop (Franco & Montibeller, 2010). The same holds for research on workshops in the social sciences more generally. Although workshops received considerable scholarly attention in the last decades (Hodgkinson et al., 2006; Jarzabkowski and Paul Spee, 2009; Vaara & Whittington, 2012; Whittington, 2006), knowledge is lacking on the impact that these workshops have. This is relevant, as studies on workshops in general report many cases in which such interventions failed to have an impact (C. Bowman, 1995; G. Bowman, 2016; Healey et al., 2015; Hodgkinson & Wright, 2002; Jarzabkowski & Seidl, 2008, Johnson, 2007).

In this study we distinguish between two ways of measuring impact: self-reports based on experiences by the participants of the workshop, and measures as observed through the comparison of pre-test and post-test questionnaires. Experimental studies indicated that a considerable gap may exist between the experienced effectiveness of dealing with a task, and the effectiveness as observed on the basis of empirical measures (Nisbett & Wilson, 1977). This mismatch between experienced and observed

* Corresponding author.
E-mail addresses: vincent.degooyert@ru.nl (V. de Gooyert), etienne.rouwette@ru.nl (E. Rouwette), hans.vankranenburg@ru.nl (H. van Kranenburg), freemane@darden.virginia.edu (E. Freeman), harry.van.breen@alliander.com (H. van Breen).

https://doi.org/10.1016/j.ejor.2021.09.007

effectiveness has also been found at the group level. After having participated in individual and group brainstorming sessions, participants are convinced that they are more productive when working as a group compared to working individually, while the opposite is true. This erroneous belief is known as the 'illusion of productivity' (Stroebe et al., 2010). While brainstorming refers to the divergent, creative phase of group decision making, overly optimistic estimation of group efficacy extends to convergence and argumentation. De Dreu & Beersma (2010) found that group members' confidence in their team's performance predicted decision quality for predictable and straightforward tasks, but not when task ambiguity was high. We explore whether a similar gap between experienced and observed outcomes exists in facilitated modelling workshops. Accordingly, our research questions are: "what is the impact of facilitated modelling on cognitive change and consensus forming?" and "what is the relation between experienced and observed outcomes of facilitated modelling?" This is in line with the behavioural OR (BOR) research agenda, which calls for a focus on actors, their actions and interactions (Franco et al., 2021). As Franco et al. (2021) note in their recent review of the BOR field, empirical research on the impact of facilitated modelling is scarce. With this study, we aim to contribute to the behavioural OR literature by examining facilitated modelling workshops and their outcomes in a field setting. We distinguish how the workshops are experienced by the participating stakeholders from the effectiveness of the workshops in terms of observed outcomes. By doing so we answer the call for more empirical studies that study the practice of OR (Franco & Greiffenhagen, 2018; White, 2016).

2. The impact of facilitated modelling

In this section we draw from the OR and social scientific literature on meetings to put the intended benefits of facilitated modelling in perspective. We first look at effectiveness studies published in the OR literature. Since studies in this domain do not provide a complete picture of intervention effects, we turn to the wider research on strategy meetings. Although we realize that many strategy meetings do not use any form of facilitated modelling, and that not all facilitated modelling workshops deal with strategic issues, we found parallels that help theorize whether and why facilitated modelling workshops are effective or not.

The recent review of BOR by Franco et al. (2021) groups behavioural studies into four groups based on two dimensions. The first dimension concerns the nature of behaviour in a given study: is it assumed to be affected by external factors such as the intervention or cognitive structures (determinist) or it is assumed that individuals have a degree of control over their own behaviour (voluntarist). The second dimension covers research methodology, which can be either aimed at testing relations between independent and dependent variables (variance) or at identifying developments over time (process). Our study is concerned with the effectiveness of facilitated modelling and therefore follows a determinist and variance approach. Among this type of studies, the review of Franco et al. included only two that address impacts of facilitated modelling which both report modest results (Joldersma & Roelofs, 2004; Cunha & Morais, 2016). While the review of Franco et al. (2020) draws on mainstream OR journals, relevant papers found in other journals offer additional insights. McCardle-Keurentjes & Rouwette (2018) also report modest effects. Rouwette (2016) summarizes research into behavioural effects of facilitated modelling in the system dynamics tradition (group model building). Two reviews cover over 130 empirical studies (Rouwette, 2002; Scott et al., 2015) and identify a range of outcomes such as communication, learning, consensus, behavioural change and implementation. The majority of these studies is based on self-assessment of outcomes but Scott et al. (2015) note a shift to more objective measures after 2001. However, for the four field experiments that assessed effectiveness of group model building relative to other conditions, it is not clear whether groups were comparable. Beyond the determinist and variance approach, there are several other studies that look at effects of facilitated modelling while also assessing mechanisms responsible for observed changes (Black and Andersen, 2012; Franco, 2013; Henao and Franco, 2016; Scott, Cavana and Cameron, 2016). A key mechanism these studies point to is the role of models as potential boundary objects: shared visual presentations of dependencies that each participant can change. Post workshop interviews are a key source of data, but as Eden (1995) notes recollections of an intervention may be biased as memory may be reconstructed to enhance once own's position or efficacy, due to limitations on recall or to wishful thinking. These studies also offer only limited evidence as they either report on single cases or use empirical data for illustrative purposes only. With Franco et al. (2021), we can conclude that at the moment there is no firm evidence of the effect of facilitated modelling on outcomes. It is therefore worthwhile to look into the generic research on management meetings.

Empirical studies into meetings more broadly and their effectiveness have generated mixed evidence. Bowman (1995) finds little evidence of impact after 40 workshops with top management teams. Some workshops feature a consultancy-led analytical process which results in a sound strategy that is however not owned or implemented by the management team. The opposite situation, commitment to a poor strategy, is also found. Bowman (1995) blames poor outcomes in part on the dual nature of many strategy meetings: an analytic process geared to developing a strategy and at the same time an occasion for team building. The latter leads to a tendency to avoid uncomfortable debates, which stops participants from surfacing and questioning core cognitions that constrain the strategy formation process. One of the theoretical foundations of facilitated modelling is that workshops induce participants to surface and share their knowledge and work out resolutions that take various positions into account (Eden & Ackermann, 2010; Franco et al., 2016). A core idea in facilitated modelling is that no single expert is able to address an issue individually because many situations are so complex that the knowledge of different stakeholders needs to be combined. Apparently, information sharing cannot be taken for granted because alternative agendas including maintaining personal relationships may lead participants to refrain from sharing and discussing sensitive issues.

In their survey of 650 strategy meetings, Healey et al. (2015) separate organizational, interpersonal and cognitive outcomes. They find little evidence for major organizational change but some effects on interpersonal relations and strategic understanding. Four design characteristics explain outcomes: clarity of workshop objectives impacts all three outcomes positively with workshops undertaken for purpose of implementation (rather than formulation) more likely to impact organizational outcomes. Detachment of a meeting from organizational practices reduces the impact on organizational outcomes. Wider participation is associated positively with interpersonal outcomes, and cognitive effort (preparation, time spent in meetings, use of challenging analytical techniques) drives strategic understanding. In contrast to the three studies discussed above, two studies by Hodgkinson et al. (2006) and Jarzabkowski & Seidl (2008) do find some evidence for organizational impact. Hodgkinson et al. (2006) conduct a large-scale survey among UK managers. On the basis of 1,337 returned questionnaires, the authors conclude that about one in three strategic meetings focusses on strategy formulation, and more than half combine strategy development and implementation. Other aims include organizational development such as teambuilding and changes in organizational structure. Hodgkinson et

al. (2006) find that there is little preparation nor use of analytical tools in these meetings. Participants are mostly senior management. Perhaps understandably, meetings are found to improve working relationships with peers but not so much with staff in other positions or external stakeholders. Meetings are used to identify and share emergent strategies. Cognitive findings are reported in about half of the meetings: a better understanding of business processes and corporate values. *'However, there is less evidence that strategy workshops lead to "harder" outputs such as increased levels of innovation, the use of technology, and/or enhanced productivity or profitability, where there is a greater reporting of "no impact", or even negative impact in a few cases'* (Hodgkinson et al., 2006: 486).

Jarzabkowski & Seidl (2008) draw on a dataset of 51 university meeting observations. They find evidence of relations between meeting practices, cognitive outcomes and real-world impact. Four types of discussion are found: administrative discussions on non-contentious strategic topics, restricted meetings (in terms of turn-taking), restricted free discussion (allowing participants to speak without going via the chair), and free discussion. Discussion type has a relation with organizational impact: free discussion has the highest likelihood of generating variations and changing strategies.

The studies surveyed so far offer some interesting parallels. Supporting Mintzberg's (1994) idea that future-oriented strategic thinking is needed to set strategic direction and create commitment, open discussion is found to foster implementation (Jarzabkowski & Seidl, 2008). This is in line with Bowman's (1995) finding that analysis without involvement does not lead to commitment. Debate on sensitive issues is needed to uncover core assumptions and work towards fundamental strategic change. Complementary to strategic thinking, Mintzberg (1994) saw a role for strategic planning in operationalization and coordination of strategies. Again, there is some support for this view. In-depth analysis of current organizational structures seems to have a role in realizing organizational outcomes, possibly via generating understanding (Kaplan, 2011). Strategy meetings on implementation (rather than formulation) are more likely to impact organizational outcomes, and detachment from organizational practices reduces the chance of impacting them (Healey et al., 2015). Cognitive effort drives strategic understanding (Healey et al., 2015) and limited analysis leads to cognitive impact in only half of the cases (Hodgkinson et al., 2006). The conclusion seems to be that informal strategic thinking and more analytical strategic planning both have a role to play in strategic change, but how they interact exactly is still an open question

3. Facilitated modelling effectiveness: cognitive change and consensus forming

As seen above, workshops can have various intended outcomes. In the remainder of this article we focus on two of these outcomes: cognitive change and consensus forming. We selected these two outcomes for three reasons: together they cover an important part of the range of outcomes that workshops can have (Hodgkinson et al., 2006), they are central aims of facilitated modelling (Franco & Montibeller, 2010; Rouwette et al., 2009) and because these two outcomes can be measured around facilitated modelling in the field, on both an experienced and observed level (Markóczy & Goldberg, 1995). We elaborate on both types of effectiveness below.

3.1. Cognitive change

Facilitated modelling provides an opportunity for workshop participants to learn about strategic issues (Healey et al., 2015). Each participant in a strategy workshop brings his or her mental model about the organization and its environment Kaplan (2011). This mental model consists of all the deeply held beliefs that individuals maintain (Walsh, 1995). Engaging in a structured dialogue can lead to instances where opposing mental models are confronted with each other. If the participants manage to surface and resolve these differences, this constructive conflict (Amason, 1996; Weick, 1995) leads to an improved understanding of strategic issues, which we call 'cognitive change'. Related notions can be found in the literature, including mental model updating (Uitdewilligen et al., 2013), mental model transformation, and belief revision Chi (2009).

Cognitive change has been shown to be crucial for organizations operating in a changing environment (Nadkarni and Narayanan, 2007; Narayanan et al., 2011). The opposite of cognitive change, cognitive inertia, has been associated with organizational downturn (Hodgkinson, 1997; Tripsas and Gavetti, 2000). Hodgkinson (1997) showed, drawing on a repeated questionnaire returned by 114 respondents from 41 firms in the UK real estate industry, that mental models can remain stable even though circumstances are changing substantially. According to Hodgkinson (1997), this implies that strategists should frequently engage in process of reflection to reconsider assumptions and beliefs (Hodgkinson, 1997, p. 940). Facilitated modelling can be seen as a platform for such a process. Barr et al. (1992) stress that "[o]rganizational renewal requires that a firm's top managers make timely adjustments in their mental models following significant changes in the environment" (Barr et al., 1992, p. 15). Kaplan shows that mental models, or frames, are contested in meetings, and that strategic decisions are a product of the mental models that have become dominant through such meetings (Kaplan, 2008). In this light, facilitated modelling provides an occasion where adjustments to mental models can be made, and it is highly relevant whether participants in facilitated modelling are able to assess whether they achieved cognitive change or not. However, to the best of our knowledge, we are the first to test whether participant's assessment of cognitive change correlates with observed measures of cognitive change or not.

3.2. Consensus forming

Meetings are occasions in which three or more people meet for a purpose related to the organizational or group function, such as coordinating different organizational perspectives and agendas. Meetings help to sustain the unity of the organization by socially validating the current order and serving as a place in which participants exchange interpretations or make sense of their situation (Jarzabkowski & Seidl, 2008). Within a meeting, a participant may use her speaking turn to offer an understanding of current practices or a future desired state of affairs. In doing so, the participants select a specific topic from the ongoing flow of events and invites others to provide their interpretation. Haug (2015) sees the testing and social validation of interpretations as one of the core processes in meetings. If participants indicate agreement on a proposition, this constitutes interactional consensus. Agreement can be determined on the basis of absence of opposition and is more or less explicit. It may take the form of the chair claiming consensus has been achieved and then moving on to another topic. Alternatively, the chair may provide an opportunity to express dissent and make sure that all participants have a chance to articulate their views. Interactional consensus is important as participants may make contributions that directly bear on their intended actions in the future. Interactional consensus is *'about reducing the uncertainty about the participants' future actions by jointly accepting a position that obliges them to act accordingly (…) and thereby coordinate collective action'* (Haug, 2015: 562). However, topics

discussed may also relate to actions in a more indirect fashion. Implementation of strategy *'requires more than a simple action plan (…). It also requires a shared grasp of the logic behind the plan. Typically, strategic decisions are not articulated in great detail, and unforeseen issues arise as events unfold (…). Details must be settled and issues resolved in a way that is consistent with the intention behind the plan'* (Kellermanns et al., 2005: 722). Thus, implied agreement on for instance the current status of the organization, or on values, may also but more indirectly constrain intended actions of meeting participants.

A separate form of consensus, more commonly studied in the literature, is mental consensus (e.g. Markóczy, 2001). This type of consensus refers to congruence or alignment of beliefs among members of an organization. Beliefs may for instance refer to priority of goals and means, competitors, or organizational strengths and weaknesses (Markóczy, 2001). Haug (2015) notes that mental consensus refers to mental states and is inferred from individual behaviour, as for instance captured in surveys. While interactional consensus is observed by meeting participants, mental consensus is observed by an outsider (e.g. a researcher) by aggregating individual measures. Franco & Greiffenhagen (2018) refer to interactional consensus when they note that participants in facilitated modelling need to infer consensus from other participants' behaviour. A line of studies relates mental consensus to organizational performance, and a review by Kellermanns et al. (2011) finds a positive and significant relationship. Interestingly, Kellermanns et al. (2011) note that studies on mental consensus often do not account for the process through which consensus was reached; they typically assume a collective appreciation of the reasons behind the strategy and a common awareness of the intended action. However, Haug (2015) finds that meeting participants may be unaware of the degree of mental consensus they have achieved. Scheff (1967) measures both individual opinions as well as what participants think most others think. In case of a mismatch between mental consensus (congruence of individual opinions) and perceived consensus, he refers to pluralistic ignorance if the majority agrees but thinks they do not. False consensus refers to the opposite: participants think they agree but in fact they do not.

We propose that facilitated modelling runs the risk of generating false consensus: the participants believe that they have achieved consensus while in practice, according to more objective measures, this is not the case. Participants exchange their views on strategic goals, the current organizational structure and which actions to prioritize, and assume that everyone attending has developed an understanding similar to theirs. Similar to strategic decisions which are typically not articulated in great detail (Kellermanns et al., 2005), the rationale for decisions also runs the risk of remaining at a conceptual and abstract level. Given the fact that meetings address complex, ambiguous issues and attending managers often have different functional backgrounds, there is a clear danger that misunderstandings arise but go unnoticed (Cronin & Weingart, 2007). In such a case the group may think they are in agreement, but in fact they are not. For true congruence between strategic beliefs to emerge, an open exchange of views, systematic processing of information and (interactive) consensus testing seem necessary. Research in social psychology points to two reasons why this form of discussion is crucial. First, Nisbett & Wilson (1977) show that people lack insight into whether a change in understanding occurred and what caused it. This means that meeting participants may have changed their opinion on the basis of information discussed, without being aware of any learning or change in preferences. In other words, the complex topics addressed in strategy meetings lead to a distinct possibility that different participants pick up different things from one and the same piece of information. Participants may however not directly be able to articulate what they picked up and how it changed their views. Second, in group meetings people erroneously feel they were very productive in generating ideas (Stroebe et al., 2010) or feel unrealistically confident about the quality of their decisions, in particular in a complex task (de Dreu & Beersma, 2010). This may lead to the situation that important aspects of the problem are overlooked and information that is exchanged is not adequately processed. A fundamental claim of facilitated modelling is that it helps avoid generating false consensus because the facilitators and the modelling efforts provide the structure that support the systematic processing of information necessary to avoid false consensus. However, also here, to the best of our knowledge, we are the first to test this premise in a field setting.

4. Method

4.1. Case description

To study facilitated modelling in the field, we performed an in-depth case study of an organization that engaged in facilitated modelling with the aim of both arriving at cognitive change and consensus forming, carried out in 2013. The context is the Dutch energy sector, where prevailing strategies were under scrutiny because then existing beliefs were confronted with dissonant information. The Dutch government derived a substantial part of its income from exploiting a large gas reserve (Kern & Smith, 2008) and the international fossil industry has a considerable representation in the Netherlands with the headquarters of the Anglo-Dutch multinational oil and gas company Royal Dutch Shell (Kemp et al., 2007). Despite its dependence on fossil energy the Dutch government was and still is increasingly implementing policies aimed at renewable energy production (Junginger et al., 2004; van Rooijen and van Wees, 2006). Back then, it had set new goals for energy conservation and renewable energy production as a part of its 2012 coalition agreement (Cabinet Rutte-Asscher, 2012). The aims implied considerable changes, with a goal for renewable energy production of 16% in 2020, compared to 4.7% in 2013 (PBL, 2012). Because of the necessary changes the government coined the term 'energy transition' (Kemp, 2010), i.e. a structural change towards a more sustainable energy system (Smith & Kern, 2009). The existing design of the Dutch energy system conflicted with the set goals and different parties were trying to make sense of what the energy transition exactly is (Hendriks, 2009: 346; Kemp et al., 2007: 316).

The organization that we follow in our study strives to play a leading role in facilitating the Dutch energy transition: the distribution system operator Alliander. Alliander is the largest Dutch distribution system operator in terms of its workforce with 5,866 full-time employees at the end of 2014. It transported 29,936 gigawatt-hours of electricity and 6,115 million cubic meters of natural gas in 2014, had a revenue of 1,696 million euros, and profits reaching 323 million euros. Its annual report around that time stated that "Alliander wants to facilitate the transition to a more sustainable energy system", and that "while the contours of the future energy system are already visible, the exact shape of things to come is still unknown". Therefore, "by working closely with customers, partners and government agencies, [they] are trying to anticipate trends and developments, wherever possible, in a responsible manner" (Alliander, 2012). Below, we follow how Alliander engages in facilitated modelling to support arriving at a shared understanding (consensus forming) of the ongoing energy transition and supporting cognitive change to try to avoid holding on to deeply held beliefs that have been rendered obsolete by the changing circumstances in the sector.

Table 1
Participant background in the eight strategy workshops.

Workshop	1	2	3	4	5	6	7	8	Total
Infrastructure	5	1	1	5	6	5	2	2	27
Services	3	3	8	1	4	3	3	0	25
Industry	3	3	2	0	2	3	2	3	18
Government	1	0	0	6	1	0	0	9	17
Other	0	2	0	0	1	1	2	3	9
Total	12	9	11	12	14	12	9	17	96

4.2. Data collection

The data consists of eight facilitated modelling workshops, carried out in September and October of 2013. The number of participants per workshop varied from eight to fifteen, with a total of 96 participants over all eight workshops, representing 19 internal stakeholders and 77 external stakeholders. The external stakeholders represented different groups: 25 represented services (consultants, bankers, lawyers), 18 industry (manufacturing, energy supply), 17 government (national, local), 8 infrastructure (transmission, distribution), and 9 other (research, NGO's). Table 1 below shows how both internal and external stakeholders were distributed over the eight workshops. In each workshop two facilitators led the dialogue on trends and developments in the energy industry. The workshops took about five hours each. The content of the workshops is discussed in more detail in de Gooyert et al., 2016.

4.3. Mental models of the energy transition

To analyse the effectiveness of the facilitated modelling workshops, we followed the procedure as put forward by Hodgkinson (2002, p. 68) and Markóczy and Goldberg (1995, p. 309): we develop a 'pool of constructs' and we use these constructs to define closed format questions with which we elicit the mental models of the participants, allowing us to measure cognitive change and consensus forming. This procedure combines the strengths of ideographic and nomothetic methods while minimizing their weaknesses (Daniels et al., 1994, 2002; Hodgkinson, 2002: 70; Liu et al., 2012). We developed a pool of constructs on the energy transition by studying newspaper articles that mention the term energy transition. The constructs found in the newspapers subsequently formed the basis for our structured questionnaires. We use Likert-type items to assess participants' interpretations, as previously done by for example Sutcliffe & Huber (1998). We asked the participants to fill out the questionnaire right before and right after the workshop, following Liu, Friedman, Barry, Gelfand, & Zhang (2012).

The pool of constructs on the energy transition is based on articles in the five largest national daily Dutch newspapers. We included all 162 articles published before 21 June 2013 (the first article mentioning "energy transition" appeared in 2003), see Table 2 below. We consecutively performed one round of open coding and one round of selective coding (Glaser & Strauss, 1967). The round of open coding was used to analyse the categories of constructs related to energy transition. The round of selective coding was used to analyse which constructs within these categories can be found.

The first round of coding resulted in three distinct categories: strategic ends that may be served by the energy transition, technological means to bring about a more sustainable energy system (e.g. solar panels, wind energy) and policy means that could be implemented by different governmental bodies to speed up the transition (e.g. carbon taxation, subsidies on innovation). The second round of selective coding started from this categorization and produced a list of constructs in these three categories. We excluded items that were mentioned only once or twice. Next, two experts from the energy sector assessed the face validity of the items during interviews. Based on the feedback of these experts we made minor changes, and this resulted in the final list of 47 items that is shown in Appendix A. The 47 items are all included in the questionnaire used for assessing the mental models of the strategy workshop participants. For each item the question was "how important is [this item] for you with regard to the energy transition?" The closed format questions were formatted as seven point Likert items, ranging from not important at all to very important. The complete questionnaire can be obtained from the authors.

4.4. Observed measures

We measure observed cognitive change by comparing the questionnaires before and after the workshop on the individual level. Cognitive change is calculated for each participant as the Euclidean distance between the answers to the seven point Likert items (i) of the questionnaires before (p) and after (q) the workshop, see the formula below (Kellermanns et al., 2005).

$$Cognitive\ change = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}$$

We measure consensus forming by comparing the answers to the closed questionnaires between the participants of the strategy workshops (based on the 86 out of 96 managers that filled out a complete pre and post questionnaire). Observed consensus forming is measured on the level of the workshops. Following Knight et al. (1999), we measure dissensus by summing up, for each workshop separately, the standard deviations of the answers on each of the 47 constructs. The higher the number, the more disagreement there is on the importance of aspects in the energy transition. Consensus *forming* then is measured by subtracting dissensus *after* the workshop from dissensus *before* the workshop.

$$Consensus\ forming = \sum_{i=1}^{n} \sigma(p_i) - \sigma(q_i)$$

4.5. Experienced measures

We measure experienced cognitive change and consensus forming using scales originally developed by Vennix, Scheper and Willems (Vennix, Scheper & Willems, 1993; Vennix & Rouwette, 2000; Rouwette, 2011; Vennix, 2016). All items are measured on a five-point Likert scale anchored by strongly disagree—strongly agree. The scale for cognitive change is based on five Likert items. Alpha reliability in an earlier study based on 49 participants (Vennix & Rouwette, 2000) is .77, alpha reliability in the current study ($n = 96$) is .73. Items address the extent to which the strategy workshop has increased insight, has led to more clarity on the interrelations between elements that play a role in the problem, has not given insight into the causes of the problem (reverse coded), has given me more insight into the feedback processes that play a role in the problem, has not given me insight into the possibilities that my organization has in managing the problem (reverse coded).

Experienced consensus forming is measured using a scale of four items. Alpha reliabilities in earlier studies were .84 (Vennix & Rouwette, 2000, $n = 55$, based on three items), .60 (Rouwette, 2011, $n = 21$, four items). Alpha reliability in the current study ($n = 96$) is .60. The items cover degree of shared vision, degree to which the workshop represented an integration of opinions, whether the group did not reach a consensus (reverse coded), and closeness of opinions.

Table 2
Articles in five large Dutch newspapers that mention "energy transition".

Newspaper	'03	'04	'05	'06	'07	'08	'09	'10	'11	'12	'13[a]	Total
AD	0	0	0	1	3	0	0	0	0	0	1	5
NRC	1	0	3	8	16	4	4	6	2	6	8	58
Telegraaf	0	0	0	1	5	2	1	3	2	3	3	20
Trouw	0	0	1	9	8	3	4	10	4	4	3	46
Volkskrant	0	0	0	5	8	5	1	9	3	2	0	33
Total	1	0	4	24	40	14	10	28	11	15	15	162

[a] up to 21 June 2013

Table 3
Observed consensus forming in facilitated modelling workshops.

Workshop	Participants	Dissensus pre	Dissensus post	Consensus forming
1	10	49,5	42,8	6,7***
2	9	55,7	48,0	7,7**
3	10	54,3	46,2	8,0***
4	9	49,2	46,0	3,2
5	14	55,5	52,3	3,2*
6	12	50,8	49,7	1,0
7	8	57,6	54,0	3,7*
8	14	56,0	52,1	3,9*

* $p < .05$,
** $p < .01$,
*** $p < .001$

5. Results

Below, we first describe the observed outcomes of the workshops in terms of consensus forming and cognitive change, then the experienced outcomes. Consequently, we describe the relation between experienced and observed measures of these outcomes.

5.1. Observed measures of cognitive change and consensus forming

Individual differences in observed cognitive change were considerable. The Euclidean distance between before and after the workshop per participant over all 47 items ranged between 3.3 and 9.6 (with an average of 6.2 and a standard deviation of 1.5). As can be seen in Table 3 below, there was a difference in the extent to which the eight facilitated modelling workshops were successful in consensus forming, according to observed measures. Based on observed measures, three of the workshops clearly resulted in higher consensus after the strategy workshop compared with before ($p < .001$), for three workshop there is some indication that consensus was formed ($p < .05$), and two workshops did not significantly result in more consensus.

5.2. Experienced measures of cognitive change and consensus forming

Our construct for experienced cognitive change proves to be reliable (Cronbach's alpha = .73). The mean value for experienced cognitive change is 3.56 ($n = 93$, standard deviation = .57). To test whether the participants in the workshops perceived facilitated modelling overall as effective in terms of cognitive change, we performed a one-sample t-test with test value 3 (the neutral answer option 'neither agree nor disagree'). This test shows that participants scored cognitive change significantly higher than 3 ($p < .000$).

Our construct for experienced consensus forming proves to be moderately reliable (Cronbach's alpha = .60). The mean value for subjective consensus forming is 3.59 ($n = 92$, standard deviation = .51). To test whether the participants in the workshops perceived facilitated modelling overall as effective in terms of consensus forming, we performed a one-sample t-test with test value 3 (the neutral answer option 'neither agree nor disagree'). This test shows that participants scored consensus forming significantly higher than 3 ($p < .000$).

5.3. Correlations between observed and experienced measures of cognitive change and consensus forming

To test whether experiences of workshop participants are in line with observed measures of workshop effectiveness in terms of cognitive change we calculated the Pearson correlation coefficient between experienced cognitive change and observed cognitive change. This showed that experienced cognitive change and observed cognitive change are not significantly correlated ($r = .08$, $n = 87$, $p = .225$ one-tailed).

To test whether the experiences of workshop participants are in line with observes measures of workshop effectiveness in terms of consensus forming we calculated the Pearson correlation coefficient between experienced consensus forming and observed consensus forming. This showed that experienced consensus forming and observed consensus forming are significantly correlated ($r = .25$, $n = 92$, $p = .008$ one-tailed). Although the correlation is significant, the extent of correlation is modest, so also here experienced and observed measures do not capture the same phenomenon. However, these results show that experienced measures of consensus forming provide an indication for observed measures of consensus forming (and vice versa), while experienced measures of cognitive change do not provide an indication for observed measures of cognitive change (and vice versa).

5.4. Robustness checks

To test the robustness of our findings, below we provide alternative measurements. In the above, we measured observed cognitive change using Euclidean distance, or the L2 norm. However, the use of this measurement is not without problems, especially because it requires considering Likert items as interval data rather than ordinal data, while we have no conclusive argument that equidistant bin lengths can be assumed. For the measures of observed consensus forming, perceived cognitive change and perceived consensus forming however this is less of a concern, because these measures aggregate individual Likert items

Table 4
Nonparametric test of observed consensus forming.

Workshop	1	2	3	4	5	6	7	8	Total
# from consensus to dissensus	3	5	5	4	6	5	5	2	35
# from dissensus to consensus	11	9	10	5	11	11	9	12	78
Binomial test p value	.03	.21	.15	.50	.17	.11	.21	.01	.00

into composites (Allen & Seaman, 2007; Brown, 2011; Maurer & Pierce, 1998).

In Appendix B we show for each item in the questionnaire whether we observe consensus or dissensus between the participants of one workshop using a coarser measurement: the interquartile range. To determine whether we observe consensus or not we use the procedure based on studies using panel data, which indicates that for seven-point Likert items an interquartile range of 1 or smaller signals consensus Von der Gracht (2012). Over all workshops we observed 37 items on which there was consensus beforehand but no consensus afterwards, versus 78 items on which there was no consensus beforehand but consensus afterwards. This gives a McNemar test statistic of 15.98, well above a critical value of 10.83 (χ^2 distribution, df = 1, $p = .001$). This provides strong support that indeed the facilitated modelling workshops helped with consensus forming in terms of observed measurements. From Appendix B it appears that most of the consensus forming took place on technological means and policy means. Consensus forming was observed on 48 technological means while a loss of consensus was observed on 13 technological means. Consensus forming was observed on 22 policy means while loss of consensus was observed on 13 policy means. This stands in contrast with strategic ends. Consensus forming was observed on 8 strategic ends while loss of consensus was observed on 9 strategic ends. A closer examination of the strategic ends shows that this finding may be caused by the high initial observed consensus on strategic ends. The participants had consensus on 80% of the strategic ends prior to participating in the strategy workshops, leaving little room for a further increase in observed consensus on strategic ends throughout the workshops.

As can be seen in Table 4 below, there was a difference in the extent to which the eight workshops were successful in forming consensus. The table can be read as follows: in workshop one there were three issues on which we observed consensus before the workshop, but not any longer after the workshop, and eleven issues on which we observed dissensus before the workshop, but consensus right after the workshop. These numbers result in a p value of .03. Workshop one and workshop eight show the strongest observed consensus forming, with a p value below .05. Workshops two, three, five, six and seven show intermediate results with about two times as many strategic issues on which consensus was formed as issues on which consensus was lost. Workshop four is the only workshop where realization of the espoused purpose is not observed at all. These findings point in the same direction as the observed measures of consensus forming based on sums of standard deviations (reported in Table 3), as that analysis also indicated that workshop 4 did not show any sign of consensus forming. A difference between the two analyses of observed consensus forming is that we now find evidence for observed consensus forming in workshop 6 in contrast to the analysis reported in Table 3.

6. Conclusion and discussion

Behavioural OR focuses on actors, their actions and interactions (Franco et al., 2021). Facilitated modelling is a tradition that extensively involves stakeholders in addressing complex issues (de Gooyert et al., 2017). However, the claims underlying facilitated modelling are widely described yet scarcely studied. We contribute to the Behavioural OR literature by testing both the observed and experienced effectiveness of facilitated modelling in terms of cognitive change and consensus forming. We found that overall our facilitated modelling workshops succeeded in bringing about cognitive change and consensus forming, supported by both observed and experienced measures of workshop effectiveness. However, on the level of individual workshops and participants, strong differences in effectiveness could be seen. We found that experienced outcomes of consensus forming were significantly correlated with observed outcomes of consensus forming. In other words, the observed measure of workshop effectiveness in terms of consensus forming gives an indication of experienced effectiveness, to some extent. However, we found that experienced outcomes of cognitive change were not significantly correlated with observed outcomes of cognitive change. In other words, participants who after the workshop argued that they did not change their mind on the topic being discussed, had about the same chance of actually having changed their mind according to observed measures as those participants who argued that they did change their mind.

Studies by Hodgkinson et al. (2006), Kaplan (2011) and Healey et al. (2015) find that formal analysis in workshops contributes to understanding. At the same time, a number of scholars indicate that workshops can fail to have substantial impact (C. Bowman, 1995; G. Bowman, 2016; Healey et al., 2015; Hodgkinson & Wright, 2002; Jarzabkowski & Seidl, 2008). An important implication of our study is that some of the most fundamental assumptions of facilitated modelling cannot be taken for granted. Facilitated modelling is based on the idea that involving participants in a structured way, using model-based and facilitated analysis, will help to elicit knowledge and resolve conflicting views. Our study showed that this is not necessarily the case, and that the effectiveness of workshops as experienced by participants does not necessarily match with observed measures of workshop effectiveness. Our finding adds to Eden's (1995) warning that recall of workshops may be biased and implies that more intensive evaluation of facilitated workshops is necessary to test whether the intended benefits were really achieved, as well as perceived, or not.

Earlier research indicates that effectiveness of workshops may be hindered if participants avoid discussing sensitive issues Bowman (1995). One way a facilitator may try to avoid this phenomenon is by preparing a workshop through individual interviews with the participants. If facilitated modelling workshops are preceded by a round of interviews where operational researchers ask individual stakeholders that are going to participate in the workshops about sensitive issues, this could play a role in helping to avoid the phenomenon where facilitated modelling workshops prove ineffective because participants steer away from the issues that matter most. This finding is also relevant to the research on mechanisms responsible for effectiveness of modelling workshops. Studies by Black and Andersen (2012), Franco (2013) and Scott, Cavana & Cameron (2016) highlighted the role of models as potential boundary objects and in particular their ability to be jointly created and modified and to transcend differences between participants in terms of language, meaning and interests. It may be the case that the model developed in the sessions has all of these characteristics, but does still not address the issues that matter most to the group of participants. This calls for adding another dimension in research on boundary objects: their substantive content.

The fact that experienced and observed measures correlate for consensus forming, but not for cognitive change, is intriguing. Our study design does not allow to infer why this is the case. Earlier studies do provide some suggestions where an explanation might be found. Eden (1995) already noted several sources of biases in forming recollections of an intervention. From research in social psychology it is known that self-assessment of learning can be considerably flawed (Nisbett & Wilson, 1977; De Dreu & Beersma, 2010). Consensus on strategic direction is formed on the basis of exchanging views and information concerning the focal organization and its environment. Views and information in turn again generate reactions by participants. Those present in a workshop witness this process and can apparently to some extent judge how this exchange swings opinions and fosters consensus Haug (2015). In contrast, judging the impact of a message on one's own opinion is a process of introspection which is difficult to human decision makers (Wilson, 2004).

An important methodological implication of our study is that the growing body of research on behavioural operational research should be very careful with using self-reported measures as proxies for outcomes. Studies often rely on self-reported measures of effectiveness (e.g. Lami & Tavella, 2019), and our findings suggest that these studies may only have an incomplete or biased image of what the outcomes of workshops really are. Facilitated modelling participants proved to be a not completely reliable source of information on the outcomes of the workshops. While Nisbett & Wilson (1977), Stroebe, Nijstad & Rietzschel, (2010), de Dreu & Beersma (2010) and Haug (2015) have found clear differences between experienced and observed perceptions of outcomes in controlled studies. When it comes to cognitive change, our study finds support for their findings in a field setting, and when it comes to consensus forming we did find a significant but limited correlation.

An important implication for practice is that it would be good to think twice about *why* facilitated modelling workshops are organized and whether the expected benefits will outweigh the costs. Our study suggests that workshop participants could suffer from an 'illusion of productivity' when considering productivity in terms of cognitive change: they might perceive the activity to be contributing to cognitive change although in terms of real-world impact little is achieved. If an illusion of productivity leads participants to think "participating in facilitated workshops is a waste of time" this may result in refusing to participate in later workshops, even if the workshop was in fact effective. And the other way around: if the workshops are ineffective, they may be perceived as effective and participants continue to waste time in workshops that have little impact. Cognitive change has been shown to be crucial for the success of an organization (Barr et al., 1992; Nadkarni & Narayanan, 2007; Narayanan et al., 2011). Our study shows that self-reported measures of cognitive change are not a reliable source to assess the effectiveness of facilitated modelling workshops in fostering cognitive change. Critically following the real-world effects of facilitated modelling workshops may help organizations in allocating scarce resources more effectively. Adopting more extensive empirical measurements for effectiveness using a pre-test post-test design, including cognitive change, might be one way of checking whether intended outcomes are achieved, although such measurements of course come with their own important disadvantages.

7. Limitations and future research

One limitation of our study that we confined our analysis to two forms of productivity: cognitive change and consensus forming. We showed that workshop participants are not a reliable source to assess workshop effectiveness, but this only holds when considering effectiveness in our narrowly defined term. Others have pointed out that participants can have other motivations to participate in workshops, including the more general goal of 'team building' (Bowman, 1995; Hodgkinson et al., 2006). Other outcomes of facilitated modelling workshops could be included in future research, for example by looking at more distal outcomes such as commitment to strategic change (Bowman, 1995), or the strengthening of interpersonal relations (Healey et al., 2015). Future research could adopt a pre-test post-test design covering a wider range of intended outcomes of facilitated modelling.

Another limitation of our study is that we have looked at one out of various settings in which facilitated modelling can take place. In Mintzberg's (1994) terms our workshops had the intention to foster strategic thinking in an ambiguous environment, rather than supporting planning in a relatively stable, less equivocal, environment. In that sense, the facilitated modelling workshops in our study resembled 'strategy workshops'. This is also why we connected to social scientific studies on strategy workshops (e.g. Healey et al., 2015; Jarzabkowski & Seidl, 2008, Johnson, 2007), while we acknowledge that many facilitated modelling workshops are not strategy workshops and many strategy workshops are not facilitated modelling workshops. Future research might investigate whether the findings of our study also hold for facilitated modelling workshops in different settings. In addition, there are different facilitated modelling traditions, future research could also investigate whether our results hold for different types of facilitated modelling workshops. Lastly, we only looked at rapid outcomes directly at the end of the workshop. Future research could also compare outcomes at more distal points in time.

As with all studies, our operationalisation has limitations. While many reports of facilitated modelling effectiveness rely on anecdotical evidence, we provided a more systematic comparison of both experienced and observed outcomes. However, even our 'more objective' observed measures are still subjective to some extent. The questionnaires that we have designed might have missed important cognitive change and consensus forming that participants were able to identify subjectively. This could provide an alternative explanation why for cognitive change we found no correlation between experienced and observed measures. We used Likert scales to assess experienced outcomes and although the accompanying Cronbach's alphas provide some signal of reliability, they are not very strong. We measured dissensus in workshops by adding standard deviations across questions that may not be independent, which can introduce double counting of dissensus information. However, our conclusions are only based on comparisons of pre- and post-workshops dissensus scores which reduces the potential double counting bias if we assume that this problem is equally present both before and after the workshop. We contrasted the findings of eight facilitated modelling workshops. We did not include any groups that did not participate in a facilitated modelling workshop. Therefore, we do not meet the standard of having a control group in our study design.

An important trend in behavioural OR is that more and more effort is put in unpacking the process of applying OR. Our study focussed on outcomes, we did not unpack the process of facilitated modelling itself. As such, we "do not reveal the events, activities or phases through which individual and group behaviour changes when engaging with OR methods, processes and tools" (Franco et al., 2021; see also Tavella & Franco, 2015). A mechanism element that recently has received considerable attention is the role of models as potential boundary objects. Our study indirectly points to the substantive content of boundary objects (in the main, their coverage of sensitive issues) as a relevant dimension to be addressed in future studies. We encourage future em-

pirical research on the effects of facilitated modelling to also include process data that allows studying the mechanisms that are responsible for the (lack of) effects, for example using ethnomethodology (Franco & Greiffenhagen, 2018).

Acknowledgements

Funding for this project was provided by Next Generation Infrastructures and Alliander (09.05.RAD). We thank all participants of the workshops for their contributions. We thank our colleagues at Radboud University, Birgit Fokkinga, Hubert Korzilius, Inge Bleijenbergh, and Marleen McCardle, for helping with facilitating the workshops. We thank Eric Jacobs and three anonymous reviewers for their valuable constructive comments.

Appendix A. Key issues in the energy transition (frequency between brackets)

Strategic ends
1. Affordability of the energy system (30)
2. Improving the job market by investing in the energy system (11)
3. Improving the Dutch natural environment (13)
4. Countering climate change (65)
5. Decreasing the dependability on other regions for energy (25)
6. Setting an example for other regions (4)
7. Honoring international agreements (6)
8. Securing possibilities of future generations to meet their needs (3)
9. Improving competitiveness by gaining a technological lead (16)

Technological means
10. Carbon capture and storage (5)
11. Hybrid vehicles (3)
12. Solar panels (9)
13. Shale gas (2)
14. Intelligent traffic management (4)
15. Concentrated solar power (4)
16. Coal plants (9)
17. Biofuels (7)
18. Geothermal energy (2)
19. Gas plants (2)
20. Electric vehicles (2)
21. Wind power on land (13)
22. Nuclear energy (5)
23. Conservation in the industry (5)
24. Wind power on sea (3)
25. Biomass (5)
26. Conservation in buildings (19)
27. Combined heat and power (2)
28. Investing in electricity grids (6)
29. Using waste heat (3)
30. Energy storage (4)
31. Smart metering (3)
32. Aquifer storage and recovery (3)
33. Hydrogen (9)
34. Blue energy (3)

Policy means
35. Subsidies for sustainable energy production (24)
36. Stricter regulations in the form of norms and obligations (22)
37. Decrease the frequency of changing energy policies (19)
38. Subsidies for energy conservation (8)
39. Improve the balancing of different stakes when defining energy policy (13)
40. Encourage new models of markets (7)
41. Fund for innovations in energy conservation and production (8)
42. Stricter certification of sustainable energy (4)
43. CO2 trading systems (10)
44. More ambitious international climate agreements (9)
45. Improving awareness by education (8)
46. Higher taxes on fossil energy (19)
47. Take away barriers in rules and regulations (9)

Appendix B: Workshop level shared cognition

Workshop:	1		2		3		4		5		6		7		8	
N:	10		9		10		9		14		12		8		14	
Before/after:	b	A	b	a	b	a	b	a	b	a	b	a	b	a	b	a
Strategic ends																
1	s	S	s	s	s	s	s	s	s	n	s	s	n	n	n	s
2	s	S	n	s	s	s	s	s	n	s	s	n	s	s	s	s
3	s	S	s	s	s	n	s	s	s	s	s	s	n	n	s	s
4	s	S	s	s	n	n	s	s	n	s	s	s	s	n	s	s
5	s	S	n	s	s	n	s	s	s	s	n	n	s	s	s	s
6	n	S	s	s	s	s	s	s	s	n	s	s	n	s	s	n
7	s	S	s	n	n	n	s	s	n	n	n	s	n	n	s	s
8	s	S	s	s	s	s	s	s	s	s	s	s	s	s	s	s
9	s	S	s	s	s	s	s	s	s	s	s	n	s	s	s	s
Technological means																
10	n	N	s	s	n	n	s	n	n	n	n	n	n	n	n	n
11	n	N	n	n	n	n	n	s	n	n	n	n	n	n	n	n
12	n	N	n	s	n	n	n	n	n	n	n	n	n	n	n	n
13	s	S	s	s	s	s	n	s	n	n	n	s	n	n	n	n
14	n	N	n	n	n	n	n	n	n	n	n	n	n	n	n	n
15	s	S	n	s	n	s	s	s	s	s	n	s	n	n	n	n
16	n	S	s	s	s	n	s	s	s	s	n	s	s	n	n	s
17	s	S	n	s	s	s	s	n	n	s	s	s	n	s	n	s
18	s	N	n	n	n	n	s	s	n	s	n	s	n	n	n	n
19	n	S	n	n	s	s	s	n	s	s	s	n	n	s	n	n
20	s	S	s	s	s	s	s	s	n	s	n	n	n	s	n	s
21	n	S	s	s	s	s	s	s	n	s	n	s	s	n	n	n
22	n	N	s	s	s	s	s	s	s	s	s	s	s	n	n	n
23	n	S	s	s	s	s	s	s	s	s	s	s	s	s	n	s
24	n	S	n	s	s	s	s	s	n	n	s	s	s	s	s	s
25	s	S	s	s	s	s	s	s	s	s	n	s	s	s	s	s
26	s	S	s	s	s	s	s	s	s	s	s	s	s	s	s	s
27	s	S	n	n	n	s	s	s	n	s	n	n	n	s	n	n
28	n	S	n	n	n	n	n	s	s	s	s	n	s	s	n	s
29	n	S	n	n	n	s	n	n	s	s	n	n	n	n	n	s
30	n	S	s	n	s	s	n	s	s	s	s	s	n	s	n	s
31	s	S	n	n	n	s	s	s	s	s	s	s	n	s	s	s
32	s	S	s	s	n	s	s	s	s	n	s	s	n	n	s	s
33	s	S	s	s	n	s	s	s	n	s	s	s	n	n	n	n
34	s	N	s	s	s	s	s	s	s	s	n	n	s	s	n	s
Policy means																
35	s	S	n	s	s	s	s	s	s	n	n	n	n	n	s	s
36	s	N	n	s	n	s	s	s	s	s	s	s	s	n	n	s
37	s	S	s	s	s	s	n	s	s	s	s	s	s	s	s	s
38	n	N	n	n	n	n	n	n	s	n	n	n	n	n	n	n
39	s	S	n	n	s	n	s	s	n	s	n	n	n	n	n	n
40	s	S	s	n	n	n	s	s	s	n	n	s	n	n	n	n
41	s	S	s	n	s	s	s	s	s	s	n	s	n	n	s	n
42	s	S	s	n	s	n	s	s	n	s	n	n	n	s	n	n
43	s	S	n	s	n	s	s	s	s	s	s	n	n	s	n	s
44	s	S	s	s	n	s	s	s	n	n	s	s	n	n	s	s
45	n	S	s	s	n	s	s	s	n	s	n	s	s	s	n	n
46	s	S	s	s	s	s	s	n	s	s	n	s	n	n	n	s
47	n	S	n	n	s	s	n	n	n	n	n	n	n	n	n	n

N = number of participants; b = before; a = after; s = shared cognition (IQR<=1); n = no shared cognition (IQR>1

References

Alliander. (2012). *Annual report*.
Ackermann, F., & Eden, C. (2011). Negotiation in strategy making teams: group support systems and the process of cognitive change. *Group Decision and Negotiation, 20*(3), 293–314.
Allen, E., & Seaman, C. A. (2007), *Likert scales and data analyses* (40, pp. 64–65). Quality Progress.
Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of management journal, 39*(1), 123–148.
Barr, P. S., Stimpert, J. L., & Huff, A. S. (1992). Cognitive change, strategic action, and organizational renewal. *Strategic management journal, 13*(S1), 15–36.
Black, L. J., & Andersen, D. F. (2012). Using visual representations as boundary objects to resolve conflict in collaborative model-building approaches. *Systems Research and Behavioral Science, 29*(2), 194–208.
Bowman, C. (1995). Strategy workshops and top-team commitment to strategic change. *Journal of Managerial Psychology, 10*(8), 4–12.

Bowman, G. (2016). The practice of scenario planning: An analysis of inter-and intra-organizational strategizing. *British Journal of Management, 27*(1), 77–96.
Brown, J. D. (2011). Likert Items and Scales of Measurement. *SHIKEN: JALT Testing and Evaluation SIG Newsletter, 15*(1), 10–14.
Cabinet Rutte-Asscher. (2012). *Coalition agreement "Building bridges."*
Chi, M. T. (2009). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. *International handbook of research on conceptual change* (pp. 89–110) Routledge.
Cronin, M. A., & Weingart, L. R. (2007), *Representational gaps, information processing, and conflict in functionally diverse teams* (32, pp. 761–773). Academy of Management Review.
Cunha, A . A, & Morais, D. C. (2016). Analysing the use of cognitive maps in an experiment on a group decision process. *Journal of the Operational Research Society, 67*(12), 1459–1468.
Daniels, K., Johnson, G., & de Chernatony, L. (1994). Differences in managerial cognitions of competition. *British Journal of Management, 5*(s1), S21–S29.
Daniels, K., Johnson, G., & de Chernatony, L. (2002). Task and institutional influences on managers' mental models of competition. *Organization Studies, 23*(1), 31–62.
Dreu de, C. K., & Beersma, B. (2010). Team confidence, motivated information processing, and dynamic group decision making. *European Journal of Social Psychology, 40*(7), 1110–1119.
Gooyert de, V., Rouwette, E., van Kranenburg, H., Freeman, E., & van Breen, H. (2016). Sustainability transition dynamics: Towards overcoming policy resistance. *Technological Forecasting and Social Change, 111*, 135–145.
Gooyert de, V., Rouwette, E., Van Kranenburg, H., & Freeman, E. (2017). Reviewing the role of stakeholders in operational research: A stakeholder theory perspective. *European Journal of Operational Research, 262*(2), 402–410.
Eden, C. (1995). On evaluating the performance of 'wide-band' GDSS's. *European Journal of Operational Research, 81*(2), 302–311.
Eden, C., & Ackermann, F. (2010). Decision making in groups: theory and practice. In P. C. Nutt, & D. C. Wilson (Eds.), *Handbook of decision making* (pp. 231–272). Chichester: Wiley.
Franco, L. A. (2013). Rethinking Soft OR interventions: Models as boundary objects. *European Journal of Operational Research, 231*(3), 720–733.
Franco, L. A., & Greiffenhagen, C. (2018). Making OR practice visible: Using ethnomethodology to analyse facilitated modelling workshops. *European Journal of Operational Research, 265*(2), 673–684.
Franco, L., Hämäläinen, R., Rouwette, E., & Leppänen, I. (2021). Taking stock of Behavioural OR: A review of behavioural studies with an intervention focus. *European Journal of Operational Research, 293*(2), 401–418.
Franco, L. A., & Montibeller, G. (2010). Facilitated modelling in operational research. *European Journal of Operational Research, 205*(3), 489–500.
Franco, L. A., & Rouwette, E. A. (2011). Decision development in facilitated modelling workshops. *European Journal of Operational Research, 212*(1), 164–178.
Franco, L. A., Rouwette, E. A., & Korzilius, H. (2016). Different paths to consensus? The impact of need for closure on model-supported group conflict management. *European Journal of Operational Research, 249*(3), 878–889.
Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Transaction Publishers.
Haug, C. (2015). What is consensus and how is it achieved in meetings? Four practices of consensus decision-making. In Joseph A. Allen, Nale Lehmann-Willenbrock, & Steven G. Rogelberg (Eds.), *The Cambridge handbook of meeting science*. New York: Cambridge University Press.
Healey, M. P., Hodgkinson, G. P., Whittington, R., & Johnson, G. (2015). Off to plan or out to lunch? Relationships between design characteristics and outcomes of strategy workshops. *British Journal of Management, 26*(3), 507–528.
Henao, F., Franco, L. A., Whittington, R., & Johnson, G. (2016). Unpacking multimethodology: Impacts of a community development intervention. *European Journal of Operational Research, 253*(3), 681–696.
Hendriks, C. M. (2009). Policy design without democracy? Making democratic sense of transition management. *Policy Sciences, 42*(4), 341–368.
Hodgkinson, G. P. (1997). Cognitive inertia in a turbulent market: The case of UK residential estate agents. *Journal of Management Studies, 34*(6), 921–945.
Hodgkinson, G. P. (2002). Comparing managers' mental models of competition: Why self-report measures of belief similarity won't do. *Organization Studies, 23*(1), 63–72.
Hodgkinson, G. P., Whittington, R., Johnson, G., & Schwarz, M. (2006). The role of strategy workshops in strategy development processes: Formality, communication, co-ordination and inclusion. *Long Range Planning, 39*(5), 479–496.
Hodgkinson, G. P., & Wright, G. (2002). Confronting strategic inertia in a top management team: Learning from failure. *Organization studies, 23*(6), 949–977.
Jarzabkowski, P., & Paul Spee, A. (2009). Strategy-as-practice: A review and future directions for the field. *International Journal of Management Reviews, 11*(1), 69–95.
Jarzabkowski, P., & Seidl, D. (2008). The role of meetings in the social practice of strategy. *Organization studies, 29*(11), 1391–1426.
Johnson, G. (2007). Ritualizing strategic thinking: The effectiveness of the strategic away day. *Strategic Direction, 24*(1), 3–5.
Joldersma, C., & Roelofs, E. (2004). The impact of Soft OR on problem structuring. *European Journal of Operational Research Res, 152*, 696–708.
Junginger, M., Agterbosch, S., Faaij, A., & Turkenburg, W. (2004). Renewable electricity in the Netherlands. *Energy Policy, 32*(9), 1053–1073.
Kaplan, S. (2008). Framing contests: Strategy making under uncertainty. *Organization Science, 19*(5), 729–752.
Kaplan, S. (2011). Research in cognition and strategy: Reflections on two decades of progress and a look to the future. *Journal of Management Studies, 48*(3), 665–695.
Kellermanns, F. W., Walter, J., Lechner, C., & Floyd, S. W. (2005). The lack of consensus about strategic consensus: Advancing theory and research. *Journal of Management, 31*(5), 719–737.
Kellermanns, F. W., Walter, J., Floyd, S. W., Lechner, C., & Shaw, J. C. (2011). To agree or not to agree? A meta-analytical review of strategic consensus and organizational performance. *Journal of Business Research, 64*(2), 126–133.
Kemp, R. (2010). The Dutch energy transition approach. *International Economics and Economic Policy, 7*(2-3), 291–316.
Kemp, R., Rotmans, J., & Loorbach, D. (2007). Assessing the Dutch energy transition policy: How does it deal with dilemmas of managing transitions? *Journal of Environmental Policy & Planning, 9*(3-4), 315–331.
Kern, F., & Smith, A. (2008). Restructuring energy systems for sustainability? Energy transition policy in the Netherlands. *Energy policy, 36*(11), 4093–4103.
Knight, D., Pearce, C. L., Smith, K. G., Olian, J. D., Sims, H. P., Smith, K. A., & Flood, P. (1999). Top management team diversity, group process, and strategic consensus. *Strategic Management Journal, 20*(5), 445–465.
Lami, I. M., & Tavella, E. (2019). On the usefulness of soft OR models in decision making: A comparison of problem structuring methods supported and self-organized workshops. *European Journal of Operational Research, 275*(3), 1020–1036.
Liu, L. A., Friedman, R., Barry, B., Gelfand, M. J., & Zhang, Z.-X. (2012). The dynamics of consensus building in intracultural and intercultural negotiations. *Administrative Science Quarterly, 57*(2), 269–304.
Markóczy, L. (2001). Consensus formation during strategic change. *Strategic Management Journal, 22*(11), 1013–1031.
Markóczy, L., & Goldberg, J. (1995). A method for eliciting and comparing causal maps. *Journal of management, 21*(2), 305–333.
Maurer, J., & Pierce, H. R. (1998). A comparison of Likert scale and traditional measures of self-efficacy. *Journal of Applied Psychology, 83*, 324–329.
McCardle-Keurentjes, M., & Rouwette, E. A. (2018). Asking questions: a sine qua non of facilitation in decision support? *Group Decision and Negotiation, 27*(5), 757–788.
Mintzberg, H. (1994). *The rise and fall of strategic planning*. New York: Free Press.
Mintzberg, H. (1994). The Fall and Rise of Strategic Planning. *Harvard Business Review, 72*(1), 107–114.
Nadkarni, S., & Narayanan, V. K. (2007). Strategic schemas, strategic flexibility, and firm performance: The moderating role of industry clockspeed. *Strategic management journal, 28*(3), 243–270.
Narayanan, V. K., Zane, L. J., & Kemmerer, B. (2011). The cognitive perspective in strategy: An integrative review. *Journal of Management, 37*(1), 305–351.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*(3), 231–259.
PBL. (2012). *Climate and energy roadmaps towards 2050 in north-western Europe*. The Hague/Bilthoven, the Netherlands.
Rouwette, E. A., Vennix, J. A., & Mullekom, T. V. (2002). FaciliGroup model building effectiveness: a review of assessment studies. *System Dynamics Review: The Journal of the System Dynamics Society, 18*(1), 5–45.
Rouwette, E. A. J. A. (2011). Facilitated modelling in strategy development: measuring the impact on communication, consensus and commitment. *Journal of the Operational Research Society, 62*, 879–887.
Rouwette, E. A. (2000). The impact of group model building on behavior. In *Behavioral operational research* (pp. 213–241). Palgrave Macmillan, London.
Scott, R. J., Cavana, R. Y., & Cameron, D. (2016). Recent evidence on the effectiveness of group model building. *European Journal of Operational Research, 249*(3), 908–918.
Scott, R. J., Cavana, R. Y., & Cameron, D. (2016). Client perceptions of reported outcomes of group model building in the New Zealand public sector. *Group Decision and Negotiation, 25*(1), 77–101.
Scheff, T. J. (1967). Toward a sociological model of consensus. *American Sociological Review*, 32–46.
Smith, A., & Kern, F. (2009). The transitions storyline in Dutch environmental policy. *Environmental Politics, 18*(1), 78–98.
Stroebe, W., Nijstad, B. A., & Rietzschel, E. (2010). Beyond productivity loss in brainstorming groups: The evolution of a question. *Advances in Experimental Social Psychology, 43*, 157–203.
Sutcliffe, K. M., & Huber, G. P. (1998). Firm and industry as determinants of executive perceptions of the environment. *Strategic Management Journal, 19*(8), 793–807.
Tavella, E., & Franco, L. A. (2015). Dynamics of group knowledge production in facilitated modelling workshops: An exploratory study. *Group Decision and Negotiation, 24*(3), 451–475.
Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic management journal, 21*(10-11), 1147–1161.
Uitdewilligen, S., Waller, M. J., & Pitariu, A. H. (2013), *Mental model updating and team adaptation* (44, pp. 127–158). Small Group Research.
Vaara, E., & Whittington, R. (2012). Strategy-as-practice: Taking social practices seriously. *The Academy of Management Annals, 6*(1), 285–336.
van Rooijen, S. N. M., & van Wees, M. T. (2006). Green electricity policies in the Netherlands: an analysis of policy decisions. *Energy Policy, 34*(1), 60–71.
Vennix, J. A. M., Scheper, W., & Willems, R. (1993). Group model building. What does the client think of it?. In *Proceedings of the 1993 International system dynamics conference*.
Vennix, J. A. M., & Rouwette, E. A. J. A. (2000). Group model building. What does the client think of it now?. In *Paper presented at the international system dynamics conference*.

Vennix, J. A. M. (2016). *Management Science: That can be quite an art*. Nijmegen.
von der Gracht, H. A. (2012). Consensus measurement in Delphi studies. *Technological Forecasting and Social Change, 79*(8), 1525–1536.
Weick, K. E. (1995). *Sensemaking in organizations*. Sage.
Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization science, 6*(3), 280–321.
White, L. (2016). Behavioural operational research: Towards a framework for understanding behaviour in OR interventions. *European Journal of Operational Research, 249*(3), 827–841.
Whittington, R. (2006). Completing the practice turn in strategy research. *Organization Studies, 27*(5), 613–634.
Wilson, T. D. (2004). *Strangers to ourselves*. Harvard University Press.