

Chapter 7

Collaboration Governance and System Dynamics Modelling: What Do Clients Want?



Rodney J. Scott and Robert Y. Cavana

Abstract System dynamics modellers sometimes involve decision makers in the modelling process, a method known as “group model building”. Group model building has been used to support group decision making and collaborative governance. Group model building has been associated with several different outcomes; it is not clear which of these outcomes are important to the clients that choose to engage with system dynamics modellers to provide group model building solutions. This chapter reports on group decision making in the context of public policy design and implementation and explores which outcomes are important to potential clients in the New Zealand public sector.

Senior management within four government agencies identified the employees who were most likely to commission and conduct group decision processes. These individuals participated in detailed semi-structured interviews, and completed a written questionnaire, exploring the contexts in which group model building may be useful and the outcomes sought in each situation. The results suggest that, even within the public sector, the importance of a particular outcome will depend upon context. However, public servants generally appear to value trust and agreement over policy quality when conducting group-decision processes. Knowledge of the outcomes sought by potential clients helps guide the outcomes measured by researchers and helps practitioners to tailor communication messages to clients.

This chapter is an update on the previously published article: 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.

R. J. Scott (✉)

School of Business, University of New South Wales, Sydney, NSW, Australia
e-mail: r.scott@adfa.edu.au

R. Y. Cavana

Victoria Business School, Victoria University of Wellington, Wellington, New Zealand
e-mail: bob.cavana@vuw.ac.nz

© Springer Nature Switzerland AG 2020

C. Bianchi et al. (eds.), *Enabling Collaborative Governance through Systems Modeling Methods*, System Dynamics for Performance Management & Governance 4, https://doi.org/10.1007/978-3-030-42970-6_7

141

Keywords Group model building · Group decision making · Collaborative governance · System dynamics · Clients · Outcomes · Public sector · Public policy

7.1 Introduction

Over almost 40 years, system dynamics practitioners have experimented with involving the client in the modelling process (Greenberger, Crenson, & Crissey, 1976). These methods are now known as “group model building” (Vennix, 1995, 1996). Group model building includes a range of approaches that can be broadly categorised on two axes: the level of participation (Kolfschoten & Rouwette, 2006) and the use of quantitative versus purely qualitative models (Coyle, 2000). In some group model building interventions, models are built by experts with some input from participants, using quantitative modelling from the outset (Kolfschoten & Rouwette, 2006). In others, the model is built in workshops with or by participants, using qualitative data. In this latter group, simulation occurs only at the end of the project (Kolfschoten & Rouwette, 2006) if at all (Cavana, Boyd, & Taylor, 2007). There has recently been a greater effort within the group model building community to provide greater clarity on exactly how interventions work (Andersen, Richardson, & Vennix, 1997), with the development of repeatably “scripts” that describe different process steps (Hovmand et al., 2012). The scripts used in the New Zealand case studies mentioned in this chapter can be found in the book “Group Model Building: Using System Dynamics to Achieve Enduring Agreement” (Scott, 2018).

Group model building practitioners and researchers (employing a range of participative approaches) noticed that group model building resulted in changes in the behaviour of participating individuals and groups. There have been over 200 published studies reporting on the effectiveness of group model building (Rouwette, Vennix, & Van Mullekom, 2002; Scott, Cameron, & Cavana, 2015). These studies note a range of outcomes which in the group model building literature are considered to be “changes in the beliefs, evaluations, intentions and behaviours of participants” (Rouwette, Vennix, & Felling, 2009, p. 582).

Expert practitioners typically conduct group model building interventions on behalf of “clients” (Vennix, Scheper, & Willems, 1993). While some studies refer to the client as the organisation or organisations that hired the group model building practitioner (Rouwette, 2003; Thompson, 2009; Vennix, 1995), others refer to the individuals who make the decision to commission or purchase the practitioners services (Andersen et al., 1997; Eden & Ackermann, 2004; Martinez-Moyano & Richardson, 2013; Rouwette, 2011; Rouwette et al., 2009; Rouwette & Vennix, 2011). In the context of this chapter, clients are assumed to be the individuals who make purchasing decisions on the group process used. This has some similarities with the “gatekeeper” role described in other papers (Luna-Reyes et al., 2006; Richardson & Andersen, 1995; Rouwette, Korzilius, Vennix, & Jacobs, 2011). This

chapter also distinguishes between “clients” (who make purchasing decisions) and “participants” (who take part in the group process).

Several recent papers have explored the use of group model building in a New Zealand public service context (e.g. Cavana et al., 2007; Cavana, Smith, Scott, & O’Connor, 2014; Scott, Cavana, & Cameron, 2013, 2016a, 2016b). These report 12 outcomes associated with group model building: insight, mental model change, enduring mental model change, mental model alignment, enduring mental model alignment, communication quality, consensus, commitment to a conclusion, strategy implementation, power levelling, persuasive content, and perceptions of workshop conclusions by non-participants. It is not clear if these outcomes are typically important to clients, or of no consequence at all.

Group model building literature suggests that different outcomes may be emphasised or ignored depending on the context (Rouwette et al., 2009; Zagonel, Rohrbaugh, Richardson, & Andersen, 2004), and implores researchers to be very clear about the outcomes sought in a particular intervention (Andersen et al., 1997). However, in many studies it is not clear how the outcomes measured by the study relate to the intended outcomes valued by the client (Dwyer & Stave, 2008; Eskinasi, Rouwette, & Vennix, 2009; Huz, Andersen, Richardson, & Boothroyd, 1997; Rouwette et al., 2011; Vennix et al., 1993; Vennix & Rouwette, 2000).

Related fields, such as “soft OR”, have asked what their clients typically value, and argued that this is a critical question for researchers and practitioners alike (Eden & Ackermann, 2004). These authors described their experiences of interacting with clients, and comment on what they believe clients value, but did not present any empirical research. This chapter seeks to address that deficiency, and thereby contribute to the evidence base for understanding what clients of group decision-making processes, like group model building, typically value.

An alternate view is that understanding what clients want is part of the client engagement process—that each intervention should begin with a detailed and explicit discussion with the client on the purpose of the intervention (Martinez-Moyano & Richardson, 2013). Although such discussion is a component of good practice, there are advantages for researchers and practitioners of knowing a priori the outcomes that clients in a particular situation are likely to value. Group model building researchers need such information in determining which outcomes warrant further attention, while practitioners can improve their initial communication with prospective clients by understanding the outcomes that are most likely to be of interest.

This chapter reports on research designed to explore client attitudes to group decision-making process outcomes. There has been an increasing trend within the public service in many countries for collaborative decision making (Ansell & Gash, 2008). As a group decision-making process (or “group-decision support system”; Andersen, Vennix, Richardson, & Rouwette, 2007), group model building has been applied in many public policy settings (Mingers & White, 2010). This chapter reports on research conducted with a sample of New Zealand public servants who were seen by their organisations as most likely to commission and conduct group decision-making processes. Their opinions were canvassed through the use of

semi-structured interviews and a numerical scale questionnaire. They were asked to rate the importance of outcomes reported in group model building studies with New Zealand public servants, and also to suggest other outcomes that were important to them. The interviews discussed when and why group-decision processes would be used, and when different outcomes were important or unimportant.

The chapter is structured into four sections after this introduction. The first section reviews the outcomes reported in the previous papers related to this topic. The second section describes the research methods. The third section reports on the results of the interviews and questionnaires. Finally, there is a discussion of what this means for group model building research and practice.

7.2 Group Model Building Outcomes

Group model building describes a range of qualitative and quantitative system dynamics methods that involve the client in the modelling process. The 12 outcomes considered in this chapter were selected from 5 case studies that examined group model building in a New Zealand public service context (Cavana et al., 2007, 2014; Scott et al., 2013, 2016a, 2016b). These studies used only qualitative group model building methods, but similar results have been reported using quantitative methods (e.g. Huz, 1999; Rouwette et al., 2011; Van Nistelrooij, Rouwette, Vestijnen, & Vennix, 2012; Vennix et al., 1993).

These case studies evaluated several public service group model building processes using three evaluation tools: a survey tool (Scott et al., 2016a), a pre-test/post-test/delayed-test questionnaire (Scott et al., 2013) and semi-structured interviews (Scott et al., 2016b). A prior meta-analysis compared the data gathered using survey tools and post-intervention interviews (Rouwette et al., 2002); this analysis revealed no difference in the outcomes reported by participants in group model building using either method.

The survey was based on a popular tool used in earlier group model building studies (Rouwette, 2011; Vennix et al., 1993; Vennix & Rouwette, 2000) that was administered immediately after participation in a group model building workshop. This was used to confirm that participants felt that the process had contributed to increased communication quality, insights, consensus and commitment to conclusions. Strategy literature reports these outcomes as being predictive of effective strategy implementation (Noble, 1999; Scott et al., 2016a; Skivington & Daft, 1991). Participants also compared the process to a hypothetical “normal” meeting and believed that group model building was comparatively more effective and more time-efficient (Scott et al., 2016a).

The survey also revealed that non-managers rated the presence of an independent facilitator as important to their experience of the workshop (Scott et al., 2016a). This was related to “power levelling” (van Nistelrooij et al., 2012), where less-powerful members are less disadvantaged in their contribution to the discussion (in this study, the positional rank was used as a proxy for power).

The pre-test/post-test/delayed-test questionnaire collected participants' recommendations for actions to address the problem at hand (Scott et al., 2013). This tool was administered immediately before, immediately after, and 12 months following participation in a group model building workshop. The results of this evaluation demonstrated that participants changed their mind during the workshop and that these decision preferences persisted for at least 12 months. Because of its enduring nature, this difference was attributed to "mental model change". Mental models are a construct from cognitive science that has attracted significant attention in the group model building literature (Doyle & Ford, 1998). Mental models are generally regarded as an "internal representation of an external reality" (Jones, Ross, Lynam, Perez, & Leitch, 2011), or "small scale models that the individual believes is analogous to how the world works" (Craik, 1943). Group model building literature distinguishes between "mental model change", that is, any alteration of a participants' mental models regarding the system modelled, and "mental model alignment", which describes when participant mental models become more alike (Scott et al., 2013). Cognitive science literature differs on how stable and enduring mental models may be (see Johnson-Laird, 1983, versus Moray, 1998). Similarly, some group model building research has explored short-term changes in mental models (see Rouwette & Vennix, 2006), and other research has explored mental model change that persists over a year (Scott et al., 2013). The pre-test/post-test/delayed-test method was able to distinguish between these various outcomes, hereafter described as "mental model change", "mental model alignment", "enduring mental model change" and "enduring mental model alignment".

Participants' new decision preferences were categorised as coming from two sources—some were persuaded by the views of other participants, and others developed new insights from their participation in the process. New insights from participating were more enduring than those developed through persuasion (Scott et al., 2013). Finally, individuals who did not participate in the workshop process did not prefer the decisions made in group model building workshops to other decision alternatives (Scott et al., 2013).

These outcomes may be interrelated. The theory of planned behaviour (Ajzen, 1991) suggests that communication quality fosters insight and consensus, and insight and consensus contribute to a commitment to conclusions (Rouwette, 2003, Fig. 7.1a). Insight, consensus, communication quality and commitment to conclusions are predictive factors supporting effective strategy implementation (Noble, 1999; Scott, 2014). Group model building is believed to support mental model change through a combination of persuasive arguments from other participants and novel insights from the modelling process (Rouwette et al., 2011, Fig. 7.1b). Where group model building has been associated with a long-lasting alignment of participants (Scott et al., 2013, Fig. 7.1c), this has been explained as related to the enduring nature of mental models of dynamic systems (Doyle & Ford, 1998). Power levelling is believed to support improved communication by providing the opportunity for more varied interactions (van Nistelrooij et al., 2012, Fig. 7.1d). The rating of workshop conclusions by non-participants could not be related to other outcomes. Various authors have attempted to combine different theories (Richardson,

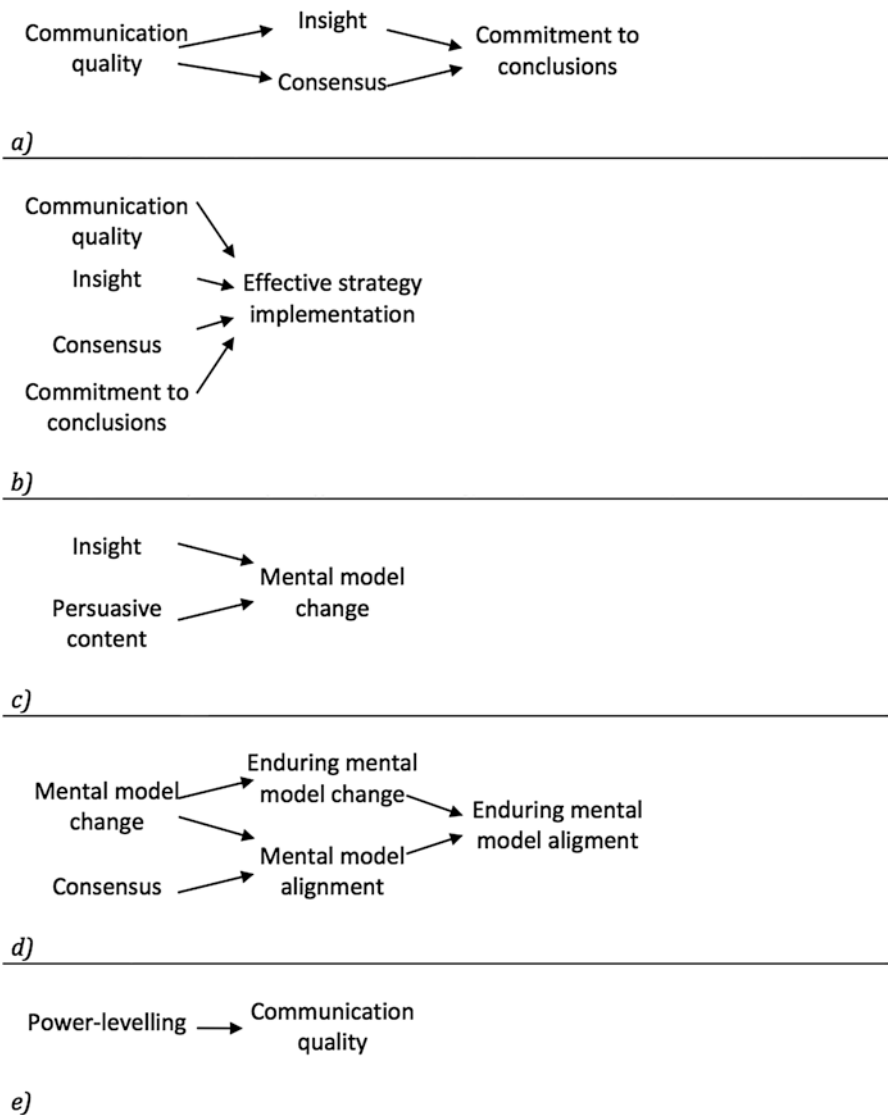


Fig. 7.1 Theoretical relationships between reported outcomes of group model building. (a) Theory of planned behaviour (Ajzen, 1991; Rouwette, 2003). (b) Success factors for effective strategy implementation (Noble, 1999). (c) Modelling as persuasion (Rouwette et al., 2011). (d) Enduring effects of group model building (Scott et al., 2013). (e) Effects of power-levelling (Van Nistelrooij et al., 2012)

Andersen, Maxwell, & Stewart, 1994; Rouwette et al., 2011; Scott, 2017) though these connections remain theoretical or speculative.

In an earlier paper (Scott et al., 2016a), an individual client was asked to describe their desired outcomes for the group model building process. They indicated that

they wanted to: create among employees a common understanding of their new organisational strategy; create agreed on implementation actions for the strategy, and increase commitment to the strategy. The prevalence of these goals is unknown, whether within other organisations, or even other problem settings (or timing) within the same organisation. The purpose of this chapter is to inform our understanding of the importance of these outcomes and to identify other outcomes that may also be important.

7.3 Methods

The study described in this chapter explored the views of public servant clients—those that regularly conduct or commission group decision-making processes (see Sect. 7.3.1)—on the importance of a variety of outcomes associated with group model building.

The study employs a mixed-methods approach to evaluation research (Blaikie, 1993). Primarily qualitative methods were chosen to explore in-depth the experiences and beliefs of the interviewees (Kvale & Brinkman, 2008, see Sect. 7.3.2), supplemented by a quantitative survey to improve the reliability of findings (Blaikie, 1993, see Sect. 7.3.3). The interviews included open questions, where interviewees identified and discussed the outcomes that were important to them and direct questions about the reported outcomes being investigated. Interviews were transcribed, and the qualitative and quantitative data compared (see Sect. 7.3.4). While this mixed-methods approach was useful in validation of the interview results, the small sample size is not ideal for quantitative or statistical analysis. This and other limitations in the study design are described in Sect. 7.3.5, which should be factored when interpreting the results.

The study in this chapter was exploratory, but the authors hypothesised that both the nature of the outcome and several contextual factors would influence the importance of that outcome (see Fig. 7.2).

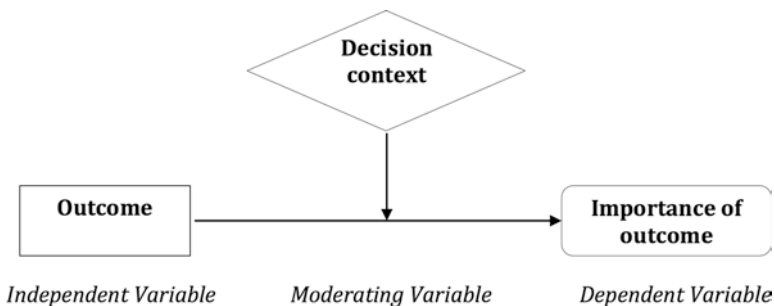


Fig. 7.2 Conceptual model for the importance of group model building outcomes in the New Zealand public sector

7.3.1 *Sample Selection*

Group model building is frequently conducted by expert “consultants” from the system dynamics community, for “clients” in the public sector and elsewhere. These clients choose between different potential methods for group decision making. Some will choose group model building, and some will choose other methods. For the group model building community to engage with and “pitch” their services to these clients effectively, they will be advantaged by understanding what these clients value. This study aims to explore the views of these clients and also potential clients—those that have not engaged with group model building before but who may do so.

As discussed below, the research involved a small number of research subjects. Consequently, the subjects chosen needed to be those who were most likely to represent the views of potential public sector clients and potential clients. The views of public sector clients and potential clients will not necessarily be the same as those of public servants more broadly.

The first author approached six New Zealand government agencies that have responsibility for developing public policy. Of these, four responded: the Ministry for Business, Innovation and Employment; the Ministry for Primary Industries; the Ministry for the Environment; the Department of Conservation. A gatekeeper (senior executive) at each agency selected individuals in their organisation who they believed most regularly commissioned or conducted group-decision processes, to aid work related to public policy. This study design uses a non-probability judgement sampling method (Cavana, Delahaye, & Sekaran, 2001), which is a form of intentional selection bias. The study design did not aim to pick a random sample that was representative of public servants, but instead to pick those public servants who were the most likely to be clients—those that most regularly commissioned or conducted group-decision processes. In a judgement sample, the risk to the validity of the study is that the gatekeepers were not able to accurately identify who in their organisation commissions or conducts group-decision processes. While it is plausible that such errors in selection occurred, it seemed likely to the authors that these gatekeepers would be better placed, with better information, to identify the correct sample than anyone else, or any other method.

Research using qualitative interviews ideally concludes when “data saturation” has been reached; the point in data collection when no new additional data are found that develop aspects of a conceptual category (Guest, Bunce, & Johnson, 2006). Conversely, logistics may require some estimate of the necessary sample size before the research has been conducted (Green & Thorogood, 2009). Francis et al. (2010) propose two steps for deciding data saturation: first, specify a minimum sample size (initial analysis sample); and second, specify how many additional interviews will be conducted without new ideas emerging (stopping criteria). The aims of the study and characteristics of the group influence the likely saturation point (Charmaz, 2006; Mason, 2010). Seven criteria have been proposed for determining an appropriate initial analysis sample size:

- The heterogeneity of the population
- The number of selection criteria
- The nesting of criteria
- Groups of special interest that require intensive study
- Multiple samples within one study
- Types of data collection methods use
- The budget and resources available (Ritchie, Lewis, & Elam, 2003)

The research in this chapter involves a selected, relatively homogenous group (public policymakers, managers, people who commission group-decision processes). There are no comparison groups, and the methods are primarily qualitative. These factors suggest a relatively small group is likely to be sufficient. Two comparable studies reported data saturation at 14 and 12, respectively (Francis et al., 2010; Guest et al., 2006).

There is no established theory on how to determine the appropriate number for stopping criteria, but three is commonly used (Francis et al., 2010). On balance, an initial sample analysis of 12 and stopping criteria of 3 was selected as most appropriate. After 12 interviews, the final 3 revealed no significant, new, unique information (i.e. achievement of data saturation). Though a robust sample for a detailed qualitative study, this is a small number on which to make meaningful conclusions on the quantitative survey data—this limitation is explored further in the Discussion section. Interviewee demographics are shown in Table 7.1.

Table 7.1 Interviewee demographics

Parameter	Value
Number of interviewees	12
Government agencies represented	4
Age	44 years
Mean	31–56 years
Range	
Length of employment in public sector	6 years
Mean	1–20 years
Range	
Gender	6
Male	6
Female	
Organisational level	2
Director	3
Group manager	1
Team manager	6
Non-manager	
Highest qualification	9
Postgraduate	3
Undergraduate	0
Completed secondary	

7.3.2 Interviews

Each research subject took part in a face-to-face interview following a semi-structured format (Kvale & Brinkman, 2008). Each interview consisted of three themes: the interviewee's experiences with group-decision processes, the interviewee's desired outcomes (and when these outcomes might be most applicable) and the interviewee's opinions of the outcomes being investigated. Each of these themes is explored further below.

The interviewee was first asked to describe the context of problem settings in which they have used group-decision processes. This included prompts on the participating parties in the group-decision process, the decision being made and the consequence of that decision. Follow-up questions further explored the tools or processes that were used. This theme was used for three purposes: to establish the relevance of the interviewee as a person who regularly commissions or conducts group-decision processes, to investigate the kinds of problem settings encountered by public servants who use these processes and to discover what tools were being employed.

The interviewee was then asked which outcomes were important in the experiences they had described, why these outcomes were important, and what aspects of the decision context contributed to their importance. This was used to validate later questions: in this theme, the interviewee did not know which outcomes interested the researcher, and so the opportunity for subject bias (Orne, 1962, where individuals report what they think researchers want to hear) was reduced. This was also used to identify outcomes other than those being investigated.

Finally, the interviewees were supplied with a set of outcomes. This method necessarily involved the authors selecting a set of outcomes to test. Group model building literature includes upwards of 30 outcomes (though some may overlap, see Rouwette et al., 2002; Scott et al., 2015). The authors selected the 12 outcomes described in the 5 New Zealand public sector case studies described in Sect. 7.2 (Cavana et al., 2007, 2014; Scott et al., 2013, 2016a, 2016b). These were selected on the basis that those initial studies had presumably tailored their design to match the outcomes that they believed were important for the specific geographic (New Zealand) and sectoral (public sector) context. This selection process meant the omission of one outcome that ultimately proved important in the results—that of process efficiency, which has been assessed in other group model building literature (Vennix et al., 1993; Vennix & Rouwette, 2000).

For each of the 12 outcomes, the interviewer asked whether it was important, when it might be important and how successful the interviewer's existing processes were in achieving this outcome. When interviewees described an outcome as sometimes important, further prompts were used to explore what factors determined whether that outcome was important or unimportant. This theme was used to evaluate each of the reported outcomes in turn. When there was any confusion about the meaning of outcomes described (for example, "mental model change"), the

interviewer provided an additional explanation based on their understanding of how the terms have been used in group model building literature.

The interviews ranged in length between 30 min and 1 h and were recorded by an audio recorder. The interview transcripts were analysed as described in Sect. 7.3.4.

7.3.3 Questionnaire

A written questionnaire was given to the research subjects after the interview. The questionnaire consisted of two parts: demographic questions and questions on the importance of each of the reported outcomes of group model building.

The demographic questions concerned parameters described in Table 7.1. Previous research had revealed that age, gender and education level did not affect participants' reported experience of group model building (Scott et al., 2016a), but the effects of different clients' demographic variables on how they valued outcomes were unknown. As noted earlier, the organisational rank (manager versus non-manager) was used as an imperfect proxy for power, which is significantly more difficult to measure directly. The question on organisational rank was included to explore whether there was a relationship between rank and outcome preference; an earlier study revealed that less powerful participants rated the importance of an independent facilitator more highly (Scott et al., 2016a).

The second part consisted of seven-point numerical scale questions to provide a quantitative indication of the importance of each of the outcomes from the literature (Cavana et al., 2001). Research subjects were asked to rate each outcome, by circling a number between 1 and 7, where 1 meant that the outcome was of no importance and 7 meant that the outcome was very important. This provides a separate measure of the subjects' views on the different outcomes, similar to the qualitative answers in the third interview theme.

The written questionnaire was used to improve the reliability of the findings. The research design was primarily qualitative because the authors wanted to understand the research subjects' experiences and beliefs. However, the interview questions have not been validated, so combining interview and questionnaire results in a mixed-method study was used to improve reliability (Blaikie, 1993). One outcome (views of non-participants) was omitted from the questionnaire in error, and this is a limitation of the data.

7.3.4 Analysis

The responses to the interview questions were transcribed, then subject to content analysis using manual coding (Cavana et al., 2001). The 12 assessed outcomes (see Sect. 7.2) were pre-determined as codes, as these were the main subjects of the study in this chapter. Any additional outcomes mentioned by interviewees were also

coded. Other codes were emergent (Holsti, 1969; Strauss & Corbin, 1990). The analysis was then constructed based on the themes that emerged in the text, illustrated with verbatim responses where these were useful in explaining each theme.

The rated outcomes were compared using commonly applied statistical methods. The seven-point numerical scales used in the questionnaire were assumed to represent interval data (Cavana et al., 2001). A Kolmogorov-Smirnov test was used to confirm normal distribution, which allows the use of a Student's t-test to determine significance (Stephens, 1974). Results for each question were compared to a neutral response (a score of 4 on the 1–7 scale), and to the overall mean (a score of 5.3 on the 1–7 scale), using a two-tailed t-test (as results could vary in either direction—Stephens, 1974).

7.3.5 *Limitations*

The study in this chapter investigates the stated beliefs of a small number of New Zealand public servant “clients”, to determine what outcomes they value as important in group decision making. These were then related to recently reported outcomes of group model building.

The individuals were selected by their agencies as those who most regularly commission or conduct group-decision processes, and so are likely to be the most relevant subjects for understanding potential group model building clients in the New Zealand public sector. Twelve individuals were interviewed. For detailed qualitative research, this number proved sufficient to achieve data saturation. For quantitative research, however, the sample size is small. The quantitative data was primarily used to support the results obtained by the interviews and should be used with caution as stand-alone measures that are representative of any broader group.

This findings presented in this chapter rely on the individuals' stated preference for different outcomes. It is possible that these do not represent individuals' actual preferences, and opens the possibility of a range of biases, including subject bias (Orne, 1962) and social desirability bias (Edwards, 1957). The results show a strong preference for agreement and efficiency, over an interest in decision quality. It is not obvious why individuals would (for example) choose to downplay their interest in improving decision quality through insight, or why it might be socially desirable to do so. These biases could potentially be addressed through an alternate study design that explored clients' revealed preferences rather than stated preferences (Samuelson, 1938), but collecting this data would be more challenging.

The framing of the interview as relating to “group decisions” may have led interviewees to focus on interpersonal (group) aspects. Perhaps asking instead about (for example) “solving complex problems” would have revealed greater preference for decision quality rather than group agreement. Different outcomes are likely to be important in different settings; however, group participation is one of the defining aspects of group model building, so framing the possible problems as “group decisions” did not seem inappropriate.

This chapter provides insights into the outcomes that are important to New Zealand public servants in commissioning and conducting group-decision processes. We would generally expect the results to be most applicable to settings that are most similar to this context. The descriptions of public servants' experiences with group decision making are consistent with international trends toward inter-agency and inter-stakeholder group decisions (Newman, Barnes, Sullivan, & Knops, 2004, and as explored in section Discussion), but this chapter does not specifically demonstrate that results of this study apply to other countries. Preferences in the private sector may vary from those in the chapter due to the different incentives of the commercial environment. The study would benefit from validation by similar studies in other contexts.

7.4 Results

Each interviewee demonstrated broad experience in commissioning or conducting group-decision processes and described multiple situations where group-decision processes had been used. This confirmed that the research subjects were well selected as potential clients or users of group model building methods.

The results come from interview and questionnaire responses and describe the importance of different outcomes in different contexts. The results were consistent with the conceptual model described earlier (Fig. 7.2), in that the importance of the outcome was affected by the nature of that outcome and several contextual factors. For some outcomes, interviewees described the outcome as important as a precondition to another more desirable outcome (for example, communication quality was seen as a pre-requisite for mental model alignment). Several interviewees described outcomes as mutually reinforcing.

A range of contextual factors influenced the importance of some outcomes: the stage of the decision process, the participating parties in the decision, and the demographics of the client, each explored below. Some outcomes were more important at different stages of the interview process, for example, insight was seen as more useful in generating new ideas at the start of a process, and consensus seen as more useful at the end of a process (see Sect. 7.4.1). The nature of the participating parties also affected the importance of some outcomes; for example, process efficiency was very important in potentially time-consuming government-stakeholder group decisions (see Sect. 7.4.2). Finally, client demographics had some impact on the results. While gender, age or education level had no significant differences, the responses varied by level of experience and organisational rank (see Sect. 7.4.3).

The results are presented in three parts: interviewees' descriptions of the importance of each outcome; how the nature of the participating parties affected the importance of each outcome; statistical analysis of the questionnaire results.

Table 7.2 Outcomes volunteered by interviewees as important in past group decisions

Organisation	1			2			3			4			
Interview subject	1	2	3	4	5	6	7	8	9	10	11	12	Total
Commitment to conclusions	✓	✓	–	✓	–	✓	✓	✓	–	✓	✓	–	8
Communication quality	✓	–	✓	✓	✓	✓	–	–	–	✓	✓	✓	8
Consensus	✓	✓	✓	–	✓	✓	✓	–	✓	✓	✓	–	9
Mental model change	✓	–	–	–	–	✓	–	–	–	–	–	–	2
Enduring change	✓	–	–	✓	–	–	–	–	–	–	–	–	2
Mental model alignment	–	–	✓	✓	–	–	–	✓	–	✓	✓	–	5
Enduring alignment	–	–	✓	✓	–	–	–	–	–	✓	–	–	3
Effective implementation	–	–	–	–	–	–	–	✓	–	–	–	–	1
Persuasive content	–	–	–	–	–	–	–	–	–	–	–	–	0
Power levelling	–	–	–	–	–	–	–	–	–	–	–	–	0
Insight	✓	–	–	–	–	✓	–	–	✓	–	–	–	3
View of non-participants	✓	–	–	–	–	–	–	–	–	–	–	–	1
Efficiency	✓	✓	–	–	✓	✓	✓	–	–	–	✓	✓	7
Further working together	–	–	–	–	–	–	–	✓	–	✓	–	–	2
Willingness to endorse	✓	–	–	–	–	–	–	✓	–	–	–	–	2
Attachment to language	–	✓	–	–	–	–	–	–	–	–	–	–	1
Participant disclosure	–	✓	–	–	–	–	–	–	–	–	–	–	1
Tiebreaking	–	–	–	–	–	✓	–	–	–	–	–	–	1
Completeness	–	–	–	–	✓	–	–	–	–	–	–	–	1

Outcomes listed above the dotted line are the 12 outcomes investigated

7.4.1 Results for Each Outcome

Three different sources were used to determine which outcomes were most important: the second theme of the interviews, where interviewees were asked to describe the outcomes that had been important in past situations (see Table 7.2); the third theme of the interviews, where interviewees were asked about the importance of specified outcomes; and the written questionnaires, where respondents were asked to rate the importance of specified outcomes on a numerical scale. These three methods showed very strong agreement, with a few exceptions noted in relevant paragraphs below, where results relating to each outcome are discussed in turn.

Commitment to conclusions was the highest-ranked outcome by the questionnaire responses. Interviewees distinguished between finding something acceptable for agreement in the meeting (consensus) and being committed to supporting and implementing those conclusions. Commitment was more important when the goal was to affect change (interagency cooperation, joint action with stakeholders), than when an agreement marked the end of the process (providing advice to a Minister or senior manager). Three interviewees mentioned that they had previously relied on voting methods to reach an agreed conclusion; however, there was concern that these methods may sometimes lead to low commitment, particularly by those whose preferred conclusions were not selected.

Communication quality was also highly rated by the questionnaire and interview responses. Communication quality was seen as “*crucial*” and “*where it all starts.*” In particular, communication quality was seen as important when working with stakeholders who did not have a “*shared language*” (“*Engineers and planners don’t even speak the same English.*”). Communication quality was seen as a pre-requisite for mental model alignment, which was seen as the ultimate outcome by one interviewee.

Consensus was generally rated as important in the questionnaire and interview responses. In many cases, coming up with “*any agreement*” was seen as a successful result. This was particularly the case in inter-stakeholder decision processes—public servants were keen that participants all agree, even if those same convenors did not see the detail of the agreement as ideal. Several responses laboured the distinction between an ideal solution and one that all participants found acceptable for agreement. Particularly in interagency processes, participants were seen as sophisticated negotiators who would trade-off different benefits to reach an acceptable agreement (in the absence of viable alternatives to a negotiated agreement). An agreement was often achieved around non-preferred but acceptable options.

Mental model change was one of the lower-ranked outcomes from the questionnaire responses, but *enduring mental model change* was one of the highest-ranked. Interview responses do not fully explain this difference. Mental model change was seen as a luxury by some interviewees—the goal was to reach an agreement, not have transformative experiences for the participants. Agreements were often seen as “*incremental*”—“*we’re not expecting big shifts in how people see the world*”. Occasionally there is a need for a “*step change*”, and in those instances, a technique for supporting mental model change would be desirable, but this applied to a minority of circumstances.

Enduring mental model change was perhaps interpreted by some interviewees as enduring agreement with the workshop conclusions; interviewees noted common delays between group-decision processes and implementation, and were particularly concerned that participants would “*go feral*” or start “*throwing stones*” at the conclusions that they had previously agreed to—“(somebody) *effectively renegeing would have been a disaster*”. While enduring mental model change may be one mechanism for reaching enduring agreement with workshop conclusions (if participants’ new models are more consistent with those conclusions), the two do not necessarily follow.

Mental model alignment was ranked moderately highly by the questionnaire responses. However, interviewees often described concepts similar to mental model alignment as their most sought-after outcomes. This was particularly true when interviewees were asked what outcomes were important to them (without being prompted with possible outcomes). Interviewees described “*shared understanding*”, being “*able to understand where each other is coming from*”, and “*seeing things from their point of view*” as especially important. One interviewee recalled his previous experience as a negotiator: “*People who are on opposite sides of the table don’t have opposite perspectives, they have different ways of looking at the same problem*”...“*What seems a perfectly logical conclusion from your starting*

point, they may come to the opposite conclusion, not because they disagree with the logic but because they're coming from a different place." Any tools or techniques that would allow participants to see the world in a more compatible way were seen as especially desirable. From these interview responses, it might be expected that mental model alignment would have been ranked more highly among the questionnaire responses. It is possible but unconfirmed that the language "mental model alignment" was unfamiliar to respondents, and that this led to lower rankings than expected. The interviewer provided clarification on the meanings of the terms in the interview, but this clarification was not available to respondents while filling in the questionnaire.

Effective strategy implementation was an outcome that did not appear well understood by some interviewees, and it was difficult to relate some answers to the questions asked. Many group-decision processes did not involve strategy implementation and therefore were not applicable. Where this was seen as important, interviewees drew distinction between talk and action ("If you don't actually implement it, then what's the point"). Applied business research struggles to evaluate system changes (Shadish, Cook, & Campbell, 2001), and this is an ongoing research challenge for group model building.

Some interviewees valued the *persuasive content* of the decision process used. Previous group model building research demonstrates that some learning occurs from other participants in the workshop, and some represent new ideas from the modelling process (Scott et al., 2013). Interviewees were asked which of these was more important or should be more emphasised. Responses were mixed and closely followed interviewees' attitudes toward the importance of insight in their processes. Those who valued new insights saw persuasion toward existing beliefs as a barrier to creation. In contrast, those who valued agreement by any means (regardless of the quality of that agreement) saw compelling persuasion as a useful means to speed the arrival of agreement. Previous studies considering persuasion did not propose how the amount of persuasion or new insight could be increased or decreased (Rouvette et al., 2011; Scott et al., 2013).

Power levelling was a concept that drew polarised responses in both the questionnaire and the interviews. Having less powerful members contribute was seen as useful in generating insight ("If it's about ideas, then you really do want to be in the situation where all participants have equal opportunity to contribute."), and in increasing a sense of "*engagement and ownership*" by those participants. Power imbalances were sometimes seen as a strong barrier to participation—"You can certainly see situations where relatively junior people are afraid to talk" and "you just get the loudest voices and the ones with the quickest tongues." Where interviewees used techniques to encourage contribution from everyone, they typically involved forcing participants to take turns in offering perspectives—interviewees talked about "*going around the room*" to elicit input individually, or using "*snowballing*" techniques to aggregate individual contributions (Thomas & Carswell, 2000). This is very different to the way group model building is thought to create power levelling, through allowing contribution and modification of the model through input from all participants (Black & Andersen, 2012; Van Nistelrooij et al., 2012).

In contrast, power levelling was sometimes seen as counter-productive. Towards the end of the group-decision process, “*when it comes close to closing the deal*”, it was seen as sometimes beneficial for those “*who don’t have authority...to sit quietly and listen to those that do.*” Some interviewees thought it represented a more durable outcome where those who had more power were more able to influence the content of the agreement—“*power is power*”. Most interviewees described power levelling as relatively unimportant, and power levelling was overall rated as one of the less important outcomes of group-decision processes.

Insight was seen as useful “*at the beginning, to open things up*” or when “*prototyping*”. However, in some cases, interviewees were more interested in coming up with “*any agreement*” than whether this agreement contained any new ideas. One positive aspect of insight was that in interagency processes, new ideas were not seen as being owned by an individual agency and, therefore, were easier for other agencies to agree with. Insight was seen as unhelpful when it complicated the parameters of the discussion and delayed progress to an agreement—“*you don’t want new ideas when you’ve trying to nail something down.*” Overall, insight was not seen as very important in group-decision processes and was the lowest-ranked outcome among the questionnaire responses.

Views of non-participants were seen as sometimes very important and sometimes not important. In many cases, particularly where the end goal of the processes was to reach an agreement, it was sufficient for only those present to agree, so long as those people had authority to do so (“*As long as you’ve got the right people in the room*”). However, in some cases described by interviewees, buy-in by broader constituencies was vital. Stakeholders were used as focus groups, with the assumption that if they agreed with a proposal, it would likely be acceptable to other stakeholders with similar interests. Previous research found that conclusions developed through group model building were compelling to those present in the workshop, but not compelling to others (Scott et al., 2013). Client acceptance of solutions developed through system dynamics modelling is a long-standing challenge (Greenberger et al., 1976). Group model building aimed to overcome this challenge by involving clients in the modelling process (Vennix, 1996). Where participants have to relay findings to a broader constituency, or where participants are assumed to be representative of non-participants with similar interests, the problem of compelling communication of system dynamics conclusions is resurrected. Further research is needed to develop better ways of communicating conclusions from the application of system dynamics methods (Serman, 2000).

Efficiency was seen as a key parameter (“*The biggest concern we have is time.*”), though participants were not specifically asked to rate its importance. Interviewees lamented that group-decision processes take considerably longer than decisions taken by individuals (“*If you were doing it by yourself, multiply the time by twenty and that’s how long it takes with a group*”). Group model building participants have previously been asked to compare the speed of progress between a group model building workshop and a hypothetical “normal meeting” (Vennix et al., 1993; Vennix & Rouwette, 2000). In these studies, participants believed that group model building led to insight, consensus and commitment to conclusions more quickly

than a normal meeting. If speed and efficiency are very important to public servants in designing group-decision processes, greater care should be taken in evaluating the speed of group model building processes compared to other group-decision processes.

Further working together was suggested by two interviewees as a key outcome of group-decision processes. In this way, participants create their own “*culture*”, “*cooperation is built incrementally*”, and future decisions have a foundation of mutual trust and “*goodwill*”. Previous research has evaluated further use of group model building tools by an organisation (Bentham & de Visscher, 1994), but not the willingness of participants to continue to work together. The boundary object mechanism for understanding group model building outcomes (Black & Andersen, 2012) proposes a reinforcing loop where “our progress fuels working together”. Empirical evidence of this loop would reassure public servants that use of group model building can be part of a process to build ongoing collaborative relationships.

Willingness to endorse was mentioned by two interviewees. This related to the inclination to publically uphold the conclusions of the decision process and referred to situations where government was co-developing a product or programme in partnership with key stakeholders. The interviewees wanted an endorsement from the group decision participants, to prevent later reputational risk to the credibility of the programme. One popular group model building research tool (the “CICC” questionnaire—Vennix et al., 1993) includes a question on willingness to endorse: “I will uphold the conclusions/findings of these meetings in front of other members of my organisation (personal communication, Etienne Rouwette, 2011).” If this outcome is important to some clients, it may be useful to report specifically on willingness to endorse in future research.

Several other outcomes were mentioned by a single interviewee only. One described a desire for a technique to overcome participants’ attachment to individual words and to focus more on the content and meaning of the agreement—*attachment to language* was seen as a barrier and delay to reaching agreement. This cannot be directly related to reported outcomes of group model building. Modelling (as a visual language) may act to interrupt any fixation on textual editing. Conversely, the act of defining variables may provide a new opportunity for language preferences to form a barrier to agreement.

One interviewee described the need for *participant disclosure*—“*we want people to put their cards on the table*.” This can be related to two findings in the group model building literature. In the group model building process discussed in Scott et al. (2013), participants literally put their cards on the table—writing the variables they believed were important on post-it notes, and sharing those with the group. Another study investigated the extent to which unique information (information only known to one person) was communicated within the group, and the extent to which participants used information received (McCardle-Keurentjes, Rouwette, & Vennix, 2008).

Another interviewee described the need for a shortcut to decision making between several choices where none is obviously better. “*If you’ve got three (options) and none is patently better than the others, then pick one*.” The need for a

mechanism for *tiebreaking* was seen as sometimes stalling otherwise-successful projects when near completion. It is unclear how group model building could be useful at this stage—applying a system dynamics perspective at this time may challenge several underlying assumptions and re-open a process that was reaching its conclusion.

Finally, one interviewee believed that it was important to ensure that no important factors or risks had been omitted from the discussion (“*How do you check you’ve got all the important stuff?*”). In context, it seemed that this focus on *completeness* was likely related to the defensibility of the decision. System dynamics practitioners may believe that their methods are more comprehensive or holistic; however, this is difficult to measure empirically.

There was limited focus on policy quality, except indirectly (as inferred through the interest in insight, power levelling, and completeness).

Those that mentioned outcomes not raised in other interviews included men and women, managers and non-managers, and those with more and less experience. No pattern was apparent, but it was likely that the sample size was too small to detect such if it indeed existed.

7.4.2 Differences Due to the Nature of Participating Parties in the Decision Process

Interviewees were asked to describe the kinds of group decisions that they commission or conduct. These were then linked to different outcomes during the interviews. The analysis of the interview transcripts revealed that the nature of the participating parties in the group decision process influenced the importance of different outcomes, although some outcomes were described as important or unimportant irrespective of the participating parties. The nature of participating parties fell mostly into five categories: political decision processes, internal decision processes, inter-agency decision processes, government-stakeholder decision processes and inter-stakeholder decision processes.

Political decision processes typically involved agencies supporting their Ministers in negotiation with their Cabinet colleagues, or with support parties. Though public servants supported these group-decision processes by providing information, it was rare that they had any influence over the decision-support process used, and therefore could not choose to use group model building. The study in this chapter was conducted from the perspective of group model building practitioners, and therefore situations, where the decision process cannot be influenced, are less useful for analysis; as one interviewee noted: “*We can’t control what they do.*”

Internal decision processes typically involved consensus decisions taken by peer groups within an agency. Where there was a disparity in the hierarchy, decisions tended to be taken by higher-ranked employees. These involved decisions on a course of action within a policy programme, or prioritisation and resource

allocation between policy programmes. These were typically convened by a member of that peer group, were either chaired by a group member or facilitated by an independent facilitator and required consensus agreement before completion—“*We were going to be locked in a room until we got this sorted.*” The exception to this pattern (mentioned by two interviewees) was when a higher-ranked employee convened a group process, and the group’s task was to arrive at a consensus recommendation—“*(The Deputy-Secretary) expects that we can come up with something...without having to bang our heads together.*” In these situations, the group included people of different rank.

Interagency decision processes involved employees of different agencies attempting to reach consensus agreement on a course of action, or on a joint recommendation to Ministers. Again, these were either chaired from within the group or involved an independent facilitator. Where Ministers had demanded a joint recommendation, processes were driven to a conclusion and often involved participants making difficult compromises. In contrast, processes to agree on a joint course of action often included alternatives to a negotiated agreement—agencies could continue to operate separately if a satisfactory negotiated agreement could not be obtained. Partial agreements or progress toward agreement were also considered acceptable outcomes. “*Sometimes it is about moving towards consensus, rather than achieving it.*” Interagency decision processes were seen as becoming more popular, with the creation of several secretariat units just to support and facilitate these discussions.

Government-stakeholder decision processes involved public servants working with stakeholders to reach an agreement. Typically public servants would begin the process with a tentative proposal, which would serve as the basis for negotiation—“*You never turn up with a blank sheet.*” Despite typically holding a monopoly or monopsony position, public servants were often disadvantaged by political or reputational drivers to achieve a negotiated agreement, or else the initiative would be considered a failure: “*There are usually win-wins, but they also know you’re not going to walk away.*” Alternately, where government was contributing funding to a negotiated agreement, it was stakeholders who had an incentive to reach an agreement or walk away empty-handed. One example was where government would fund the production of an educational programme if stakeholders and government could agree to the content of that programme.

Inter-stakeholder decision processes involved public servants acting as convenors to facilitate an agreement between other parties. These processes aimed to arrive at consensus agreements, such that government did not need to act as a referee between competing interests. These processes were seen as increasing in popularity as they helped government avoid making contentious decisions, and were believed by interviewees to lead to less discord between opposing parties.

Interview responses commonly related the importance of each outcome to a particular decision context (as described throughout Sect. 7.4.1). For each decision context, content analysis was used to provide a simple count of how often each outcome was mentioned as particularly important or unimportant (see Table 7.3). In several cases, multiple interviewees described an outcome as particularly important in a decision context, notably: consensus in internal decisions; mental model

Table 7.3 Important and unimportant outcomes for decisions involving different participating parties

Participating parties	Outcomes mentioned as important ^a	Outcomes mentioned as unimportant or detrimental ^a
Internal decisions	Consensus (3), insight (2), commitment to conclusions (2), power levelling (1)	None mentioned
Interagency decisions	Mental model alignment (4), further working together (2), consensus (1), insight (1)	Power levelling (2), insight (1)
Government-stakeholder decisions	Efficiency (3), commitment to conclusions (2), willingness to endorse (2), enduring mental model change (1), consensus (1)	None mentioned
Inter-stakeholder decisions	Communication quality (1), enduring mental model change (1), mental model alignment (1), efficiency (1)	Insight (1), views of non-participants (1)

^aNumbers in parentheses refer to the number of interviewees who mentioned this outcome as particularly important or unimportant for decisions involving these participating parties

alignment in inter-agency decisions; process efficiency in government-stakeholder decisions. This last finding is of particular interest as process efficiency was not an outcome investigated.

The importance of the different participating parties was not anticipated. It may have been useful to ask separate interview questions about each type of decision group, as this would have allowed a more thorough examination of the relationship between participating parties and outcome importance. This could form the basis for further study.

7.4.3 Statistical Analysis of Questionnaire Results

The written questionnaire was primarily used to verify the conclusions of the interviews, as explored in the discussion of each outcome above. However, a comparative analysis of the questionnaire results revealed some interesting findings.

All of the outcomes assessed were rated as equally or more important than the neutral response (a score of 4 on the 1–7 scale), and some significantly more important (Table 7.4). This suggests that all outcomes assessed were viewed as somewhat important, and several were viewed as very important. There was a wide range of responses—only “communication quality” and “commitment to conclusions” were always rated at 5 or higher.

Outcomes were then compared against each other. Some outcomes were viewed as more important than others. “Communication quality” and “commitment to conclusions” were both viewed as significantly more important than the other outcomes, and “insight” and “power levelling” were viewed as significantly less important. Significance was determined by comparing scores for that outcome with the overall mean score (see Sect. 7.3.5).

Table 7.4 Ratings of the importance of each outcome, relative to neutral and mean responses (n = 12)

Outcome	Mean	Range	Standard deviation	Difference from neutral score ^a	Difference from overall mean ^b
Commitment to conclusions	6.3	5–7	0.78	+2.3**	+1.0**
Communication quality	6.0	5–7	0.74	+2.0**	+0.7**
Consensus	6.0	4–7	0.95	+2.0**	+0.7*
Enduring mental model change	6.0	4–7	1.04	+2.0**	+0.7*
Mental model alignment	5.8	4–7	1.03	+1.8**	+0.6
Effective strategy implementation	5.7	3–7	1.30	+1.7**	+0.4
Enduring alignment	5.3	4–7	0.98	+1.3**	+0.1
Mental model change	4.4	3–7	1.31	+0.4	-0.9*
Persuasive content	4.3	2–7	1.71	+0.3	-1.0*
Power levelling	4.2	2–6	1.11	+0.2	-1.1**
Insight	4.0	2-6	1.35	+0.0	-1.3**

*p < 0.05

**p < 0.01

^a“Neutral score” is a score of 4 on a 1–7 numerical scale^bOverall mean = 5.3, difference rounded to one decimal place

The results from the numerical scale questions were also compared to each demographic field. The greatest differences were between the responses of managers (n = 6) and non-managers (n = 6) and between interviewees who had been in the public service for more than 5 years (n = 6), and those who have been in the public services for 5 years or fewer (n = 6).

There was no significant difference ($p > 0.10$) in the overall mean for managers (mean = 5.4) versus non-managers (mean = 5.2). Where the groups diverged was in their rating of the importance of persuasive content, this was ranked higher by managers than non-managers (5.0–3.5, $p < 0.05$). The authors had considered that non-managers might place a higher value on power levelling, as they had less institutional power, but there was no significant difference between the responses of managers and non-managers for this question (4.3–4.0, $p > 0.05$).

It had been considered that the outcomes valued by public servants might vary through their careers. There was no significant difference ($p > 0.10$) in the overall mean for those with more than 5 years of experience (mean = 5.2) and those with 5 years or fewer (mean = 5.4). However, experienced public servants were significantly more likely to value mental model alignment as a very important outcome (6.7–5.0, $p < 0.05$). In the interviews, more experienced public servants described “shared understanding” (possibly equivalent to mental model alignment) as critically important in group decision making.

7.5 Discussion and Conclusions

This chapter explores the views of New Zealand public servants who regularly commission or conduct group decision-making processes, and are thus potential group model building clients. This section describes some of the implications of these results. Section 7.5.1 identifies the public sector as a potentially growing market for group model building interventions. Section 7.5.2 explores the implications of these findings for group model building research. Moreover, in Sect. 7.5.3, the authors reflect on their own challenges with the findings of this research and their attitudes toward group model building and system dynamics modelling in general.

7.5.1 *A Growing Market?*

Many problems faced by public sector organisations are highly complex, with multiple actors, multiple stakeholders and conflicting outcomes (White, 2002). This makes public policy questions obvious targets for the problem-solving and problem-structuring applications of system dynamics (Rose & Haynes, 1999).

Two trends appear to be increasing the use of group-decision processes in the public sector. Instances of failed policy on issues that span organisational boundaries have driven demand for greater connectivity between agencies (Treisman, 2007). In New Zealand, this has manifested in calls for greater interagency coordination by the “Better Public Service” initiative (Scott & Bardach, 2018; Scott & Boyd, 2019; State Services Commission, 2011). Decisions based on consensus between stakeholders are thought to be more enduring than those arbitrated by government decision, leading to increased use of collaborative governance (Ansell & Gash, 2008; Emerson, Nabatchi, & Balogh, 2012; Newman et al., 2004)—in New Zealand this is being trialled through the consensus-based “Land and Water Forum” (Eppel, 2013). This growing field lacks agreed and accepted methods for supporting group decision making (Eden & Ackermann, 2013; Kim, 2008; Plottu & Plottu, 2011). The opportunity for group model building in the public sector appears large and is likely to be growing even larger (Bayley & French, 2008).

7.5.2 *Implications for Research*

If group model building has the potential to fill this opportunity, it is important to develop a sound empirical basis for the use and selection of group model building techniques. This empirical base should relate to the outcomes that potential clients are looking for.

The findings presented in this chapter suggest that, in most settings, public servants who commission group decision processes are primarily interested in

efficiently reaching an agreement between participants (consensus). Participants should be willing to endorse these agreements publically and to act on them when appropriate (*commitment to conclusions*). These are areas where there is strong evidence to support the effectiveness of group model building (Dwyer & Stave, 2008; Eskinasi et al., 2009; Huz, 1999; Rouwette, 2011; Scott et al., 2016a; Vennix et al., 1993; Vennix & Rouwette, 2000).

It is important that these agreements last. Government can move slowly, and commitment to these agreements must persist until the agreement can be put into action. While some group model building research evaluates enduring mental model change and alignment (Huz, 1999; Scott et al., 2013), further research is needed to evaluate the enduring agreement and the durability of commitment. It may be difficult to evaluate these outcomes due to problems of attribution (McCart & Rohrbaugh, 1989, 1995; Rohrbaugh, 1987; Shadish et al., 2001).

Public servants who commission group decision processes are also interested in exploring outcomes where there is limited evidence. They are concerned by the speed it takes to reach a decision, for which group model building literature can provide only indirect evidence (participants making comparisons to hypothetical meetings, Vennix et al., 1993; Vennix & Rouwette, 2000; Scott et al., 2016a). They are also interested in building trust and goodwill between participants that in turn fuels future cooperation, an area that requires evaluation in group model building literature.

The lukewarm attitudes to achieving new insights were somewhat surprising, as was the general lack of interest in policy quality. Interviewees often seemed so focussed on reaching any agreement that policy quality seemed a lesser concern. This is likely to be important as group model building practitioners think about how to describe the potential benefits of their techniques to potential customers. However, there may be a need for some caution in applying this finding, as discussed in Sect. 7.5.3.

The findings presented in this chapter demonstrate that different outcomes are valuable in different contexts. The group model building literature is currently missing practical guidance on how to vary the processes used to emphasise or enhance different outcomes. Three areas of literature provide helpful but incomplete clues in this regard: experimental studies on learning outcomes; a meta-analysis of the outcomes of qualitative versus quantitative processes; participants' rating of the contribution of different process elements. Each is explored in the following paragraphs.

Several experimental studies compare the presence or absence of group model building components and how these contribute to various outcomes. These studies have evaluated the importance of the presence of a facilitator (Borštnar, Kljajić, Škraba, Kofjač, & Rajkovič, 2011; Shields, 2001), the creation of causal loop diagrams (Fokkinga, Bleijenbergh, & Vennix, 2009) and the opportunity for group feedback and discussion (Borštnar et al., 2011; Škraba, Kljajić, & Borštnar, 2007; Škraba, Kljajić, & Leskovar, 2003). Unfortunately, these studies were conducted in experimental settings unlikely to be representative of real-world behaviours (Scott, 2014).

A meta-analysis found that quantitative modelling processes are associated with more commitment to conclusions, consensus and system change than qualitative only processes (Rouwette et al., 2002). However, this analysis did not compare like interventions, as the quantitative processes involved far greater time commitment by participants (Scott, 2014).

Other studies ask participants to rate the contribution of different components to the success of the intervention (Eskinasi et al., 2009; Scott et al., 2016a; Vennix et al., 1993; Vennix & Rouwette, 2000). There are limitations to the ability of individuals to describe their own learning (Doyle, 1997; Nisbett & Wilson, 1977), and, further, the study design did not allow each component to be related to individual outcomes. Further guidance is required to allow practitioners to tailor their practice toward particular outcomes.

Despite broad variance across different decision contexts, the findings presented in this chapter generally showed strong support for interpersonal outcomes relating to trust and agreement, and relatively less support for outcomes relating to policy quality. A similar distinction is evident in two contrasting perspectives of group model building sessions (Andersen et al., 2007). One perspective considers the model as an allegedly realistic representation of the external policy environment (“microworld”—Zagonel, 2002; “virtual world”—Serman, 2000). The second perspective considers the model as a socially constructed artefact for building trust and agreement (“boundary object”—Black & Andersen, 2012; Black, 2013; Franco, 2013; Scott et al., 2016b; Zagonel, 2002; “transitional object”—Ackerman et al. 2005). The findings presented in this chapter support the “boundary object” perspective as most representative of the views of public sector clients.

7.5.3 Reflections

Both authors have facilitated group model building interventions for public service clients in New Zealand. These include several published studies (Cavana et al., 2007, 2014; Cavana & Clifford, 2006; Cavana, Davies, Robson, & Wilson, 1999; Cavana & Tobias, 2008; Rees, Cavana, & Cumming, 2017; Scott, 2014, 2017, 2018; Scott et al., 2013, 2015, 2016a, 2016b; Scott, Cavana, & Cameron, 2016c; Tobias, Cavana, & Bloomfield, 2010), as well as countless others that were not part of formal research projects. Something that was striking to both authors is how little relation the findings presented in this chapter bore to our own experiences.

In our professional experience, clients most often present as having a policy problem for which they would like a method for arriving at the technically “best” solution. Rarely have we heard “we don’t much care about the quality of the policy recommendations, so long as everybody agrees quickly and maintains a commitment to that agreement over time.” Nevertheless, the research presented in this chapter suggests that the latter sentiment represents the most important outcomes in many cases.

Why might that be? How do we explain the variance between this research and our professional experience? We have no way of knowing, based on the data, but propose two possible explanations: preferences about future decisions may differ from those about the past, or clients may be (intentionally or unintentionally) misleading group model building practitioners. The following paragraphs briefly explore each explanation and how these hypotheses could be tested empirically.

The methodology used in this chapter asks research subjects to think about concrete past examples of group decision-making processes (through the interviews), as well as general, abstract or hypothetical group decision making (through the survey questionnaire). There was a reasonably strong agreement between both data sets. In a consulting context, the authors would typically ask at the outset what the client was hoping to achieve. It may be that the client's outcome preferences change over time. We speculate that clients may begin with the desire to use a group to reach a decision that is technically superior to one that any individual could arrive at on their own. Then, over time, they find that the process drags on, timelines slip, consensus is elusive or else individuals appear to reach agreement only to disagree later. It may be that group decisions tend to be frustrating. At the end of such a process, the client might then be more likely to reflect the sentiment from our data and claim that the most important outcomes are an efficient process that reaches a lasting agreement. The hypothesis that outcome preferences change over the course of a decision process could easily be tested; a researcher could gather data, using either the interview or survey questionnaire methods used in this chapter, both before and after the group model building intervention, and compare both data sets.

A second possible explanation is that the clients do not tell group model building practitioners the truth. This explanation seems counterproductive—when purchasing a consulting service, it seems likely that the service will be more beneficial if the client honestly communicates their needs. Nonetheless, we cannot rely on clients to always be rational (Munro, 2009). Instead, the variance between the data presented in this chapter and the authors' own professional experience may be explained by a social desirability bias. Social desirability bias describes the tendency of respondents to answer questions in a way that will be viewed favourably by others (Fisher, 1993; Grimm, 2010). Earlier in this chapter, we explain why the social desirability bias did not explain the results of this study, where most responses did not emphasise the importance of policy quality. Nevertheless, we can speculate that in a client-consultant discussion, social desirability bias may come into play. Clients may, in justifying an expenditure to themselves or others, prefer to describe outcomes that directly benefit society (better policies), than those that benefit a bureaucratic process (quicker, more stable decisions). To resolve this, we would need better ways of engaging with and understanding clients' actual needs. In general, it is difficult to fully control for social desirability bias, though several approaches have been proposed (Gittelman et al., 2015; Nederhof, 1985).

As practitioners as well as researchers, the findings in this chapter presented the authors with another challenge: if all clients want is to reach an agreement, does modelling quality matter? Like the previous question, this cannot be resolved using the data in this study. We consider two possible perspectives: that quality doesn't

matter, and instead practitioners should be looking to maximise any cognitive biases that tilt participants towards supporting conclusions derived from the model (see Scott, 2014, 2017, 2018); or alternatively, that group model building improves consensus and commitment to conclusions *because* participants can rationally be more confident about the quality of those conclusions through developing models that explain the behaviour of systems over time. Where the right decision is difficult to ascertain even in retrospect, the line can be blurry between a *bias* away from rational choice and a *heuristic* for increasing choice confidence (Tversky & Kahneman, 1974). We present no direct evidence to support either perspective but note that both authors also use system dynamics modelling when on their own, to analyse complicated problems. This suggests that, at least implicitly, we believe system dynamics modelling can improve decision quality for at least some problems types, and feel more confident in decisions that we have reached in this manner.

7.5.4 Conclusions

Collaborative governance is a complex field, and success can be measured in many ways, from the success of the processes involved (Carey & Harris, 2016) right through to the resulting change in society (Scott & Boyd, 2017a, 2017b). Similarly, the success of system dynamics modelling can be measured in ways equally diverse. The core focus of this book is to relate the two—to show how system dynamics modelling can and should contribute to the practice of collaborative governance.

This chapter contributes to that discussion by exploring which of these outcomes are most valued by public sector clients. Our data shows that, even within the public sector, there exist a broad range of different group-decision contexts with different aims. In general, the research subjects preferred consensus and commitment to conclusions to cognitive change, which suggests the “boundary object” perspective of group model building may be most relevant to their needs. Potential clients value most outcomes reported in group model building literature, but more research is required to compare the process efficiency of group model building with other methods. Further, this chapter raises broader questions about the “boundary object” versus “microworld” perspectives on the use of system dynamics in a group decision-making context: how system dynamics practitioners think about group model building, how clients think about group model building, and how these two “worlds” communicate.

References

- Ackermann, F., Eden, C., & Brown, I. (2005). *The practice of making strategy. A step-by-step guide*. London: Sage.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211.

- Andersen, D. F., Richardson, G. P., & Vennix, J. A. M. (1997). Group model building: Adding more science to the craft. *System Dynamics Review*, 13(2), 187–203.
- Andersen, D. F., Vennix, J. A. M., Richardson, G. P., & Rouwette, E. A. J. A. (2007). Group model building: Problem structuring, policy simulation and decision support. *Journal of the Operational Research Society*, 58(5), 691–694.
- Ansell, C., & Gash, A. (2008). Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory*, 18(4), 543–571.
- Bayley, C., & French, S. (2008). Designing a participatory process for stakeholder involvement in a societal decision. *Group Decision and Negotiation*, 17(3), 195–210.
- Bentham, J. B., & de Visscher, A. G. (1994). Systems thinking and its influence on operational culture. In *Proceedings of the 1994 International System Dynamics Conference*. System Dynamics Society, Albany, NY.
- Black, L. J., & Andersen, D. F. (2012). Using visual representations as boundary objects to resolve conflicts in collaborative model-building approaches. *Systems Research and Behavioral Science*, 29, 194–208.
- Blaikie, N. (1993). *Approaches to social enquiry*. Cambridge: Polity.
- Borštnar, M. K., Kljajić, M., Škraba, A., Kofjač, D., & Rajkovič, V. (2011). The relevance of facilitation in group decision making supported by a simulation model. *System Dynamics Review*, 27(3), 270–293.
- Black, L. J. (2013). When visuals are boundary objects in system dynamics work. *System Dynamics Review* 29 (2):70–86.
- Carey, G., & Harris, P. (2016). Developing management practices to support joined-up governance. *Australian Journal of Public Administration*, 75(1), 112–118.
- Cavana, R. Y., Boyd, D. M., & Taylor, R. J. (2007). A systems thinking study of retention and recruitment issues for the New Zealand army electronic technician trade group. *Systems Research and Behavioral Science*, 24(2), 201–216.
- Cavana, R. Y., & Clifford, L. V. (2006). Demonstrating the utility of system dynamics for public policy analysis in New Zealand: The case of excise tax policy on tobacco. *System Dynamics Review*, 22(4), 321–348.
- Cavana, R. Y., Davies, P. K., Robson, R. M., & Wilson, K. J. (1999). Drivers of quality in health services: Different worldviews of clinicians and policy managers revealed. *System Dynamics Review*, 15(3), 331–340.
- Cavana, R. Y., Delahaye, B. L., & Sekaran, U. (2001). *Applied business research: Qualitative and quantitative methods*. Brisbane: Wiley.
- Cavana, R. Y., Smith, T., Scott, R. J., & O'Connor, S. (2014). Causal mapping of the New Zealand natural resources sector system. In *Proceedings of the 2014 International System Dynamics Conference*. System Dynamics Society, Albany, NY.
- Cavana, R. Y., & Tobias, M. (2008). Integrative system dynamics: Analysis of policy options for tobacco control in New Zealand. *Systems Research and Behavioral Science*, 25(5), 675–694.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks: Sage.
- Coyle, R. G. (2000). Qualitative and quantitative modelling in system dynamics: Some research questions. *System Dynamics Review*, 16(3), 225–244.
- Craik, K. J. W. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.
- Doyle, J. K. (1997). The cognitive psychology of systems thinking. *System Dynamics Review*, 13, 253–265.
- Doyle, D. K., & Ford, D. N. (1998). Mental model concepts for system dynamics research. *System Dynamics Review*, 14(1), 3–29.
- Dwyer, M., & Stave, K. (2008). Group model building wins: The results of a comparative analysis. In *Proceedings of the 2012 International System Dynamics Conference*. System Dynamics Society, Albany, NY.

- Eden, C. E., & Ackermann, F. (2004). Use of “soft OR” models by clients—What do they want from them? In M. Pidd (Ed.), *Systems modelling theory and practice* (pp. 146–163). Chichester: Wiley.
- Eden, C. E., & Ackermann, F. (2013). ‘Joined-up’ policy-making: Group decision and negotiation practice. *Group Decision and Negotiation*, 23(6), 1385–1401. <https://doi.org/10.1007/s10726-013-9375-1>
- Edwards, A. (1957). *The social desirability variable in personality assessment and research*. New York: The Dryden Press.
- Emerson, K., Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. *Journal of Public Administration Research and Theory*, 22(1), 1–29.
- Eppel, E. (2013). *Collaborative governance: Framing New Zealand practice*. Institute for Governance and Policy Studies Working Paper. Victoria University of Wellington, New Zealand.
- Eskinasi, M., Rouwette, E., & Vennix, J. (2009). Simulating urban transformation in Haaglanden, the Netherlands. *System Dynamics Review*, 25(3), 182–206.
- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2), 303–315.
- Fokkinga, B., Bleijenbergh, I., & Vennix, J. (2009). Group model building evaluation in single cases: A method to assess changes in mental models. In *Proceedings of the 2009 International Conference of the System Dynamics Society*. System Dynamics Society, Albany, NY.
- Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., & Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health*, 25(10), 1229–1245.
- Franco, L., A. (2013). Rethinking Soft OR interventions: Models as boundary objects. *European Journal of Operational Research*, 231(3), 720–733.
- Gittelman, S., Lange, V., Cook, W. A., Frede, S. M., Lavrakas, P. J., Pierce, C., & Thomas, R. K. (2015). Accounting for social-desirability bias in survey sampling: A model for predicting and calibrating the direction and magnitude of social-desirability bias. *Journal of Advertising Research*, 55(3), 242–254.
- Green, J., & Thorogood, N. (2009). *Qualitative methods for health research* (2nd ed.). Thousand Oaks: Sage.
- Greenberger, M., Crenson, M. A., & Crissey, B. L. (1976). *Models in the policy process: Public decision making in the computer era*. New York: Russell Sage Foundation.
- Grimm, P. (2010). Social desirability bias. In: Sheth JN, Malhotra NK, eds. *Wiley international encyclopedia of marketing*. John Wiley and Sons Ltd. <https://doi.org/10.1002/9781444316568.wiem02057>.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59–82.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading: Addison-Wesley.
- Hovmand, P. S., Andersen, D. F., Rouwette, E., Richardson, G. P., Rux, K., & Calhoun, A. (2012). Group model-building ‘scripts’ as a collaborative planning tool. *Systems Research and Behavioral Science*, 29(2), 179–193.
- Huz, S. (1999). *Alignment from group model building for systems thinking: Measurement and evaluation from a public policy setting*. New York: State University New York.
- Huz, S., Andersen, D. F., Richardson, G. P., & Boothroyd, R. (1997). A framework for evaluating systems thinking interventions; an experimental approach to mental health system change. *System Dynamics Review*, 13(2), 149–169.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.
- Jones, N. A., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental models: An interdisciplinary synthesis of theory and methods. *Ecology and Society*, 16(1), 46.
- Kim, J. (2008). A model and case for supporting participatory public decision making in e-democracy. *Group Decision and Negotiation*, 17(3), 179–193.

- Kolfschoten, G. L., & Rouwette, E. A. J. A. (2006). Choice criteria for facilitation techniques. In R. O. Briggs & J. F. Nunamaker (Eds.), *Monograph of the HICSS-39 symposium on case and field studies of collaboration* (pp. 35–44). Hawaii: Hawaii International Conference of System Sciences.
- Kvale, S., & Brinkman, S. (2008). *InterViews* (2nd ed.). Thousand Oaks: Sage.
- Luna-Reyes, L. F., Martinez-Moyano, I. J., Pardo, T. A., Cresswell, A. M., Andersen, D. F., & Richardson, G. P. (2006). Anatomy of a group model-building intervention: Building dynamic theory from case study research. *System Dynamics Review*, 22(4), 291–320.
- Martinez-Moyano, I. J., & Richardson, G. P. (2013). Best practices in system dynamics modeling. *System Dynamics Review*, 29(2), 102–123.
- Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 11(3). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/1428/3027>
- McCardle-Keurentjes, M. H., Rouwette, E. A. J. A., & Vennix, J. A. M. (2008). Effectiveness of group model building in discovering hidden profiles in strategic decision-making. In *Proceedings of the 2008 International System Dynamics Conference*. System Dynamics Society, Albany, NY.
- McCart, A., & Rohrbaugh, J. (1989). Evaluating group decision support effectiveness: A performance study of decision conferencing. *Decision Support Systems*, 5, 243–253.
- McCart, A. T., & Rohrbaugh, J. (1995). Managerial openness to change and the introduction of GDSS. *Organizational Science*, 6, 569–584.
- Mingers, J., & White, L. (2010). A review of the recent contribution of systems thinking to operational research and management science. *European Journal of Operational Research*, 207, 1147–1161.
- Moray, N. (1998). Identifying mental models of complex human-machine systems. *International Journal of Industrial Ergonomics*, 22, 293–297.
- Munro, A. (2009). *Bounded rationality and public policy: A perspective from behavioural economics*. New York: Springer Science & Business Media.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15(3), 263–280.
- Newman, J., Barnes, M., Sullivan, H., & Knops, A. (2004). Public participation and collaborative governance. *Journal of Social Policy*, 33(2), 203–223.
- Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Noble, C. H. (1999). The eclectic roots of strategy implementation research. *Journal of Business Research*, 45, 119–134.
- Orne, M. (1962). On the social psychology of the psychology experiment. *American Psychologist*, 17, 776–783.
- Plottu, B., & Plottu, E. (2011). Participatory evaluation: The virtues for public governance, the constraints on implementation. *Group Decision and Negotiation*, 20(6), 805–824.
- Rees, D., Cavana, R. Y., & Cumming, J. (2017). Using cognitive and causal modelling to develop a theoretical framework for implementing innovative practices in primary healthcare management in New Zealand. *Health Syst (Basingstoke)*, 7(1), 51–65. <https://doi.org/10.1057/s41306-017-0029-4>. (On-line first).
- Richardson, G. P., & Andersen, D. F. (1995). Teamwork in group model building. *System Dynamics Review*, 11(2), 113–137.
- Richardson, G. P., Andersen, D. F., Maxwell, T. A., & Stewart, T. R. (1994). Foundations of mental model research. In *Proceedings of the 1994 International System Dynamics Conference*. System Dynamics Society, Chestnut Hill.
- Ritchie, J., Lewis, J., & Elam, G. (2003). Designing and selecting samples. In J. Ritchie & J. Lewis (Eds.), *Qualitative research practice. A guide for social science students and researchers* (pp. 77–108). Thousand Oaks: Sage.

- Rohrbaugh, J. (1987). Assessing the effectiveness of expert teams. In J. L. Munpower, L. D. Phillips, O. Renn, & V. R. R. Uppuluri (Eds.), *Expert judgment and expert systems* (Vol. 35, pp. 251–267). Berlin: Springer.
- Rose, J., & Haynes, M. (1999). A soft systems approach to the evaluation of complex interventions in the public sector. *Journal of Applied Management Studies*, 8(2), 199–216.
- Rouwette, E. A. J. A. (2003). *Group model building as mutual persuasion*. Nijmegen: Wolf Legal Publishers.
- 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., Korzilius, H., Vennix, J. A., & Jacobs, E. (2011). Modeling as persuasion: The impact of group model building on attitudes and behavior. *System Dynamics Review*, 27(1), 1–21.
- Rouwette, E. A. J. A., & Vennix, J. A. M. (2006). System dynamics and organizational interventions. *Systems Research and Behavioral Science*, 23(4), 451–466.
- Rouwette, E. A. J. A., & Vennix, J. A. M. (2011). Group model building. In *Complex systems in finance and econometrics* (pp. 484–496). New York: Springer.
- Rouwette, E. A., Vennix, J. A., & Felling, A. J. (2009). On evaluating the performance of problem structuring methods: An attempt at formulating a conceptual model. *Group Decision and Negotiation*, 18(6), 567–587.
- Rouwette, E. A. J. A., Vennix, J. A. M., & Van Mullekom, T. (2002). Group model building effectiveness: A review of assessment studies. *System Dynamics Review*, 18(1), 5–45.
- Samuelson, P. (1938). A note on the pure theory of consumers' behaviour. *Economica*, 5(17), 61–71.
- Scott, R. J. (2014). *Group model building and mental model change*. PhD Thesis. University of Queensland.
- Scott, R. J. (2017). Explaining how group model building supports enduring agreement. *Journal of Management & Organization*, 25(6), 783–806. <https://doi.org/10.1017/jmo.2017.12>
- Scott, R. J. (2018). *Group model building: Using systems dynamics modelling to achieve enduring agreement*. Melbourne: Springer.
- Scott, R. J., & Bardach, E. A. (2018). Comparison of management adaptations for joined-up government: Lessons from New Zealand. *Australian Journal of Public Administration*, 78(2), 191–212. <https://doi.org/10.1111/1467-8500.12348>
- Scott, R. J., & Boyd, R. (2017a). *Interagency performance targets: A case study of New Zealand's results programme*. New York: Business of Government Series, IBM.
- Scott, R. J., & Boyd, R. (2017b). Joined-up for what? Response to Carey and Harris on adaptive collaboration. *Australian Journal of Public Administration*, 76(1), 138–144.
- Scott, R. J., & Boyd, R. (2019). Determined to succeed: Can goal commitment sustain inter-agency collaboration? *Public Policy and Administration*. Accepted for publication. <https://doi.org/10.1177/0952076720905002>.
- Scott, R. J., Cameron, D., & Cavana, R. Y. (2015). A meta-analysis of quantitative evidence for group model building. *European Journal of Operations Research*, 249(3), 908–918.
- Scott, R. J., Cavana, R. Y., & Cameron, D. (2013). Evaluating immediate and long-term impacts of qualitative group model building workshops on participants' mental models. *System Dynamics Review*, 29(4), 216–236.
- Scott, R. J., Cavana, R. Y., & Cameron, D. (2016a). Group model building and strategy implementation. *The Journal of the Operational Research Society*, 66(6), 1023–1034.
- Scott, R. J., Cavana, R. Y., & Cameron, D. (2016b). Mental model change among group model building participants. *Systems Research and Behavioural Science*, 33(1), 100–118.
- Scott, R. J., Cavana, R. Y., & Cameron, D. (2016c). Client perceptions of reported outcomes of group model building in the New Zealand public sector. *Group Decision and Negotiation*, 25(1), 77–101.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Wadsworth: Cengage Learning.

- Shields, M. (2001). An experimental investigation comparing the effects of case study, management flight simulator and facilitation of these methods on mental model development in a group setting. In *Proceedings of the 2001 International Conference of the System Dynamics Society*. System Dynamics Society, Albany, NY.
- Skivington, J. E., & Daft, R. L. (1991). A study of organizational “framework” and “process” modalities for the implementation of business-level strategic decisions. *Journal of Management Studies*, 28, 45–68.
- State Services Commission. (2011) *Better public services advisory group report*. New Zealand Government.
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347), 730–737.
- Sterman, J. D. (2000). *Business dynamics—systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park: Sage.
- Škraba, A., Kljajić, M., & Borštnar, M. K. (2007). The role of information feedback in the management group decision-making process applying system dynamics models. *Group Decision and Negotiation*, 16(1), 77–95.
- Škraba, A., Kljajić, M., & Leskovar, R. (2003). Group exploration of system dynamics models—Is there a place for a feedback loop in the decision process? *System Dynamics Review*, 19(3), 243–263.
- Thomas, P., & Carswell, L. (2000). Learning through collaboration in a distributed education environment. *Educational Technology and Society*, 3(3), 1–15.
- Thompson, J. P. (2009). *How and under what conditions client learn in system dynamics consulting engagements*. PhD Thesis, Strathclyde Business School, Glasgow.
- Tobias, M., Cavana, R. Y., & Bloomfield, A. (2010). Application of a system dynamics model to inform investment in smoking cessation services in New Zealand. *American Journal of Public Health*, 100(7), 1274–1281.
- Treisman, D. (2007). *The architecture of government*. New York: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.
- Van Nistelrooij, L. P. J., Rouwette, E. A. J. A., Vestijnen, I., & Vennix, J. A. M. (2012). Powerlevelling as an effect of group model building. In *Proceedings of the 2012 International System Dynamics Conference*. System Dynamics Society, Albany, NY.
- Vennix, J. A. M. (1995). Building consensus in strategic decision making: System dynamics as a group support system. *Group Decision and Negotiation*, 4(4), 335–355.
- Vennix, J. A. M. (1996). *Group model building: Facilitating team learning using system dynamics*. Chichester: Wiley.
- Vennix, J. A. M., & Rouwette, E. A. J. A. (2000). Group model building. What does the client think of it now? In *Proceedings of the 2000 International System Dynamics Conference*. System Dynamics Society, Albany, NY.
- 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*. System Dynamics Society, Albany, NY.
- White, L. (2002). Size matters: Large group methods and the process of operational research. *Journal of the Operational Research Society*, 53(2), 149–160.
- Zagonel, A. A. (2002). Model conceptualization in group model building: A review of the literature exploring the tension between representing reality and negotiating a social order. In *Proceedings of the 2002 International System Dynamics Conference*. System Dynamics Society, Albany, NY.
- Zagonel, A. A., Rohrbaugh, J., Richardson, G. P., & Andersen, D. F. (2004). Using simulation models to address “what if” questions about welfare reform. *Journal of Policy Analysis and Management*, 23(4), 890–901.

Rodney Scott is an Adjunct Associate Professor at the University of New South Wales, Australia. He was a 2018 Fellow in Practice at the Blavatnik School of Government (Oxford University), and 2017 Innovations Fellow at the Kennedy School of Government (Harvard University). He has authored over 60 publications on system dynamics and public administration, including “Group Model Building: Using Systems Dynamics to Achieve Enduring Agreement” (Springer 2018).

Rodney is also the Chief Policy Advisor for the State Services Commission of New Zealand, previously holding senior policy and management roles across the New Zealand public sector. Current projects include a new Public Service Act for New Zealand and a “Public Service Fale” to support public sector reform in Pacific Island Countries.

Robert Y. Cavana, MCom (Econ), PhD (System Dynamics), ANZAMF, is a Reader in Systems Science with the Wellington School of Business and Government, Victoria University of Wellington, New Zealand. Previously he was the Corporate Economist with NZ Railways Corporation. Bob is a past President of the NZ Operational Research Society, a past Vice-President of the International System Dynamics Society and a former Managing Editor of System Dynamics Review. He is a co-author of Systems Thinking, System Dynamics: Managing Change and Complexity 2nd ed (Pearson Education, Auckland, 2007) and Applied Business Research: Qualitative and Quantitative Methods (Wiley, Brisbane, 2001; Chinese edition 2004). He recently co-edited the MDPI “Systems” online special issue and book on “Systems Education for a Sustainable Planet”.