Group model building: adding more science to the craft

David F. Andersen[a]*, George P. Richardson[a] and Jac A. M. Vennix[b]

David Andersen is Professor of Public Administration, Public Policy, and Information Science at the University at Albany, State University of New York. He holds an A.B. in Mathematics and Social Sciences from Dartmouth College and a Ph.D. in Management from Massachusetts Institute of Technology's Sloan School specializing in System Dynamics.

George P. Richardson is Professor at the Rockefeller College of Public Affairs and Policy in the State University of New York at Albany. He is the author of *Introduction to System Dynamics Modeling with DYNAMO* (1981), *Feedback Thought in Social Science and Systems Theory* (1991) and edited the collection *Modeling for Management: Simulation in Support of Systems Thinking* (1996).

Jac Vennix is Associate Professor of research methodology at Nijmegen University. He holds a Ph.D. from Nijmegen University and his current research interests focus on methods for group model-building and empirical evaluation of its effectiveness. His most recent publication is a book on *Group Model-Building*.

Abstract

This article discusses the issue of making group model building interventions more of a science than an art by outlining a number of requirements of a research program. Important elements that are discussed are the various goals of group model building interventions and the components and scripts of an intervention. Then the problem of theory development is discussed, together with a number of hypotheses which the authors suggest need more investigation. The article also discusses issues related to the selection of an appropriate research design, as well as a number of thorny measurement problems. © 1997 by John Wiley & Sons, Ltd. Syst. Dyn. Rev. **13**, 187–201, 1997

(No. of Figures: 0   No. of Tables: 0   No. of Refs: 36)

The nature and scope of the problem

The articles in this special issue have presented a number of standard procedures of group model-building approaches, as well as results from empirical studies aiming to identify the success of modeling with client groups. From these articles it becomes clear that group model building is still more art than science. Research on the effects of group model building is scarce; it focuses on a wide variety of outcomes and variables, and research designs differ quite considerably. Instead of a solid research program creating replicable[1] and cumulative results, we seem to have series of presumptions and hunches being repeated in a descriptive literature with little empirical evidence, certainly lacking any sense of competing propositions or refutability of the claims being made (most often by the practitioners who are using the system-intervention approach). The norm for research seems to be to posit an intuitively grounded hunch about what will work with a group and then to design a facilitated conference process around that hunch. If the interventions are successful (in the sense that paying clients like them and are willing to fund them being repeated), then the hunch is substantiated and the best intuitive practice continues. It seems that legends about what is working grow up around these interventions in what can only be described as superstitious behavior.

An example may help here. Those of us who work in the field of system dynamics have long held the belief that one of the great benefits of our modeling craft is that

[a]Nelson A. Rockefeller College of Public Affairs and Policy, The University at Albany, State University of New York, Albany NY 12222
[b]Department of Policy Studies, Nijmegen University, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands. Email: J.Vennix@maw.kun.nl
*Corresponding Author.

System Dynamics Review Vol. 13, No. 2, (Summer 1997): 187–201   *Received November 1996*
 CCC 0883-7066/97/020187–15 $17.50   *Accepted February 1997*

it can make explicit the implicit mental models of our clients. Somehow, we have come to believe that explicit and discussible mental models (whatever they would turn out to be) are better, more accurate, and more useful than the old-fashioned, implicit mental models. We even have some more or less precise ideas about what these mental models are. Somehow, we came to believe that decision makers had an implicit understanding of system structures loaded in their cognitive capacities and that the causal structures that we build into formal simulation models could improve on and make more precise these implicit mental models. As a result policy makers would be better able to manage a complex dynamic system.

Several empirical studies have revealed that this is not the case. When given funding from the National Science Foundation to investigate more precisely the nature of managerial mental models, Maxwell *et al.* (1994; see also Richardson *et al.* 1994) could not find any strong evidence that decision makers actually carried around implicit or explicit elaborated causal structures in their minds. What the researchers thought mental models were could not be readily observed or measured. Indeed, when they gave decision makers explicit training in what they had presumed mental models to be, they found that this training had no statistical impact on their ability to manage a dynamic system.

However, they did discover that when decision makers were given what they called “causal chunks” or simple “If you do thus and so the outcome will be such and such” type of statements, decision makers were able better to manage a dynamic system. In short, it was thought that they were giving decision makers deeper and more elaborate understandings of system causal structure (a hunch that had developed into superstitious behavior), when what really seemed to matter was that the modeling team leave decision makers with highly chunked strategic insights. It took the authors a year and quite a little effort to disabuse themselves of their initial superstitious hunch. Other empirical studies largely confirm these results. Vennix (1990) and Verburgh (1994) found that even after extensive training in modeling, although individual learning occurred, no real improvement of participants’ mental models, in terms of entertaining more feedback loops or more elaborate causal relationships, could be established.

As these examples reveal, empirical research into the success of group model building is dearly needed. The focus of a research program into the role of facilitated group model building should, in our opinion, center on making these interventions more successful. Specifically, such a research program should work to identify what aspects or combinations of activities within facilitated group model-building conferences lead to more successful outcomes for individual clients, groups of clients, or the host organization. This type of a broadly focused research program gets stalled right from the start because we do not have a good description of what is being done in the diversity of activities that call themselves strategic or systems thinking conferences. In addition, the sets of outcomes are many and diverse, often not clearly specified by researchers, and are less often measured in any sort of a rigorous fashion. The goal of this paper is to sketch the basic requirements and the outlines of such a research program on group model building. This framework might be used in the future to align research efforts, to build on each others' work and to add more science to the craft of group model building.

We begin by making explicit our assumptions about what constitutes a solid research program. Next we will turn to a number of important subjects for research and a number of thorny research problems.

Properties of a good research program

If we, as a research community, intend to move away from superstitious reinforcement of *a priori* hunches, then we will have to move toward a systematic research program that has several important properties.

Replicable

We must begin to replicate each others' work. For example, in a dissertation supervised by Andersen, Richardson and Stewart (Maxwell 1995) on the effect of cognitive style on mental models, it seemed that the literature is full of unreplicated findings that appear to contradict each other. But the problem with replication is that those of us who are working in facilitated group model-building conferences lack a common description or taxonomy of approaches for what we do (and no-one gets tenure for replicating the work of others). Hence, we are all doomed to publish small studies about what we do in our own small domains with little hope of replicating what our colleagues are doing. How can we learn from each other?

Cumulative

Closely tied to the notion of replication is the idea that this research program can be cumulative. We need to get better at citing each others' results and building on those results. Again a clear problem arises because we now lack a common vocabulary for describing how our interventions are the same or are different, both in terms of process and in terms of intended and actual outcomes. Since we now lack a common set of conceptual dimensions by which we can describe the diversity of our practice, we lack a solid ability to build on each others' results.

Refutable

Perhaps the most important aspect of a scientific research program is the ability of the research to unseat accepted hunches about what is really working and to truly challenge our notions of what is or is not working in our interventions. We must be open to the idea that our cherished notions about what is important are simply not correct. In principle, a research program must be able to refute our various sacred cows, but, in order to create a refutable research program, we must be able to observe a wide enough variety of types of interventions to sort out whether or not our pet theories about what works are really correct. However, if we always use our preferred techniques and approaches (out of a strong belief that they are what is best and will produce best outcomes), then our research program will lack the counter-factual examples that can refute those same pet theories.

By definition a good research program should challenge some of the central beliefs and premises of those who are practicing group-facilitated strategic or systems thinking. Since many of those practicing the craft are those who would do the research, we have to ask seriously how much do we really want to know what works and what doesn't?

Goals of group model building

If we are truly serious in this research program about making our facilitated interventions more successful, then we must become crystal clear about the intended outcomes of our interventions. At what level do we believe that our interventions are making a difference — for individuals, for the group that we are working for, or for the host organization(s)? These are quite different levels of outcomes (and units of analysis in empirical research) and, unless we can be clear about what we are intending to accomplish through our interventions, then we will have no hope of measuring and describing our successes in a way that can lead to insights about what will be more successful in the future.

One goal that has been specified at the individual level is learning, i.e., improving mental models. More specifically, we may believe that our group modeling efforts should help individual participants gain more insight into the structure and behavior of a system. As stated above, empirical research carried out thus far seems to contradict the notion of the potential for mental-model improvement when such improvement is defined as increasing participants' ability to correctly perceive relationships between system structure and system behavior at a detailed level. Given these research results, we should consider discarding the hypothesis that group model building will lead to better mental models of the details of system structure and system behavior. A second goal which has recently been posited is a change of attitude towards a proposed policy. The goal of many organizational interventions is to change people's behavior. According to Ajzen's well-known theory of planned behavior (Ajzen, 1991), one necessary prerequisite for behavioral alteration is a change of attitude. Exploratory research shows that group model building can aid in bringing about a change in attitudes towards a proposed policy (Vennix *et al.* 1996).

At the group level the goals of group model building have been described as:

1. mental model alignment (Huz *et al.* 1997);
2. creating agreement (consensus) about a policy or decision;
3. generating commitment with a decision (Rohrbaugh 1992; Senge 1990a; Vennix et al. 1993; Winch 1993).

At the organizational level the goals of group model building have been specified as system process change (Do we do things differently?) and systems outcome change (Are customers or clients impacted differently?) (see Cavaleri and Sterman 1997).

At first sight, the potential outcomes of a group model-building project look quite diverse. As stated by Cavaleri and Sterman, the most important goal is system improvement. Does this mean that we can ignore the other goals? No, because we have to take into account that some of these goals may be interrelated. For instance, if there is no agreement on how to do things and different people in an organization do things differently, this may negatively impact the performance of the system. A change in behavior may be required, which in turn will result in improving performance of the system. Alternatively, if people do not understand the effects they bring about by certain decisions or behavior, then group model building may be helpful to create this awareness, thereby altering a person's behavior and improving system outcome.

The intervention: components and scripts

Thus far we have been talking about "intervention" or "group model building" as if it were clear what this is. One has to recognize that the intervention itself is a complex system, consisting of a number of elements together and, in interaction, producing the outcome of the intervention. One prerequisite is the formulation of theories and hypotheses is a more precise description of what happens in behavioral terms when we as consultants and facilitators are locked up in those rooms with our clients. Do we all use similar brainstorming techniques? Do we use computing support? If so, how and when? Once we can comprehensively classify the diversity of our practice, we can begin to probe more deeply into what matters in that practice. We see two major components to this descriptive process:

1. identifying the basic components of a group model-building project;
2. the behavioral description of group model-building components.

Components of group model building

We should draw up some sort of a taxonomy of group-facilitated meetings that we all agree is broad enough to encompass what we all do. All of these conferences seem to share a number of common components (or at least dimensions along which they differ systematically). Here we make a distinction between three stages in a group model-building intervention: pre-meeting activities, the actual meetings and the after-care or follow-up activities. In addition, the context in which the project takes place may be important.

With regard to pre-meeting activities a distinction can be made between (a) contracting and client–consultant relationship, (b) participants and (c) contacts with participants prior to actual meetings. For each of these categories, our interventions can vary in a number of important ways.

- Pre-project client–consultant relationship:
  - How contact and entry and contracting are handled (e.g., who initiated the contact, modeler or client?).
  - Type of problem being addressed and goals of project.

- Participants:
  - Size and composition of team: who is involved and who decided on this?
  - Level of top management support.

- Contacts with participants:
  - Were pre-meeting interviews scheduled?
  - What introduction to system dynamics is given?

When it comes to the meetings themselves we need to distinguish a number of components: (a) actual composition of group and meetings, (b) modeling procedure, (c) aspects of facilitation, and (d) meeting logistics.

- Meetings and participants:
  - Participants (number and characteristics of attendees);
  - Meetings: number of meetings and average duration;

- How much (and what kind of) work was done off site and how much with the group?;
- Participant satisfaction with process and outcome.

• Modeling procedure:
  - What type and process of modeling was used (flow diagrams or causal loop diagrams and quantitative modeling and simulation) and how were policies assessed?
  - Support: supporting techniques used in the process.
  - Was a preliminary model used or did the meetings start from scratch?
  - Were questionnaires/workbooks used?

• Facilitation aspects:
  - Number of facilitators and their roles.
  - Degree to which facilitator steers the discussions.

• Meeting logistics:
  - Were meetings held away from the office?
  - Room design and layout.

Finally, we also have to recognize that no group model-building intervention takes place in a vacuum. There will always be contextual variables to take into account, e.g., type of organization, organization culture and history.

What are the "scripts" for these components?

In their contribution to this special issue Andersen and Richardson (1997) use the term "scripts" to stand for small behavioral descriptions of pieces of a facilitated group exercise that move a group forward in a systems thinking intervention. We suspect that all practitioners script their facilitations much like an improvisational theater group plans a vignette. What are these scripts and how are they the same or different across varying systems thinking approaches? Is it too much to hope that we could devise a coding scheme for scripts in group-facilitated meetings analogous to the coding scheme employed by choreographers to record the complex dance patterns of many performers in a full ballet?

We have also become interested in the question of how to string a set of these scripts together to make a complete experience for our clients. It is difficult to execute a script fully because the exigencies of each conference dictate some amount of improvisation and *ad hoc* adjustment. These processes are still the most artful in our work and cry out for more systematic description and analysis.

Undoubtedly, we all have quite different theoretical bases to predict what about our behaviors in facilitated meetings makes a difference, but, before we get to these very interesting questions, we need to generate a commonly agreed, and we suspect behaviorally grounded, description of what we actually do in the meetings.

Theory development

It seems to us that any systematic research program must first solve the big problems of describing what we do and describing intended and actual outcomes for our clients and their host organizations. Only when we have constructed this basic groundwork, can we move onto cumulative and replicable research that is capable of challenging in a refutable fashion some of our pet theories and hunches.

Describing the components of group-facilitated meetings, i.e., descriptive research, is not sufficient, nor is empirical research that is not grounded in theories. We have have to strive for explanatory theories, i.e., sets of hypotheses that explain why a particular result will be produced by this particular intervention.

Theories on the potential effects of group model building will differ across the goals to be accomplished. For instance, a theory explaining why consensus will be reached more easily in group model building (as opposed to ordinary meetings) will differ from a theory explaining why peoples' attitudes have changed. In both cases the intervention (i.e., independent variable) might be the same, but the dependent variable (as well as intermediate variables) will differ considerably. We do have to point out that matters can be further complicated, because the outcome of a group model-building process may differ considerably from what was expected at the outset. This does not necessarily have to be caused by a faulty theory; rather it results from the difficulty to diagnose readily and fully a client's problem in advance of the group model-building intervention. Sometimes the "real" problem does not emerge until the group model-building process is underway (cf. Rosenhead 1989; Vennix 1996). In those cases, the outcome can only be explained in hindsight with all the well-known pitfalls.

Although there are no clear-cut theories on the potential effects of group model-building interventions, relevant literature, exploratory research, and personal experience have given rise to some propositions that we would dearly like to see investigated more carefully and completely.

Systems thinking sub-components hypothesis: What matters is that practitioners use causal loops, or word-and-arrow diagrams, or fishbone diagrams, or computer support, etc. (see, for instance, McCartt and Rohrbaugh 1989).

Facilitation hypothesis: Is success contingent on the present of a facilitator, or can ordinarily accomplished teams have successful conferences given the right scripts? (e.g., Conyne and Rapin 1977; Maier and Thurber 1969; Phillips and Phillips 1993; Vennix *et al.* 1993; Vennix 1995).

Group structure hypothesis: What matters is that top management is together with the "doers" for an extended period (see Akkermans 1995; Akkermans and Vennix 1997).

Chunking hypothesis: What matters is getting big chunks of insight — the details that lead up to the insights are largely means to acquire group confidence and are forgotten (cf. Andersen *et al.* 1994).

The Gifted Practitioner Hypothesis: Can any person learn how to effectively facilitate groups or is it just a matter of talent? Or stated differently: some persons will always be successful, no matter what type of intervention, while others will mostly be unsuccessful, no matter how good their intervention method and tools.

Group communication climate hypothesis: What really matters is the quality of the communication process in a group (cf. Schein 1987; Argyris 1990).

Hawthorne effect hypothesis: What matters is that something special or out of the ordinary was done with the group and the problem at hand — it really does not matter much what the special process is.

Research design

Of course, even if we can describe the diversity of our practice and more precisely define desired outcomes, a number of troubling research questions will remain. One of these is related to appropriate research design. In the social science research methodology various purposes of research are identified (see, for instance, Babbie 1995): exploration, description, and explanation. In the previous section we have already emphasized the importance of descriptive research into the components and scripts of a group model-building intervention. We have also indicated that, although we have a number of hypotheses we would like to see tested, there is a need for systematic exploratory research, leading to the formulation of new, better, and systematically interrelated hypotheses into a systematic theory of group model building. Exploratory research can also be helpful to develop and try suitable methods for assessment. Explanatory research is required to answer questions about why a particular intervention is (or will be) successful.

In addition to these three purposes of research, the methodological literature also identifies three basic research designs: experiment, survey, and case study (alternatively called field research) (cf. Babbie 1995). Below we will roughly indicate how these research designs can be helpful in systematically studying group model-building interventions.

Case study

Apart from testing hypotheses, it might also be useful to take a closer look at what actually happens in organizations when a group model-building intervention takes place. Carefully conducted case studies and descriptions will be helpful in:

1. describing what actually happens;
2. determining whether particular research and data collection methods actually work;
3. generating relevant and useful hypotheses which can be tested in a more rigorous fashion.

A number of case descriptions have appeared in the literature (e.g., Akkermans 1995; Morecroft and Sterman 1994; Lane 1993; Richardson and Andersen 1995; Senge 1990b).

One problem with these case descriptions is that there is no coherent conceptual model to guide the researcher in determining what to record and what to ignore; hence their incomparability on a number of aspects (see the components of group model building discussed earlier in this article). A second problem is that most of these descriptions do not come up with either an improved method for follow-up research or tangible hypotheses that may add to our understanding of the processes and which may be tested in controlled experiments.

Finally, we have to point out that, although Yin (1989) argues otherwise, case studies are only suitable to generate hypotheses, not to test them rigorously. For the latter, we will have to rely on carefully controlled experiments in order to be able to rule out the effect of confounding factors. However, as we will see in the section on experiments, we will also have to pay a price for this.

Survey research

A survey design can be quite useful in gaining more insight into the effectiveness of our model-building interventions, particularly when individuals are the unit of analysis. Survey refers to a research design in which many data are gathered on a large number of units. This results in a large database, which, in our case, would include data on a number of variables per project, e.g., number of participants, number of meetings, type of modeling procedure employed, etc. — exactly the elements described in the section on components of group model building. The database can be used to statistically test a number of hypotheses. One might, for instance, test the significance of the relationship between participants' satisfaction with the process and project success, controlling for the number of meetings.

Building up such a database and using statistical analysis might also help solve another problem, the requisite variety question — how to deal with getting enough variety to test for an important factor or influence (i.e., if everyone gets top leadership support, how can you test for its importance?). Having a database containing data on various projects with possible differences in top management support may help to test the significance of this factor in project success. It may also shed some light on the factors which do significantly affect the success of a group model-building intervention. One would also have to pay attention to the problem of meta-design — how to combine data from a number of specialized and more focused studies.

One clear prerequisite for this is that the database contains enough data (typically at least one hundred records) to reliably employ statistical techniques. (For an example, see McCartt and Rohrbaugh 1989, 1995; Schuman 1995.) This is one more reason for group model builders to gather similar data on their projects in order to get this database filled rapidly.

Experiments

Naturally, if one wants to test causal hypotheses, the experimental approach is the way to go. This, however, presents a number of problems. The first is related to consulting ethics and client confidentiality. Consultants are supposed to deliver the "best product", making systematic control virtually impossible for real client groups, let alone randomization of participants over research conditions. A second well-known problem is the lack of external validity of laboratory experiments. Results found under controlled conditions in a laboratory do not necessarily have to hold in real-life situations in organizations. Research into the effectiveness of group model-building interventions (and probably interventions in general) seems to be trapped in a dilemma. Results of laboratory research seem to be hampered by external validity problems (on this see, in particular, Eden 1992). On the other hand there is the "burden of proof" problem. If one relies on field research and case studies and system improvement is empirically established, how can one prove that this improvement is the result of the group model-building intervention? (See Cavaleri and Sterman 1997, as well as Huz *et al.* 1997 in this issue).

Each of the above mentioned designs has its strengths and weaknesses. It seems that more than one research design is needed in this research program, in order to increase our understanding and to cancel out the inherent weakness of employing only one design.

Measurement problems

Once we have defined what we intend to accomplish and at what level, and have chosen an appropriate research design, we will have to confront a number of thorny measurement questions. The first is related to operationalization of concepts. Testing hypotheses requires not only that the goals of the intervention be specified in advance, but also that these goals be operationalized into measurable variables. What exactly are mental model alignment, consensus and commitment and how do we measure them? How do we establish improvement of system performance (see Huz *et al.* 1997)?

Closely related to this is the question whether to rely on self-reported measures by participants (e.g., self-reported cognitive changes) or to make an attempt to "objectively" establish cognitive changes. Empirical research indicates that the former is not without dangers (cf. Naftulin *et al.* 1973).

The third problem refers to a Heisenberg-like uncertainty principle. In particular, measuring mental-model improvement is risky. It gives rise to what might be called the mental-model uncertainty principle: efforts to capture an individual's mental model are likely to distort what they seek. The principle applies equally as well to group mental-model measurement, such as measuring progress toward alignment.

The fourth issue concerns measurement level. Should this be done at the reaction level, individual behavior level, group level, or at the level of the organization?

The fifth issue has to do with the time at which to measure the effects of the model-building intervention. Should this be done immediately after the interventions or should retention effects also be taken into account?

Finally, there is the problem of what data gathering techniques best to use: content analysis, interviews, questionnaires, or observation? As was the case with research designs, it seems wise to employ a variety of data-gathering techniques to increase the robustness and validity of the research results.

Practical problems

Finally, we will have to think about how to staff and fund such a research program. We believe that one component of such a program would be the simple first step of having those who practice various versions of this craft attend and participate in each other's facilitated meetings. This would help to work toward the taxonomy of behaviors and to help clarify outcome questions.

The challenge

It is thus reasonably clear what we have to do to advance the science of group model building. It is also clear that there are natural forces — the academic's need for new results rather than replications and the consultant's lack of time or incentives to measure outcomes — that will continue to make it difficult to pursue the necessary research. It remains for us to rise above the inherent challenges and begin the tasks of learning what really helps groups think systematically and strategically.

Notes

1. We intend this neologism to mean "able to be replicated".

References

Ajzen, I. 1991. The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50: 179–211.

Akkermans, H. A. 1995. *Modelling with Managers: Participative Business Modelling for Effective Strategic Decision Making*. Ph.D. dissertation, Eindhoven Technical University, the Netherlands.

Akkermans, H. A. and J. A. M. Vennix. 1997. Clients' Opinions on Group Model-Building: An Exploratory Study. *System Dynamics Review*, 13(1): 3–31.

Andersen, D. F., T. A. Maxwell, G. P. Richardson and T. R. Stewart. 1994. Mental Models and Dynamic Decision Making in a Simulation of Welfare Reform. *Proceedings of the 1994 International System Dynamics Conference*, 11–18. System Dynamics Society, Rockefeller College of Public Affairs and Policy, University at Albany, State University of New York, Albany, NY 12222, U.S.A.

Andersen, D.F. and G. P. Richardson. 1997. Scripts for Group Model Building. *System Dynamics Review* 13(2): 107–129.

Argyris, C. 1990. *Overcoming Organizational Defenses, Facilitating Organizational Learning*. Boston, MA: Allyn and Bacon.

Babbie, E. 1995. *The Practice of Solid Research*. Belmont, MA: Wadsworth Publishing Co.

Cavaleri, S. and J. D. Sterman. 1997. Towards Evaluation of systems thinking Interventions: A Case Study. *System Dynamics Review* 13(2): 171–186.

Conyne, R. K. and L. S. Rapin. 1977. Facilitator- and Self-directed Groups: A Statement-by-Statement Interaction Study. *Small Group Behavior* 8: 341–350.

Eden, C. 1992. A Framework for Thinking about Group Decision Support Systems (GDSS). *Group Decision and Negotiation* 1(3): 199–218.
Huz, S., D. F. Andersen, G. P. Richardson and R. Boothroyd. 1997. A Framework for Evaluating Systems Thinking Interventions: An Experimental Approach to Mental Health System Change. *System Dynamics Review* 13(2): 149–169.
Lane, D. C. 1993. The Road Not Taken: Observing a Process of Issue Selection and Model Conceptualization. *System Dynamics Review* 9(3): 239–264.
Maier, N. R. F. and J. A. Thurber. 1969. Limitations of Procedures for Improving Group Problem-solving. *Psychological Reports* 25: 639–656.
Maxwell, T. 1995. *Decisions: Cognitive Style, Mental Models and Task Performance*. Ph.D. dissertation, Rockefeller College of Public Affairs and Policy, University at Albany, State University of New York, Albany, NY 12222, U.S.A.
Maxwell, T., D. F. Andersen, G. P. Richardson and T. R. Stewart. 1994. Mental Models and Dynamic Decision Making in a Simulation of Welfare Reform. *Proceedings of the 1994 International System Dynamics Conference*, July 1994, Stirling, Scotland, 11–28. Vol: Social and Public Policy, Albany, NY: System Dynamics Society.
McCartt, A. T. and J. W. Rohrbaugh. 1989. Evaluating Group Decision Support Effectiveness: A Performance Study of Decision Conferencing. *Decision Support Systems* 5: 243–253.
McCartt, A. T. and J. W. Rohrbaugh. 1995. Managerial Openness to Change and the Introduction of GDSS: Explaining Initial success and Failure in Decision Conferencing. *Organization Science* 6(5): 569–584.
Morecroft, J. D. W. and J. D. Sterman. 1994. *Modelling for Learning Organizations*. Portland, OR: Productivity Press.
Naftulin, D. H., J. E. Ware and F. A. Donnely. 1973. The Doctor Fox Lecture: A Paradigm of Educational Seduction. *Journal of Medical Education* 48: 630–635.
Phillips, L. D. and M. C. Phillips. 1993. Facilitated Work Groups: Theory and Practice. *Journal of the Operational Research Society* 44(6): 533–549.
Richardson, G. P. and D. F. Andersen. 1995. Teamwork in Group Model Building. *System Dynamics Review* 11(2): 113–137.
Richardson, G. P., D. F. Andersen, T. A. Maxwell and T. R. Stewart. 1994. Foundations of Mental Model Research. *Proceedings of the 1994 International System Dynamics Conference*, Stirling, Scotland. Albany, NY: System Dynamics Society.
Rohrbaugh, J. W. 1992. Collective Challenges and Collective Accomplishments. In *Computer Augmented Team Work: A Guided Tour*, ed. R. P. Bostron, R. T. Watson and S. T. Kinney, 299–324. New York: Van Nostrand Reinhold.
Rosenhead, J. ed. 1989. *Rational Analysis for a Problematic World: Problem Structuring Methods for Complexity, Uncertainty and Conflict*. Chichester: John Wiley & Sons.
Schein, E. H. 1987. *Process Consultation*, Vol. II. Reading, MA: Addison-Wesley.
Schuman, S. P. 1995. *Valuing and Using Data in Group Decision Making: An Examination of Decision Conferences and the Effect of Decision Makers' Perceptions of Data and Empirical Process on Outcomes*. Ph.D. dissertation, Rockefeller College of Public Affairs and Policy, University at Albany, State University of New York, Albany, NY 12222, U.S.A.
Senge, P. M. 1990a. *The Fifth Discipline: The Art and Practice of the Learning Organization*. New York: Doubleday.
———. 1990b. Catalysing Systems Thinking Within Organizations. In *Advances in Organization Development*, Vol. 1, ed. F. Massarik, 197–246. Norwood, NJ: Ablex.

Vennix, J. A. M. 1990. *Mental Models and Computer Models. Design and Evaluation of a Computer-Based Learning Environment for Policy Making*. Ph.D. Dissertation, University of Nijmegen, the Netherlands.

———. 1995. Building Consensus in Strategic Decision Making: System Dynamics as a Group Support System. *Group Decision and Negotiation* 4(4): 335–355.

———. 1996. *Group Model-Building: Facilitating Team Learning Using System Dynamics*. Chichester: John Wiley & Sons.

Vennix, J. A. M., H. A. Akkermans, and E. A. J. A. Rouwette. 1996. Group Model Building to Facilitate Organizational Change: An Exploratory Study. *System Dynamics Review* 12(1): 39–58.

Vennix, J. A. M., Scheper, W., and Willems, R. 1993. Group model-building: what does the client think of it? *The role of strategic modelling in international competitiveness, Proceedings of the 1993 International System Dynamics Conference*, Cancun, Mexico, ed. E. Zepeda and J. Machuca, 534–543.

Verburgh, L. D. 1994. *Participative Policy Modelling: Applied to the Health Care Insurance Industry* Ph.D. dissertation, Nijmegen University, The Netherlands.

Winch, G. W. 1993. Consensus Building in the Planning Process: Benefits From a "Hard" Modeling Approach. *System Dynamics Review* 9(3): 287–300.

Yin, R. K. 1989. *Case Study Research: Design and Methods*. London: Sage.