Eindhoven University of Technology

Clients' opinions on group model-building : an exploratory study

Document status and date:
Published: 01/01/1997

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Clients' opinions on group model-building: an exploratory study

Henk A. Akkermans[a] and Jac A. M. Vennix[b]*

Henk Akkermans currently works as a senior consultant with Origin. He undertook his doctoral research as an assistant professor at Eindhoven University of Technology. His current research interests include system dynamics modelling in managerial consultation and empirical evaluation of its effectiveness.

Jac Vennix is associate professor of research methodology at Nijmegen University. 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 (John Wiley).

Abstract

Group model-building is increasingly used to support strategic decision-making in organizations. However, little is known about its effectiveness, apart from anecdotal evidence and statements by consultants that it works. This article reports on an assessment study of six group model-building projects. Since few tested theories are available, case studies and a qualitative research approach were used to shed more light on the effectiveness of group model-building projects in real organizations working on real strategic problems. The results show that a number of hypotheses known from experience or textbook theory were 'confirmed', while others were only partially confirmed or should be rejected on the basis of the six cases investigated. © 1997 by John Wiley & Sons, Ltd. Syst. Dyn. Rev. **13**, 3–31, 1997

(No. of Figures: 22   No. of Tables: 3   No. of Refs: 24)

Building models with a group of stakeholders has become an established approach to support strategic decision-making. Involving these stakeholders helps to generate relevant information regarding the issue and at the same time creates ownership of and consensus on the resulting group recommendations for dealing with the issue. In this way, group model-building creates managerial commitment to implement these recommendations.

In the last decade, the system dynamics community has made considerable progress in developing tools and techniques to support a group model-building process. Graphical facilitation techniques, such as causal loop diagrams, stocks-and-flows diagrams and graphical functions are used in combination with guidelines for structuring and facilitating group sessions, group knowledge-elicitation techniques and appropriate consulting roles (e.g. Morecroft and Sterman 1994, Richardson and Andersen 1995; Vennix 1996). Success stories abound on the application of these refined approaches to support corporate decision-making; the subject is also widely discussed in scientific journals.

Nevertheless, we have rarely asked our clients if they are as enthusiastic about group model-building as we, the consultants, are. That is to say, we have not often bothered to do so in a systematic and rigorous manner. There are, of course, many anecdotes of managerial appreciation of group model-building approaches using system dynamics. Unfortunately, those are insufficient for at least two reasons. First, consultants remain suspect sources of information on clients' perceptions

[a] Origin/IPS, P.O. Box 6374, 5600 HJ Eindhoven, The Netherlands. E-mail: Henk.Akkermans@nlehvips.origin.nl
[b] Dept. of Policy Sciences, Nijmegen University, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands. E-mail: J.Vennix@maw.kun.nl
* Corresponding Author.

since they have a personal interest in emphasizing good news and downplaying bad news. Second, superficial answers to questions like "how did you like it?" are bound to miss much of the richness of information that systematic interviewing and analysis can deliver.

In this paper an exploratory study is described in which clients' opinions were asked in extensive, structured, post-project interviews. These interviews were transcribed and the texts were analysed systematically. The findings themselves seem, in general, to confirm some of the assumptions commonly made in the field, but simultaneously seem to contradict other parts of textbook theory as well. Moreover, several interesting, but less frequently mentioned, assumptions were found to hold up in this study.

In this study the researcher (i.e. the first author) also acted as the consultant in all of the cases. Hence, in this study a number of precautions were taken to increase the study's validity and reliability (cf. Yin 1989). First, the researcher was assisted by another person and differences in analyses between these two persons were discussed throughout the study in order to increase reliability. Second, member checks were carried out to ensure that conclusions drawn from the material were at least recognized by those interviewed. Third, peer reviews were conducted by the three supervisiors of this doctoral reseach project (including the second author of this paper) as a check upon the analyses and conclusions drawn from the data.

Research methodology

Multiple case study design

Over a period of two and a half years, six commercial model-building projects were conducted by the first author; the second author collaborated in the second project. Each of these case studies has been described in separate publications, mainly for a system dynamics audience, so they will not be discussed at length here. These case studies varied widely in scope, content matter, client type and many other characteristics, but in all six case studies the same specific version of group model-building was used, which is known as Participative Business Modelling or PBM within Origin (Akkermans 1995a). As in most group-model building approaches, the PBM method blends system dynamics modelling with a non-expert mode of process consultation (Schein 1969; Richardson and Andersen 1995; Vennix 1996) to ensure maximum client participation and ownership of results. The projects were on the following topics:

1. cycle-time reduction in newspaper distribution (Akkermans 1994);

Fig. 1. Research model

STRATEGIC DECISION MAKING   STRATEGIC DECISION IMPLEMENTATION
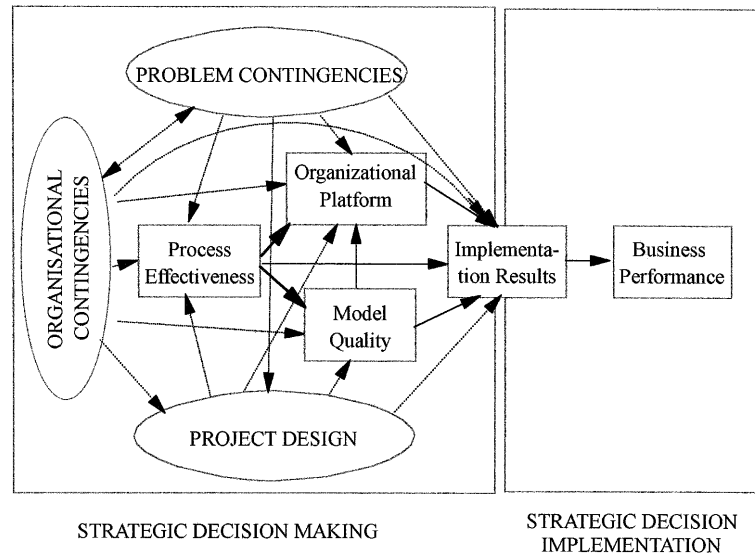
2. creating a more collaborative attitude between independent business unit managers in an IT company (Vennix *et al.* 1996);
3. a logistics strategy for a pharmaceutical company (Akkermans 1995b);
4. an implementation plan for a corporate strategy in the service industry (Akkermans and Bosker 1994);
5. rationalization of branch office networks in banking (Akkermans 1995c);
6. supply chain management strategy in electronics (Akkermans 1995a).

Research model

Figure 1 shows the overall theoretical framework, or research model, for the evaluation study. It illustrates that implementation results will depend both on the quality of the model that is being used and on the level of organizational support for that model and its recommendations. It also shows that both will depend critically on the quality of the modelling/decision-making process ("process effectiveness").

Figure 1 also shows that there are a number of contingencies at play. "Data availability" and "problem tangibility" are examples of problem-related contingencies, while "political sensitivity" and "problem ownership" are examples of organizational contingencies. The third category of contingencies ("project design") is one that is within the control of the modeller; this is the specific design of the modelling method that is used. Aspects of that design include, amongst others, the

Table 1. Aspects of strategic decision-making effectiveness

Process effectiveness	Organizational platform	Model quality	Implementation results
Focus	Awareness	Completeness	Implementation of decision
Speed	Consensus	Thoroughness	Business performance
Involvement	Commitment	Theory-basedness	Insight
Communication	Ownership	Usability	Organizational learning
Willingness to cooperate	Confidence		

usage of quantified simulation, what graphical modelling techniques are employed and the expertise of the process facilitators.

In the model it is assumed that business performance is affected by the implementation of strategic decisions, which are in turn the result of a particular group model-building intervention. In practice, it is difficult, if not infeasible, to actually "prove" that business performance has changed or improved as a result of a group model-building intervention (see, for instance, Cavaleri and Sterman 1997). In addition, particularly for the later cases, the time frame between intervention and assessment was too short to actually observe any changes. Hence, this study is primarily limited to the other factors in the research model. However, in order to get some insight into the potential effects of the intervention interviewees in the first two cases of the study were asked for changes in business performance as a result of the intervention.

Figure 1 contains a top-level overview of the research model. If the model were given in more detail, each of the overall factors identified, would be seen to contain some four to six so-called "indicators", or aspects of this overall factor. Table 1 provides listings of these indicators for aspects of strategic decision making effectiveness.

Most of the labels in Table 1 will be self-evident, but some crucial ones require additional explanation. For instance, "involvement", which meant in general "the degree to which the organizational stakeholders participated in the decision-making process", has been divided into project participation and workshop participation. Also "communication", which was the catch-all for "the quality of the conversational process between the various participants", was actually subdivided into five different aspects, notably:

Exchange of ideas/viewpoints: the degree to which participants felt they were able to present their ideas;
Openness: the degree to which discussions were felt to be open, without hidden agendas;
Common language: the degree to which a shared language was used and participants understood one another;

Table 2. Contingencies influencing stategic decision-making effectiveness

Problem contingencies	Organizational contingencies	Project design elements	
Problem scope	Top management support	Pre-interviews	Data analysis
Problem tangibility	Hierarchical diversity	Hexagon brainstorming	Simulation
Data availability	Problem ownership	Causal loop diagrams	Final report
Problem urgency	Group size	Stocks-and-flows diagrams	Central/graphical presentation
Political sensitivity	Working relations	Graphical functions	Facilitator skills
		Workbooks	Abstraction level
		Propositions	Project size

Fig. 2. Evaluation process
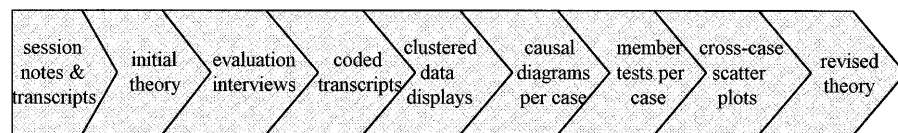
(Lack of) verbal dominance: the degree to which participants were able to contribute equally to the discussions;

Freedom: the degree to which participants felt free to introduce their ideas and opinions.

In addition, the objects of the nouns can be semantically ambigious. For instance, "willingness to cooperate" refers to participants' attitude towards the modelling *process*, whereas "ownership" points to participants' feelings toward the *output* of that process, i.e. the model and the recommendations that arise from it.

Table 2 presents the indicators for the various contingencies distinguished in Figure 1. For a complete description of all the indicators, the reader is referred to Akkermans (1995a).

Evaluation procedure

The evaluation procedure for this study was both exploratory and extensive. Exploratory, because very little similar research had been conducted in the past, which also led to a large number of variables that had to be taken into account. Extensive, because of its broad focus and the huge amount of text material that had to be processed. The evaluation process lasted more than two years and has taken up at least one and a half man-years, in which approximately one hundred hours of tape recordings were analysed. Figure 2 shows the main steps taken in this evaluation process, or rather, the outputs of each step.

Fig. 3. Example of an
observation from
Case 5

1/2/1994 [after a learning-wheel workshop with one of the "test banks"]
Bank manager at the end of workship, after question on organisational platform for model:
Original advice from internal consultant was mainly economical; it did not contain these
kind of considerations [see session tape for rest].
Non-verbal reaction of internal consultant: Looks surprised at manager, bends over to look
at thick portfolio with research data in front of him and lifts it partially as if to say "And
what about this here, then?"

Session notes and tape recordings were the direct output of the conduct of the cases
themselves. The researchers on site, the first author and a research assistant,
recorded their observations and memos during the process and most of the group
model-building sessions were taped; these recordings were transcribed afterwards.
Figure 3 contains an example of such a memo (see also Akkermans 1995c).

An initial theory of what determined strategic decision-making effectiveness in
these modelling projects was constructed by the researchers. This theory was based
upon a survey of the relevant existing literature on the one hand, and, on the other,
upon their experiences and discussions during the cases. This theory was
formulated as a causal diagram. Its final top-level representation is contained in
Figure 1, but the original version was far less refined or detailed.

Evaluation interviews were conducted, guided by this theory. On the basis of the
concepts and hypotheses distinguished by the researchers, interview questions
were formulated. Some 80% of the participants in the six cases were interviewed,
and their answers were also taped and transcribed. As the reader may understand,
very much data was gathered in this way (a total of some 70 hours of spoken word).
One central problem in the qualitative research was to reduce these data in such a
way that clear and reliable conclusions can be drawn. This includes the process of
coding the transcripts.

Coded transcripts were thus derived from the interview data. This was done by
cutting up the interview transcript into smaller so-called "scenes". These scenes
were content-analysed and one or more codes (or labels) were attached to a scene.
These labels in turn corresponded to the aspects and indicators presented in Tables
1 and 2 (e.g. communication, involvement, ownership). The result is a so-called
"coded transcript", for one particular interviewee.

Clustered data displays further reduced the data. After interview transcripts had
been coded the results were clustered into an overview of relevant statements by all
the respondents with regard to each of the aspects or indicators distinguished in
Tables 1 and 2. This produced so-called "clustered data displays". These data

displays could again be summarized into overall statements, both verbally and quantitatively (with pluses and minuses), and then grouped into again a higher level table or data display (Miles and Huberman 1984).

Figure 4 shows a display of Level 1. Here the case data on "willingness to cooperate" are shown for one particular case, i.e. Case 4. This display was constructed by selecting from the coded transcripts of each interview the relevant scenes with regard to "willingness to cooperate". (R1 indicates "respondent 1", R2 "respondent 2", etc.)

A crucial step in the analysis, reflected in the bottom two rows of Figure 4, is the *assignment of values* (+, -, ++, --) to verbal summaries. This step is crucial, because these values will be subsumed to the higher-level displays and will be used in the cross-case analysis. The reason for this is that such values "are less ambiguous and may be processed with more economy" (Miles and Huberman 1984, 54). However, the verbal summaries are also retained in higher-level displays, because "although words may be more unwieldy than numbers, they also enable 'thick description', that is, they render more meaning than numbers alone, and should be retained throughout data analysis." (Miles and Hukerman 1984, pp. 54–55).

Another thing worth noticing in this display is the fact that the two researchers who worked on these displays, the author and his research assistant, cross-checked each others' assessments. The research assistant would make up the initial display, the author would review this and suggest changes he felt were appropriate (indicated by strikethrough and italic fonts in the display). These modifications were discussed and, at times, yielded further changes. The reasons for all such alterations were documented on the displays.

As stated, Figure 4 shows a display Level 1 for "willingness to cooperate", which is one aspect of the variable "process effectiveness" presented in Table 1. In the next data-reduction step, results for "willingness to cooperate" are combined with other indicators for process effectiveness (i.e., focus, speed, involvement and communication). As can be seen in Figure 5, this is done by taking the "data reduction" from the bottom rows of Figure 4 and transferring these to the cells in the column "willingness to cooperate" in Figure 5. A similar procedure is carried out for each of the other four indicators of process effectiveness that make up Figure 5.

As can be seen in the two bottom rows, the results of the interviews were then summarized per indicator. Figure 5 thus shows the overall results for one case with regard to one of the seven important research concepts (i.e. process effectiveness) as distinguished in Figure 1. Hence, for each of the six projects, seven of such displays were constructed: one for each measure of strategic decision-making effectiveness (i.e., "process effectiveness", "model quality", "organizational platform" and

Fig. 4. Display level 1
for "willingness to
cooperate" in Case 4

R1	R2	R3	R4	Memoranda/ Documents
SC7: We have only done a part of the process and perhaps that is because some people did not support the process and don't believe in it, that I don't know. SC20: We stopped halfway, partially also because one or two people didn't believe in the approach. SC22: I certainly do believe that everyone tried to contribute positively, but during the process I got the feeling "This is not going right, we are not achieving our goal", then after a certain time you quit.	SC4: (on grouping hexagons): I shifted a few and the dominant players were shifting them back, so I say forget it. SC8: I think there were, let's say, at least one absolutely for sure and one or two, I would say, dubious who did not agree with the objectives. SC11: I got the impression that it was kind of commissioned by one guy without really anyone around him wanting it particularly... When one say says "I think it is a very good idea and we"re gonna do it" and everyone else around would prefer to be somewhere else, it is not actually destined to give the right result.... A lot of people thought it was a waste of time from the beginning. SC14: I think the fundamental problem was the way the thing was initially started by one man, saying this is a good idea without really considering how to do it, and it wasn't a management team kind of decision. SC16: I know a number of people pushed for it to be cancelled. SC18: E. [the project sponsor — HA] wanted to carry on at the end of the day; everyone else really said stop it. SC22: It kind of never worked because people were getting frustrated by it. SC31: I think nobody there got committed to it at all SC32: If one is gonna do this, one of the first things is to make sure you have the right commitment from the right people. SC33: ...You just get a memo saying you will be there SC34: I know a number of people was kind of ...well, how do we get out of this thing? Some even categorically cancelled it out of their diary.. Lots of last minute bookings.... SC37: You really need to have at least a commitment generally amongst the people SC39: The commitment [to the process] has to be right	SC5: (People did not know the method...) Now there was a disbelief in the technique and therefore people who said: "Why am I sitting here?" and got irritated, and so on. SC10: I somewhat had the impression that this was a sale made by B. [the author's consulting company —HA] (Wasn't it supported by the people?) No.. there had been internal tension regarding that session.. There has to be a large willingness to do this. SC11: I think that the original setting already indicated the impossibility of success. I did not believe in it myself, I have tried one week before to have it cancelled.	SC13: I think that everyone who was present there had a very positive attitude to give that input that could be expected from him at that time [insinuating that this was not the case by me resulting in irritation, interviewer]	HA: There were three kinds of unwillingness: 1. People did not want the new strategy, and did not want to talk openly about their objections (R2, R4) 2. People did not want this participatory process, this method (R6 - not interviewed -according to R3) 3. People did not want to discuss this issue with their subordinates in this way (R3)
3 scenes SUMMARY: 2 people did not support the process, and because things were going so bad R1 also quit himself. RATING: +/- Most wanted to contribute positively, but not everybody	13 scenes. SUMMARY: Many people did not want this project, they were forced to participate RATING: - - Referred to very often	3 scenes. SUMMARY: Project was forced, people did not want it and did not grasp it. RATING: minus Could also be double minus RATING: - - Others did not want it, even R3 did not want it himself either HA	1 scene SUMMARY: attitude was positive Gives vague/diplomatic reply, no reply really. HA RATING: plus RATING: NA (HA)	

21 references.
OVERALL: - - Several people did not want this project at all.

Fig. 5. Display level 2 for various aspects of process effectiveness in Case 4

PROCESS EFFECTIVENESS	Focus	Speed	Involvement	Communication	Willingness to cooperate
Mindmaps pre-interviews	NA	NA	NA	no open communication, verbal dominance, old gang dominant coalition (2 ref)	NA
Sessions	--: there is too little steering, there's no direction and it's not about the project goal (13 ref)	NA	+: good that everybody is sitting round the table (1 ref)	-: a lot of discussion, partly because of a lack of common language and a dominance of two people (13 ref)	NA
R1 post-interview	--: bad focus due to lack of steering facilitators (11 ref)	-: no speed because of bad focus (4 ref)	+: all important people were there and had a part in the project (4 ref)	+: mutual discussion was good, though sometimes some dominance and lack of common language. Enough room to participate and open atmosphere (14 ref)	±: most people tried to cooperate although some people weren't behind the process. In the end it went wrong and I was not behind it any more (3 ref)
R2 post-interview	--: wrong subject and bad steering (20 ref)	-: too quick, no time to come to an agreement (2 ref)	--: wrong balance in national and international people (8 ref)	--: hardly any discussion, no common language, dominance of some people and lack of openness (21 ref)	--: most people didn't want the project, they were told to cooperate (14 ref)
R3 post-interview	--: too many different goals and too much talking about subjects already known by most participants (9 ref)	-: too quick for some and too much repetition for others (7 ref)	-: wrong setting because of background of people and group size (7 ref)	-: no common language and little openness because of hierarchy and internal goals (8 ref)	-: people were forced to cooperate, they didn't want it and didn't understand it (3 ref)
R4 post-interview	--: different goals and no steering (7 ref)	-: no speed because the difference in goals made it difficult to come to agreement (2 ref)	+: right people were there, other factors meant that they couldn't contribute totally (5 ref)	±: no common language, atmosphere was open (5 ref)	NA
Memoranda	-: people talk about method not about project goals, model coach takes over facilitating (2 ref)		right people were there, but didn't participate equally	±: gradually more discussions during sessions, dominant people don't overrule discussions (2 ref)	--: who *did* want the project, now I am beginning to doubts about the sponsor. HA
Documents	NA	NA	NA	+: philosophy of company is openness and respect, and no hierarchical dominance (1 ref)	NA
# ref	62	15	26	66	20
OVERALL	--: different goals, wrong subject, insufficient steering	-: for some too much speed, for others too little, both because of bad focus	+: all important/relevant people, background of the participants caused problems	-: no common language, some dominance and lack of openness; free thinking and mutual discussion was possible	--: some people didn't want the project
OVERALL	RATING: minus RATING: -- (double minus) *If this was not a bad process then when are processes bad? It was a nightmare! HA* Involvement was okay but bad focus and communication caused lack of speed and not everybody was willing to cooperate				191 ref

"implementation results"), two for organizational and problem contingencies and the seventh for the various aspects of project design. These then were grouped together once more, according to the same methodology, into a single data matrix that showed the overall assessment of the various elements of strategic decision-making effectiveness for this particular case (see Akkermans and Bosker 1994 and Akkermans 1995a). From there it is only a small step to a fourth level of data reduction, i.e. an overall assessment of the effectiveness of the six cases together, as shown in Table 3 in the section on the main findings of the study.

Causal diagrams per case were constructed simultaneously with the data matrices as a separate stream of analysis. The reason for this is that data matrices as described above may give a good *description* of the "inputs" to the strategic decision-making process and its "outputs", but they do not provide an *explanation* of why these particular results were obtained, e.g., explanations of questions like: "What caused commitment to be high in this case?" or "Why was completeness only moderate?". For that purpose causal diagrams are employed.

In qualitative research, a causal network is described as "a visual rendering of the most important independent and dependent variables in a field-study and of the relationships between them" (Miles and Huberman 1984, 132). It is considered as "the analyst's most ambitious attempt at an integrated understanding of a site." (Miles and Huberman 1984, 142).

In developing a causal network, one can choose between two kinds of generic analytic approaches. In the *deductive* approach, the researcher starts with a preliminary causal network, based upon existing theory, and looks for data that will confirm or, even better, refute this network. In the *inductive* approach, the researcher looks for mention of causal links in the case data and based on these references constructs a causal network "bottom up", leaving the confrontation of this causal network with existing theory until afterwards (see Miles and Huberman 1984). In this study, a mixture of both approaches was used. The variables from the research model and several relations between them were already available at the time of the causal case analysis. In that sense, the approach was deductive: research started off from a (partial) preliminary causal network. On the other hand, initial versions of these causal networks were constructed "from the ground up", from actual clues in the case material. To that degree, the approach was inductive.

This bottom-up construction was performed by collecting from the case data direct references to causal reasoning. The researchers scanned all the cells in Display Level 1 for mentions of causal relations involving elements of the research model. Sometimes these remarks are very direct, sometimes they are little more than clues. (For a more detailed description of Case 2, see Vennix *et al.* 1996 or Akkermans 1995a.). These relations were then plotted in a causal network. Figure
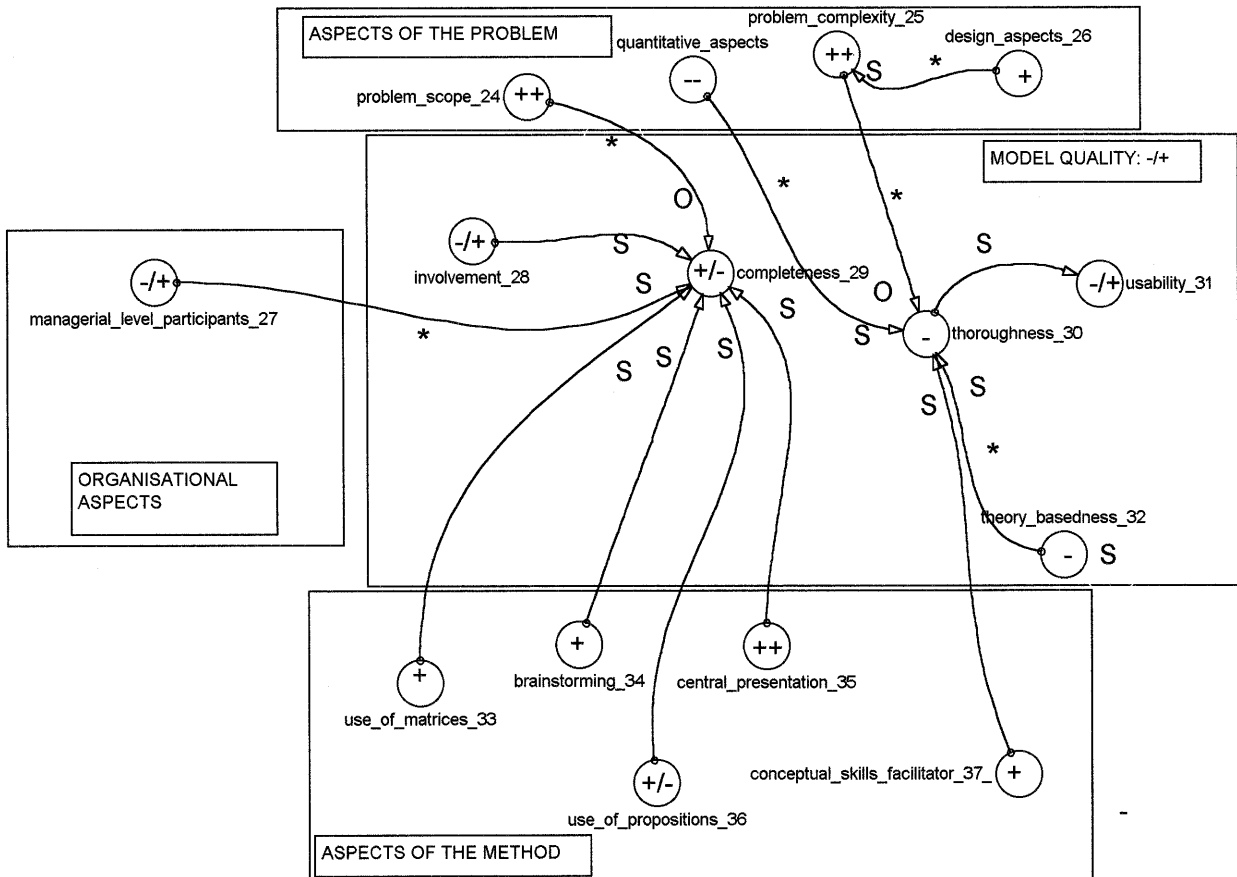
6 shows the results of this plot for Case 2. Every relation that is not marked with an asterisk (*) was inferred directly from case evidence (we will go into these marked relations a little later).

The information density of such a causal network is very high: a considerable number of variables are displayed, together with their scores. They are grouped in boxes, each of which represents an overall concept of the research model. The score for the overall concept "model quality" is also shown. Finally, the directions of the causal relations are also indicated with "S" for "same" and "O" for "opposite".

In qualitative research a causal diagram also has associated text describing the meaning of the connection between the factors. Such a textual description for the causal network in Figure 6 is shown in Figure 7. Each variable, each relation in the network, has a number; each relation is described separately and refers to that specific number. One can study the diagram, one can read the text, or one can do both.

Fig. 7. Verbal description of the "model quality" network in Figure 6

The quality of the model was inadequate. Formulated in general terms, the model-based <u>analysis</u> of the problem was OK, but we didn't succeed in finding good <u>solutions</u> for the problem in the model-building process.

"Model Quality" is built up from a number of aspects: "completeness", "throughness" (of analysis), "theory-basedness" (the degree to which existing theory is used) and (practical) "usability". Regarding these aspects, the following can be said:

|	
29	The *completeness* of the analysis was acceptable
*24→29	This was especially so considering the wide scope of the problem. A number of the techniques used contributed to this completeness. Mentioned specifically are:
34→29	brainstorming with hexagons:
33→29	the usage of matrices and, more generally:
35→29	the central presentation, that enabled it to capture and keep track of most of the discussion.
36→29	Respondents are less positive about the usage of propositions, one feels that this obscures aspects of previous discussions.
28→29	Broad involvement in the project, in itself, improves completeness, but as indicated this involvement left something to be desired sometimes.
30	The *thoroughness* of the analysis was insufficient. (By thoroughness we mean the degree to which all the required analyses have been conducted.)
37→30	The conceptual modelling skills of the facilitators did contribute,
*25→30	but the problem itself was so complex
*21→30	and intangible in nature that no good solutions came from the analysis.
31	This made the practical *usability* of the model limited. Why were there no good solutions found? The interviews give no clear explanations for this. Possible explanations are:
*32→30	Insufficient use was made of normative theories on how to design new organisational structures (if these exist)
*26→25	This problem was so complex because it is a <u>design</u> issue: designing a new organisational structure is a far more difficult assingment than analysing the current structure. This is because there are so many options.
*27→29	Finally, this might not have been the right management level to which to address this question. These managers were pilots, not aircraft builders, left alone aircraft designers; only top management of the company can make such choices.

Member checks per case is a procedure to improve internal validity of the case (see Yin 1989). In a member check, "one presents facts and interpretations to participants to establish whether the reconstruction of reality as the researcher sees it is also recognisable to them" (Hutjes and van Buuren 1993, p. 212). In this research project, both member checks and peer reviews were conducted for the case material. In the member checks, each respondent received a copy of the causal case analysis. Each was asked to look at the interpretations described, paying special attention to the marked (*) relations, because these contained the researcher's own interpretations. This suggestion may have helped to reduce the amount of data respondents had to check, but many respondents indicated that they were more or less overwhelmed by the amount of data that they were confronted with and gave

Fig. 8. Cross-case scatter plot for simulation → thoroughness
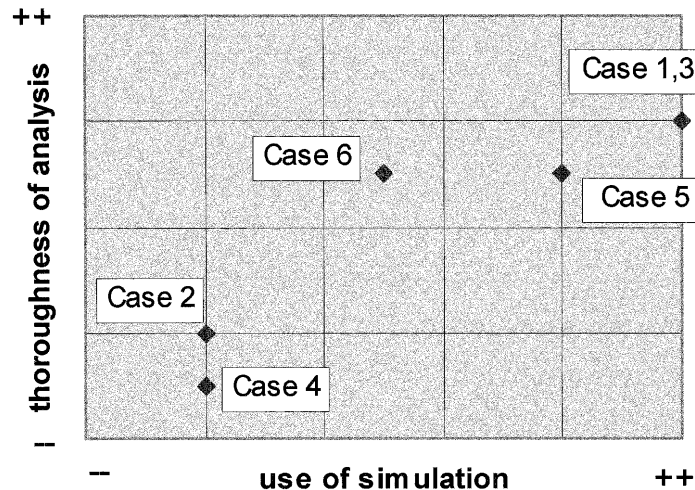
fairly general answers. Sometimes the replies were more specific, mainly when the interpretations presented in the analysis were regarded as controversial. In general, therefore, the researchers saw the member check as little more than a fairly crude "sanity check" on case analysis: if the researcher was really talking nonsense, the respondents would tell him so, whereas if the analysis looked plausible, they would tend to agree with it.

Cross-case scatter plots were a key element in the cross-case analysis process that started next. Cross-case scatter plots are "figures that display data from all sites on two or more dimensions of interest that are related to one another. Data from the sites are carefully scaled, and laid out in the space formed by the respective axes." (Miles and Huberman, 1984, 181). Figure 8 shows such a "scatter plot" of the data on the relation: "the use of simulation improves thoroughness", one of the textbook theories in system dynamics, if not in all quantitative modelling disciplines. In this research, scatter plots were constructed by translating the "--" to "++" values for each of the model variables into a five-point scale.

Figure 8 suggests that, at least in the six cases investigated, the use of simulation seems to relate positively with perceived thoroughness of the analysis. (Please note though that although the method superficially resembles bivariate statistical analysis, no statistical generalization is implied. We are merely looking at how well the case data fit our research model.)

For this particular relation, we can probably say that the relation was confirmed by the data from the six cases. We can do this with some confidence because the ranges of values are fairly broad for both variables. However, this does not mean that projects without simulation cannot result in thorough analyses; this scatter

plot only visualizes some case evidence for the hypothesis — which is well grounded in the literature — that simulation modelling is instrumental in achieving a thorough analysis of a problem.

A final remark is on the direction of causality. In this relation, it seems obvious that the use of simulation leads to thoroughness and not the other way round. With some relations, however, this is less straightforward. Does focus improve communication, or does good communication lead to focus? In this respect, the following points are relevant:

Because everything affects everything else, we have to make a selection of what relations we feel are most relevant. This selection is made on the basis of a review of the literature and/or information from the evaluation interviews.
For several relations, causality is probably bi-directional: good focus leads to better communication and in a process with good communication it may be easier to achieve good focus. However, this was only assumed in those instances where both causal directions were mentioned in the literature or were made explicit in the case data, such as with "communication leading to insight" and "insight leading to communication".

Revised theory. The last step in the evaluation procedure was to revisit the causal relations that were initially defined on the basis of a sampling of the literature and the experiences of the researchers as consultants and confront these relations with the cross-case evidence that had been accumulated for them. This analysis focussed on two specific categories of causal relations in particular:

Textbook theory: relations that were found to be often cited in the literature sample that was taken (e.g. Morecroft and Sterman 1994 for the field of system dynamics, but see Akkermans (1995a) for a complete list).
Exploratory theory: causal relations that were not part of the textbook theory but that were nevertheless found to be supported strongly by the cross-case evidence.

Results

In this section we will discuss the main findings of this study, and present them in three parts:

1. an evaluation of overall effectiveness of the group model-building process;
2. an evaluation of "textbook theory" relations on the basis of the cross-case evidence;

Table 3. Strategic
decision-making
effectiveness by case

Overall concept	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Process effectiveness	+	+	−/+	−	+	+/−
Model quality	+	−/+	+	−	+	+/−
Organizational platform	++	+	++/+	−	++/+	++/−
Implementation results	++	+/−	+	−/+	+	−/+

3. an assessment of the exploratory relations that were strongly confirmed by the cross-case evidence.

Evaluation of overall effectiveness of the group model-building process

In general, the results of these six projects were fairly positive. Table 3 shows how the six cases scored on the four key elements of strategic decision-making effectiveness that were distinguished. It shows that only Case 4 was a clear failure, albeit a failure that generated important insights for the consultants and improvements to the PBM method. It also shows that, in general, these projects scored very high on organizational platform, the overall concept in which concepts like consensus, ownership and commitment are grouped.

Evaluation of "textbook theory" relations

Of the 66 relations distinguished in the original research model, only a limited number (11) were found to be frequently discussed in the literature. These relations are depicted in Figure 9. This figure also indicates which relations were confirmed or partially confirmed by cross-case analysis. Each of these relations has been set against the case data results. In particular, a comparison has been made between the scores for each of the variables mentioned in each of the cases. If a relation between, for instance, "involvement" and "ownership" was present in these cases, then one would expect that whenever involvement was high, ownership would also be high, and vice versa.

   Figure 9 shows only three "smiling face" icons, which means that only three of the 11 "textbook theories" were clearly confirmed by the data from our six studies. Three additional relations were moderately confirmed, as is indicated by the balanced seesaws. The remaining five were all refuted (that is, we found cases where the relation did not hold up). We will discuss each category in some more detail.

CONFIRMED TEXTBOOK THEORY RELATIONS.   In this category we get what we would expect according to the standard theory, so there are not too many surprises.
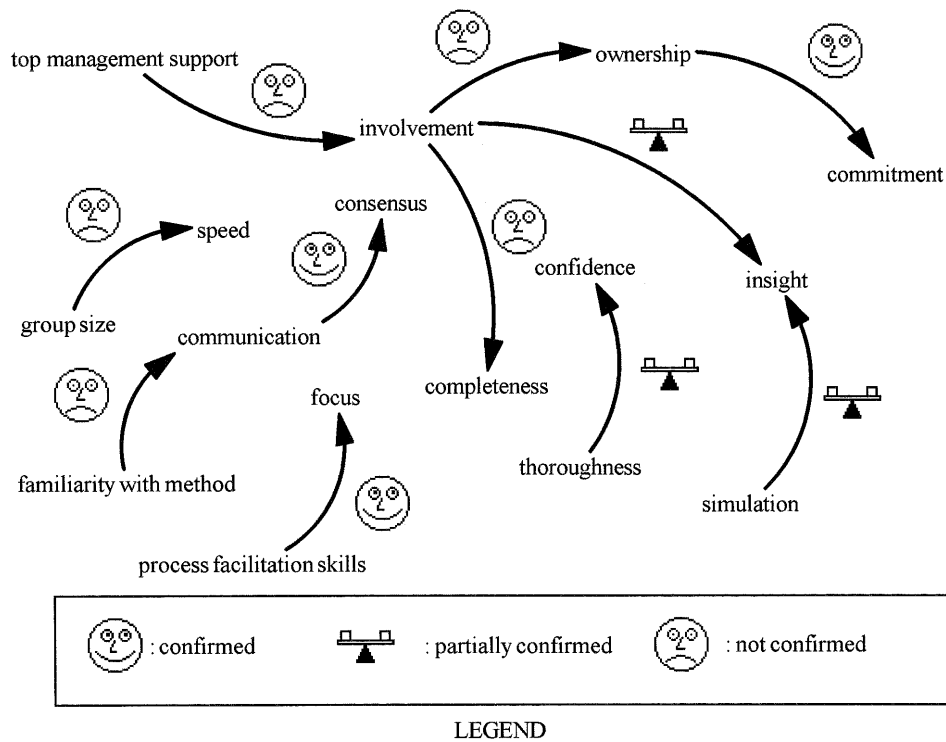
Fig. 9. Textbook
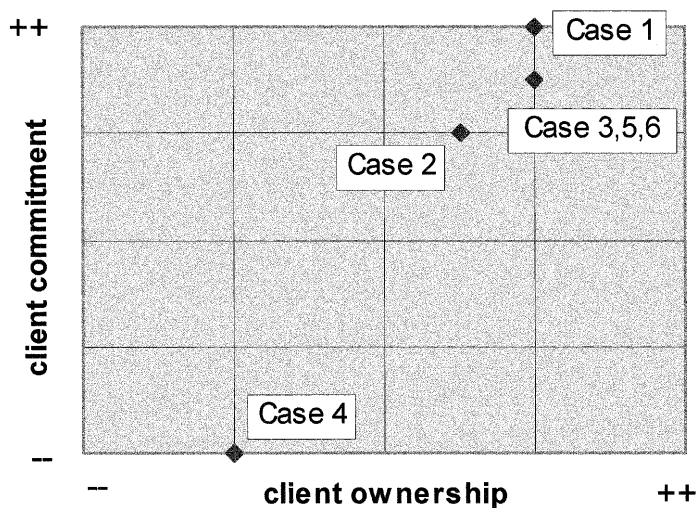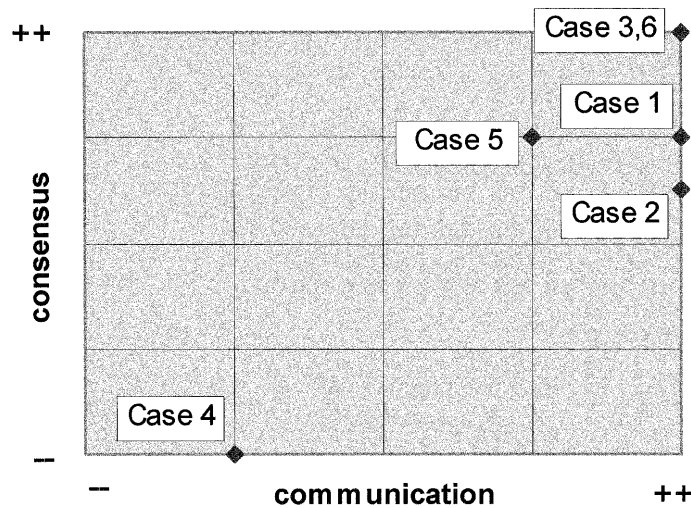theory

LEGEND

Fig. 10. Ownership →
commitment

Fig. 11.
Communication →
consensus

Ownership → commitment. Textbook theory declares that it is important to get maximal client ownership of a model because client ownership will lead to commitment from the client to implement the model's recommendations. If clients do not feel that a model is theirs, they are unlikely to act upon it. This is strongly confirmed by the case data, as Figure 10 shows. Incidentally, this also illustrates the academic relevance of unsuccessful case studies; for had it not been for the disaster of Case 4, only a very narrow range of values would have been observed.

Communication → consensus. The relation between communication and consensus is often discussed in the literature, but is not universally regarded as a strictly positive one. In politically sensitive situations, it has been suggested that intensive communication may have a negative impact on consensus (McGrath 1984; Mintzberg 1994). For the data of our study we have to take into account that we measured the quality of communication rather than its intensity. The better the quality, the easier it is to reach consensus (Figure 11). This idea would seem to be confirmed by the case data, for in Case 4, the only highly political sensitive project, consensus did indeed reach an all-time low. In all other cases good communication coincided with fair to high levels of consensus.

Process facilitation skills → focus. For neither the researcher nor the practitioner will it come as a surprise that this relation was upheld: whenever process facilitation skills were perceived as good, focus was good, and vice versa. However, as stated, the direction of causality may be less clear.

PARTIALLY CONFIRMED TEXTBOOK THEORY RELATIONS. Relations were described as partially confirmed if the relation was upheld in four or five of the six cases.

Fig. 12. Thoroughness → confidence
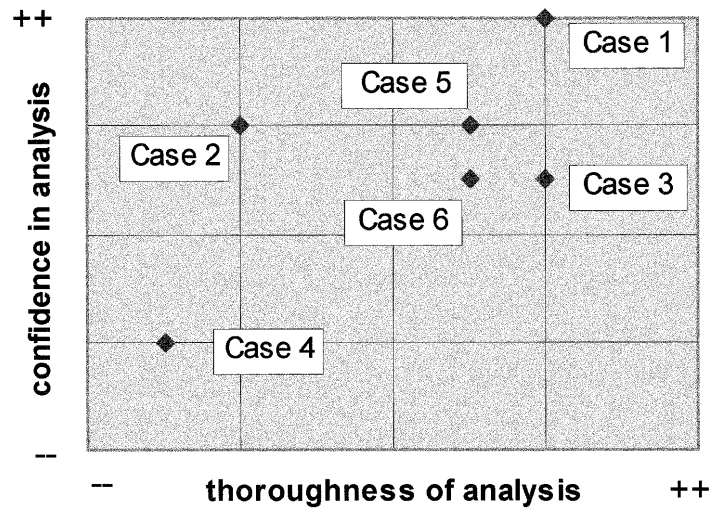
Especially because of one or two outliers in the case evidence, the results here start to display somewhat more "surprise behaviour".

- *Thoroughness → confidence.* For instance, it would seem obvious that, the more thorough participants perceive the analysis to have been, the more confidence they would have in the resulting recommendations. This is the overall picture that comes across from glancing at Figure 12. However, there are two obvious exceptions to the rule. The first is Case 3, where the analysis was considered to be fairly thorough, yet confidence was only modest because hardly any real-world data had been available to test the model recommendations (this was a new company with a unique product in a new market). The other is Case 2, where there was no quantitative validation, but participants still felt they had the right model.
- *Simulation → insight.* The claim is an obvious one and has long been a cornerstone of system dynamics modelling: through conducting simulation experiments one learns about a problem. If we look at Figure 13, we see that this relation was upheld in five out of six cases. In Case 2, participants indicated that they had learned a great deal about the issue at stake, i.e., reasons for lack of collaboration between business units, although there was no simulation modelling employed. So perhaps we should conclude that this relation is confirmed, but that it works only in the positive sense: simulation does lead to insight, but there are other ways to gain insight from a model than quantified simulation. As Case 2 illustrates, a qualitative model can also lead to substantial learning. One really would need more qualitative cases than the two shown in
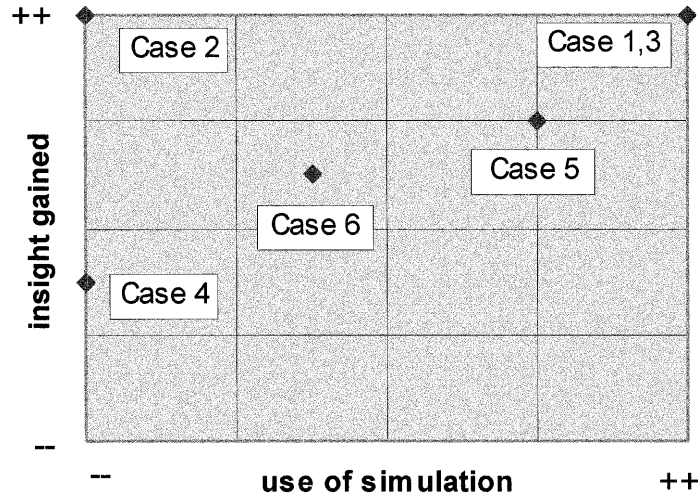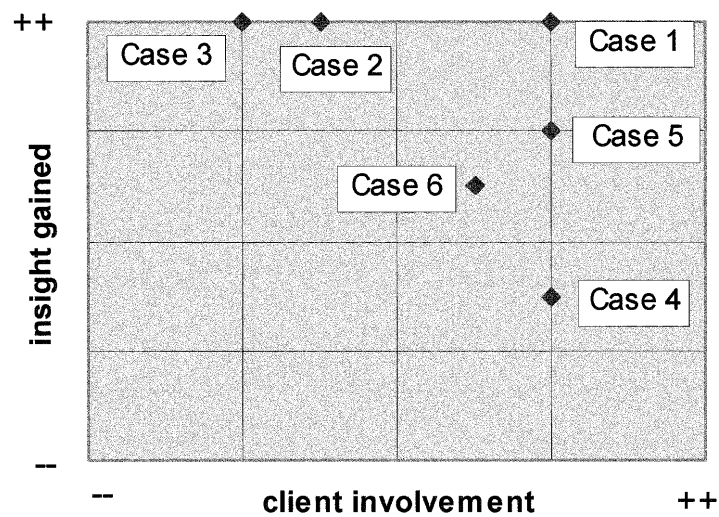
Fig. 13. Simulation →
insight

insight gained (vertical axis, ++ to --)

Case 2

Case 1,3

Case 5

Case 6

Case 4

use of simulation (horizontal axis, -- to ++)

Fig. 14. Involvement
→ insight

insight gained (vertical axis, ++ to --)

Case 3

Case 2

Case 1

Case 5

Case 6

Case 4

client involvement (horizontal axis, -- to ++)

this graph to assess the validity of this relation properly (and, of course, far more quantitative cases as well).

- *Involvement leads to insight.* The idea that involvement in a modelling process generates learning and insight is another cornerstone of all participatory modelling techniques. One first observation from Figure 14 is that the scores for involvement are disappointing. One interpretation of these data could be that the client will generally learn a great deal, whether everybody is deeply involved or not. In the majority of cases high scores for insight were found, regardless of involvement levels. For instance, in Cases 1–3 the score for insight is " + + ", but this high level of insight is achieved with involvement levels ranging from "-" to

Fig. 15. Top manage-
ment support →
involvement
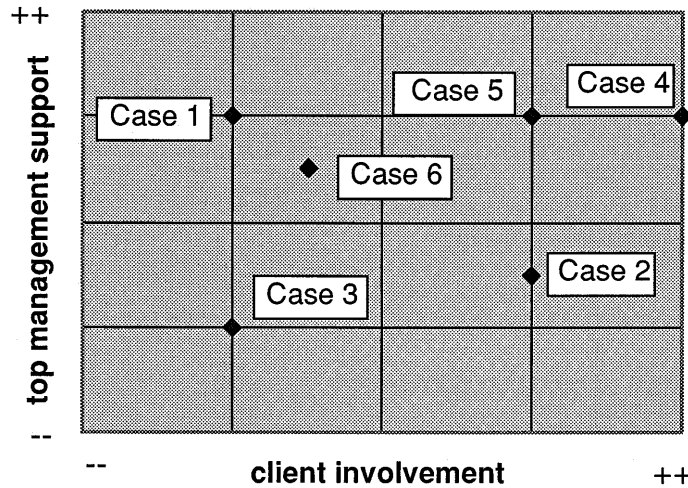
"+", i.e., across the entire range of involvement encountered in the cases.
Another interpretation is related to the way involvement was defined and
measured in this study. We will refer to this in the next section, when we discuss
the relationship between involvement and ownership.

REFUTED TEXTBOOK THEORY RELATIONS. Relations can fall in this category because of
two reasons: either the research was faulty or the textbook theory needs
modifications. Of course, it is much easier, and in most cases also more correct, to
assume the former for an exploratory research project like this one. Nevertheless,
let us look at the evidence.

- *Top management support leads to involvement.* A top manager can force all
  stakeholders to the table. The best illustration of this is provided by Case 4, but
  Case 5 is another example of the textbook theory, as Figure 15 indicates. Cases 3
  and 6 exemplify the negative version of the theory: low top management support
  leads to unsatisfactory involvement. Unfortunately, Cases 1 and 2 do not fit the
  theory at all. Therefore, it has to be accepted that we can draw no conclusions on
  the basis of these data: involvement can be good ("+") for values for top
  management support across the whole range from "-" up to "++". There may
  still be a positive relationship between these two variables in general, but this
  effect does not appear to have had an overriding influence in these six specific
  cases.
- *Group size → affects speed.* Another textbook classic is that with larger groups,
  sessions in particular and projects in general tend to proceed more slowly.
  However, we find no support for this relation from the data as shown in Figure

Fig. 16. Group size →
speed
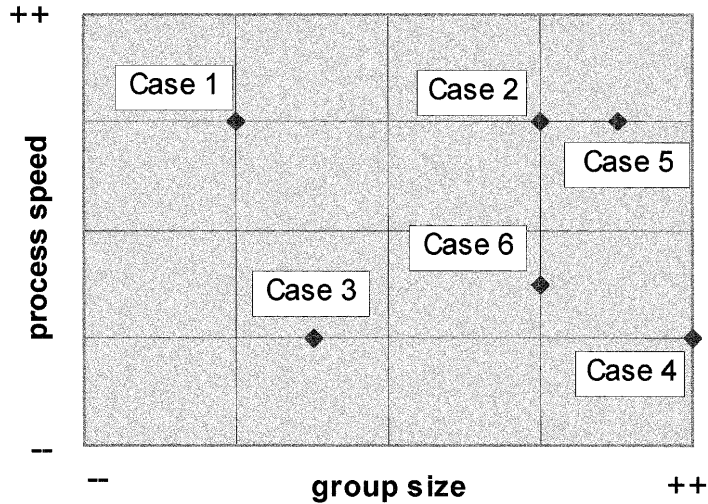
16. Speed can be high in large groups (Cases 2 and 5) as well as in small groups (Case 1). It can also be low in large groups (Case 4) and in small groups (Case 3). The case data offer no evident explanation for this, except that speed was probably more strongly influenced by other factors. If we turn to the individual case evaluations, we find that in Case 3 speed was low because of low involvement, which led to repetitions; in Case 4 it was low owing to a variety of reasons, unwilling participants in a political environment being one of them. In Case 6 speed was perceived as somewhat low, no doubt partly because of the complexity of the issue and probably because the project was still at an early stage. So we find different reasons, none of them related to group size.

   Another possible explanation is that bigger groups do indeed slow down the progress, but that in Cases 2 and 5 the consultants may have found a way of overcoming this handicap.

• *Familiarity with method improves communication.* Several distinguished practitioners in the business modelling world stress the need to make participants familiar with the particular method they use, be it system dynamics or any other methodology, prior to the real modelling process. This is supposed to improve the subsequent process considerably. It sounds like good advice, and probably is too, if one can get the client sufficiently motivated to comply.

   Unfortunately, the case data from this research were of limited value simply because hardly any of the participants were familiar with a PBM method. At most, one or two participants out of a group of eight in Case 4 had encountered similar methodologies, whilst in Case 6 techniques stressing participation and teamwork were said to be common practice within the client company. So the data tell us little, other than that it does seem possible to have excellent
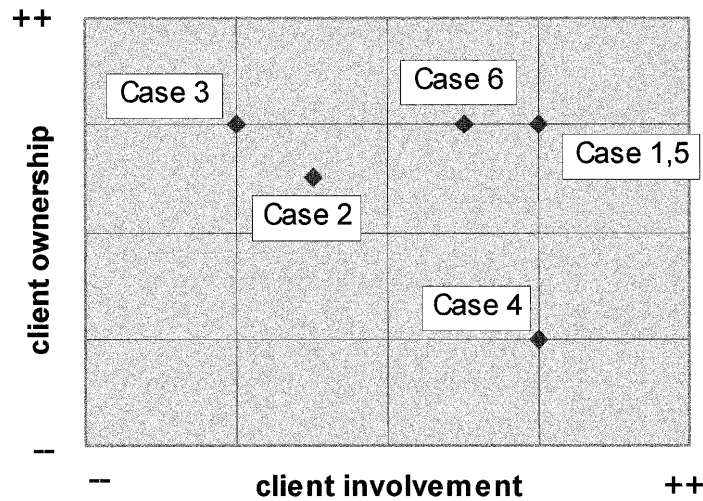
Fig. 17. Involvement
→ completeness
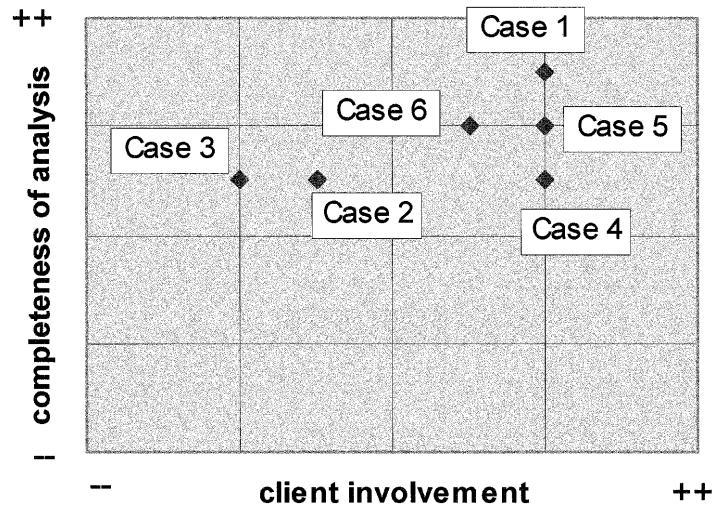
communication in the modelling process without substantial prior exposure to the various aspects of the modelling method.

- *Involvement leads to completeness.* The idea behind this relation was that one needs inputs from all stakeholders to arrive at a complete picture of a problem. This relation cannot be confirmed by our case data, as Figure 17 indicates. For this, we can see two reasons. First, the values for both variables fall within narrow ranges, which impedes a proper assessment of the relation. Second, completeness is not just influenced by involvement, but by other variables such as "problem scope" and "project size". This explains why, within those narrow ranges, there are still wide variations.

- *Involvement leads to ownership.* Full client involvement is an essential ingredient of all group model-building methods. One of the main reasons for this emphasis on client involvement is precisely this relation: if managers are not closely involved in the modelling process, they will feel little ownership for the model; it will not be *their* model. Figure 18 summarizes the top-level case data for this relation.

If we look at the data on this cornerstone of participatory modelling, then we cannot help but to be initially disappointed, for a reasonable inference would be that involvement may lead to ownership, but certainly not always. In Cases 2 and 3 there was (relatively) low involvement but (relatively) high ownership, whilst in Case 4 involvement was high yet ownership was low. So half of the cases contradict this relation, which hardly endorses the assertion that involvement leads to ownership. This is a situation similar to the effect of involvement on insight, as we have seen in the previous section.

Fig. 18. Involvement
→ ownership

The main explanation for this surprising result may lie in the specific interpretation that was given to the terms in this research. The reader will recall that we defined "involvement" as "being part of the project team" or "being present at the modelling sessions". The figure shows that in Cases 2 and 3 involvement (i.e. attendance) was low on average, because not all participants were present at every modelling session, for various reasons. The data at the individual level, however, reveal that those who did attend frequently also felt high ownership. On the other hand in Case 4 attendance was high but ownership is low. That is because in Case 4 all stakeholders were forced to attend the workshops by their CEO, but being present is not the same as active participation or involvement. In other words, Cases 2 and 3 seem to suggest that attendance leads to ownership, while Case 4 reveals that attendance in itself is not a sufficient condition to create ownership. Given these contradictory results, it seems that another (third) variable must be at play. Based on the results of the exploratory relations, which will be discussed in the next section, we assume that it is the quality of communication which explains the contradiction.

An assessment of strongly confirmed exploratory relations

The preceding section investigated cross-case data for 11 of the 66 relations contained in our research model. These 11 were selected for their special interest as representing relations that were more or less "textbook theory". But what about the other 55 relations? They too ought to be worthy of interest, especially any
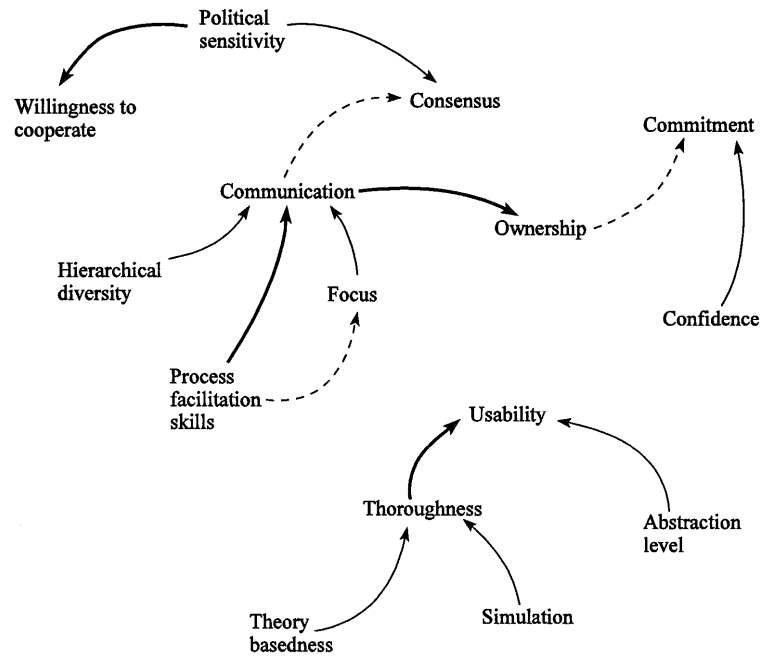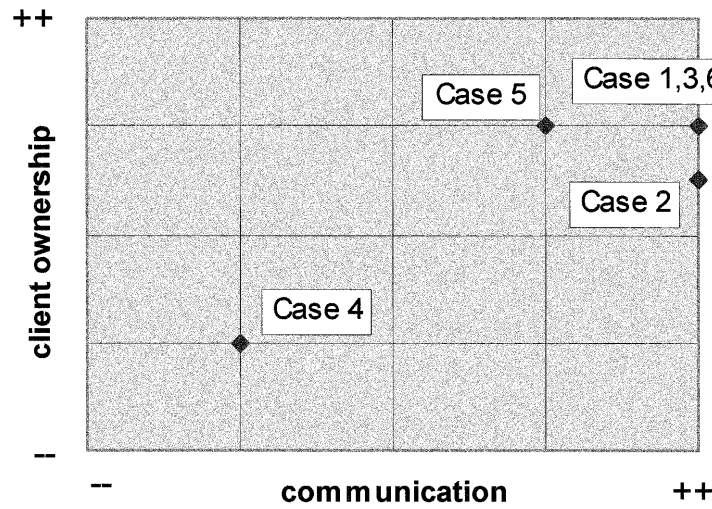
Fig. 19. Exploratory
relations

Fig. 20.
Communications →
ownership

relations that might be substantiated by cross-case analysis. In all, 14 such
relations, shown in Figure 19, were clearly confirmed.

For reasons of completeness Figure 19 also contains dashed arrows of
relationships that were discussed in the previous sections. In the remainder of this

Fig. 21. Thoroughness
→ usability
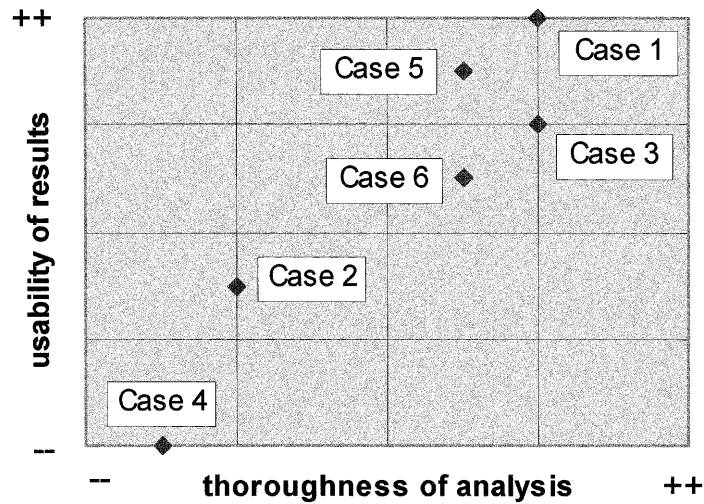
section we will discuss four of the remaining relationships (thick arrows), because they seem important to better understand the effects of group model-building.

- *Communication → ownership.* Our revised version of "textbook theory" asserts that involvement in the form of being present at the model-building process is not enough; only if people communicate openly and effectively in the course of the modelling process, will they develop a feeling of ownership for the resulting model. The case data shown in Figure 20 appear to confirm that relationship.

    If we look at the data case by case, we notice that whenever communication was good, so was ownership and vice versa. In Case 4 participation was forced, communication quality was low and there was no feeling of ownership. In Cases 2 and 3 communication quality was high and there was also a high level of ownership. In Case 2 ownership was somewhat lower than in Case 3. This may have been due to the somewhat lower level of consensus that was reached in that project, but it may also be due to the lower feeling of control over the decision to be made by someone else than the group (see Vennix *et al.* 1996).

- *Thoroughness leads to usability.* The general idea behind this relationship is that, the better your model is, the greater the utility of its findings will be: "model quality improves usability". But because usability is part of the overall concept of "model quality" in the research model, a surrogate variable had to be found, and the two key components, "completeness" and "thoroughness" were introduced as substitutes (see Table 1). The effect of completeness was only partially confirmed, but the effect of thoroughness was strongly confirmed, as Figure 21 reveals.

    Thoroughness was low in two instances (Cases 2 and 4), neither of which led

Fig. 22. Political sensi-
tivity → willingness
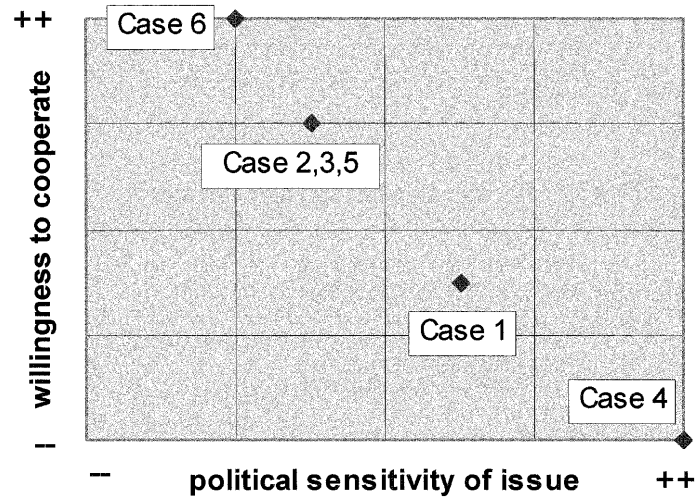
to directly usable models. In three other cases thoroughness was good ("+"), which coincided with fairly usable up to very usable results. Case 5 illustrates that it is possible to have a model that is not thoroughly analysed and yet is fairly usable.

- *Political sensitivity decreases willingness to cooperate.* Politically sensitive problems pose career risks for participants and, as a consequence, this relation postulates, they will tend to be unwilling to cooperate in the modelling process. Figure 22 shows the case data on this negative relation. The data plot is especially nice, showing an almost precise linear negative relation. Moreover, since the range of values is quite broad for both variables, one can safely say that these data provide strong supportive evidence for this relationship. Finally, it is worth noting that in the majority of cases political sensitivity was not very high.

- *Process facilitation skills improve communication quality.* This is also an interesting relationship, which emerges from the case data. It seems to be in line with other research which suggests the importance of the facilitator role in group model-building projects (see Vennix *et al.* 1993; Phillips and Phillips 1995). The relationship is important because it (in)directly also affects consensus building and commitment (see also Korsgaard *et al.* 1995).

Conclusion and discussion

The study reported in this paper has revealed some interesting results with regard to two issues, first with regard to relationships that were rejected by the case study

data, and second with respect to more insight into factors responsible for accomplishing goals of group model-building interventions, i.e., learning, consensus and commitment. When it comes to the first point, this study has cast doubt on the following "textbook" relationships:

- Group size impedes the speed of the group model-building process. Obviously there are ways to mitigate this effect by using effective procedures and increasing focus.
- Top-management support is maybe a necessary but not a sufficient condition to guarantee real involvement. Participation in the process does not necessarily lead to ownership of the model. Rather ownership seems to be created through high-quality communication.
- Familiarity with the modelling methodology does not seem to be a prerequisite for effective communication.

Future research may attempt to find out whether these conclusions are valid and to gain more insight into the underlying processes.

When we turn to the second issue, the most important claims that have been made with regard to group model-building are that it increases learning, helps to build consensus and commitment and improves system performance. The six cases from this study indicate that, with one exception, all cases were succesful in creating insight and in building consensus and commitment. Actual changes or improvements in system performance due to the intervention could not really be studied, although in the first two cases people report that results have been implemented and system performance has improved. However, we do not have hard empirical evidence for that.

When it comes to learning and insight, with the exception of Case 4, in all cases participants say that they gained more insight into their problem. In the present study, the question whether learning occurred and new insights were generated was established by participants' self reports. These may be questioned for at least two reasons. First, it may be the case that people are under the impression that they did learn, while in fact this is dubious (see, for instance, Naftulin *et al.* 1973). Second, people may report that they did learn, because they invested time in the project and it is shameful to admit that it was a waste of time. However, in studies in which learning from simulation models is established in a more objective way (e.g., by multiple choice tests) similar results are found (see Vennix 1990; Verburgh 1994). On the other hand, we do have to point out that neither of these studies was able to find significant improvement in mental models in any formal sense, e.g., people entertaining more feedback loops or creating more explicit causal argumentations (see also Andersen *et al.* 1994).

In addition, as we have seen, even when it is assumed that learning and insight take place, it is not quite clear what factors cause insight to occur. In general, it is assumed that learning is enhanced when simulations are carried out and when involvement is high. Both these relationships received only partial support from the data. Partial support for involvement may have been due to the way it was defined and measured. Partial support for simulation suggests that substantial learning may also occur in cases when no simulations are conducted.

So, far from being sorted out, the loosely employed concepts of "learning", "insight" and "change in mental models" are quite opaque. Conceptual clarification is clearly needed. Future studies should thus be quite specific about what they mean when they suggest that people learn from simulation models, or that simulation changes their mental models. Moreover, research should focus on the factors that may be held responsible for insight to occur.

With regard to consensus and commitment to the decision, this study suggests that both are in a significant way related to the process and the quality of communication within the group. Reaching consensus seems to be directly related to communication quality, while commitment is affected by the creation of ownership over the model through good communication. In turn, it seems that the skills of the group facilitator significantly affect the quality of communication. Hence, when it comes to learning, the simulation model may seem more important, but, when it comes to building consensus and commitment, the role of the group facilitator may be more significant. Given the importance of consensus and commitment for effective decision-making (Senge 1990; Vennix 1996), future studies should, among other areas, focus on the role of the facilitator and attempt to identify (in)effective facilitation behaviour with regard to building consensus and commitment.

References

Akkermans, H. A. 1994. Participative Business Modelling to Support Strategic Decision Making in Operations — A Case Study. *International Journal of Operations & Production Management* 13(10): 34–48.

———. 1995a. Modelling with Managers. Participative Business Modelling for Effective Strategic Decision-Making. Doctoral Dissertation, Eindhoven University of Technology.

———. 1995b. Developing a Logistics Strategy through Participative Business Modelling. *International Journal of Operations & Production Management* 15(11): 100–112.

———. 1995c. Quantifying the Soft Issues. A Case Study in the Banking Industry. *Proceedings of the 1995 International System Dynamics Conference, Tokyo, Japan* 313–322.

Akkermans, H. A. and J. Bosker. 1994. Design Guidelines for Participative Business Modelling Projects; Lessons from an Unsuccessful Case Study. *Proceedings 1995 International System Dynamics Conference. Problem Solving Methodologies, Stirling, Scotland*, 15–24.

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, Social and Public Policy*, 11–18.

Cavaleri, S. and J. D. Sterman. (1997). Towards Evaluation of Systems Thinking Interventions: A Case study. *System Dynamics Review,* 13(2). In the press.

Hutjes, J. M. and J. A. Van Buuren. (1992). *De Gevalsstudie: Strategie van Kwalitatief Onderzoek* (Case study: strategy for qualitative research). Meppel, The Netherlands: Boom.

Korsgaard, M. A., D. M. Schweiger and H. J. Sapienza. 1995. Building Commitment, Attachment, and Trust in Strategic Decision-making Teams: the Role of Procedural Justice. *Academy of Management Journal* 38(1): 60–84.

McGrath, J. E. 1984. *Groups: Interaction and Performance.* Englewood Cliffs, NJ: Prentice-Hall.

Miles, M., and A. M. Huberman. 1984. *Qualitative Data Analysis. A Sourcebook of New Methods.* London: Sage.

Mintzberg, H. 1994. *The Rise and Fall of Strategic Planning.* New York: Prentice Hall.

Morecroft, J. D. W. and J. D. Sterman, (ed.) 1994. *Modeling for Learning Organizations.* Portland, OR: Productivity Press.

Naftulin, D. H., J. E. Ware, F. A. Donnely. 1973. The Doctor Fox Lecture: A Paradigm of Educational Seduction. *Journal of Medical Education* 48: 630–695.

Phillips, L. D. and M. C. Phillips. 1995. 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.

Schein, E. H. 1969. *Process Consultation: Its Role in Organization Development.* Reading, MA: Addison Wesley.

Senge, P. (1990) *The Fifth Discipline: The Art and Practice of the Learning Organization.* New York: Doubleday.

Vennix, J. A. M. 1990. *Mental Models and Computer Models: Design and Evaluation of a Computer-Based Learning Environment for Policy Making.* Doctoral Dissertation, Catholic University Nijmegen.

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

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

Vennix, J. A. M., W. Scheper and R. Willems. 1993. Group Model-building: What Does the Client Think of it? *Proceedings of the 1993 International System Dynamics Conference*, Cancun, Mexico: 534–543.

Verburgh, L. 1994. *Participative Policy Modelling, Applied to the Health Care Insurance Industry.* Doctoral Dissertation, Catholic University Nijmegen.

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