A method for eliciting and comparing causal maps

Lívia Markóczy[1]
Cranfield School of Management
Cranfield University
Cranfield, Bedfordshire MK43 0AL
England
Tel: (44-234) 751 122
Fax: (44-234) 751 806

Jeff Goldberg
Theoretical Linguistics Department
Budapest University

February 9, 1995

[1]Correspondence should be directed to Lívia Markóczy.

Abstract
Causal mapping has long been valued as a tool for exploring individuals' idiosyncratic beliefs. In this paper we do three things to make these benefits more utilizable by researchers. First, by detailing a replicable elicitation technique, we lay the ground for systematic and objective comparison of these beliefs: Systematic analysis and comparison require systematic elicitation. Secondly, through a careful analysis of what a causal map means, we have been able to refine a way of comparing maps which uses all of the available information within the maps and is objective in the sense that all subjective aspects of the method are made explicit (e.g., the pool of constructs, the interpretation of missing information, etc). Thirdly, we discuss and extend various tools for analyzing the distance based measures which the comparison yields. In addition, a brief example of the use of the method is sketched.

Bibliographical Note
Lívia Markóczy is starting as a Research Fellow at Cranfield School of Management, Cranfield University. She is currently completing her PhD at the Judge Institute of Management Studies, University of Cambridge. She has a 'doktori' degree (1986) in Business Economics from the Karl Marx University of Economics, Budapest (since renamed Budapest University of Economics), where she has also been a member of the faculty since 1986. She is best reached at `kvgt001@ursus.bke.hu`.
Jeff Goldberg teaches theoretical syntax and computational linguistics in the Theoretical Linguistics Department of Eötvös University (Budapest), and is also employeed by the Research Institute for Linguistics of the Hungarian Academy of Sciences. He has a BA in linguistics from the University of California, Santa Cruz (1984), and persued a number of years of graduate study in linguistics at Stanford University. He is best reached at `goldberg@nytud.hu`.

Acknowledgements
This work was supported by a grant F107 from the Hungarian Scientific Research Fund (OTKA). The authors would also like to thank those who participated in the reliability tests, István Bartók and Ferenc Tóth; those who gave us detailed comments on previous drafts, anonymous reviewers and Matthew Jones; and those who gave us specific pointers to the statistical literature, Erzsébet Kóvacs, and Márton Gyombolai. All errors remain are own however.

"What's the good of Mercator's North Poles and Equators,
Tropics, Zones, and Meridian Lines?"
So the Bellman would cry: and the crew would reply
"They are merely conventional signs!

"Other maps are such shapes, with their islands and capes!
But we have our brave Captain to thank"
(So the crew would protest) "that he's bought us *the best—*
A perfect and absolute blank!"

— LEWIS CARROLL *The hunting of the snark* (1876)

1 Introduction

Causal mapping techniques have been used in management research for gaining insight into the belief systems of managers, but have yet to be used (with a few notable exceptions) for studies directly linking managerial cognition either to other characteristics of the manager (such as age, work experience, etc) or to action (decisions made, organizational performance, etc). Hall (1984) and Barr, Stimpert and Huff (1992) are among the exceptions and have tried to link causal maps to behavior.

One barrier to this kind of work is the lack of techniques for systematically comparing causal maps in a way that uses all of the information contained in such a map. Previous comparisons were either made based on subjective researcher judgement (e.g., Barr et al., 1992), or used only a small amount of the information contained in maps (e.g., Ford and Hegarty, 1984) who compare the out-degrees (relative influence) of particular constructs).

Langfield-Smith and Wirth (1992) have made an important advance for work with causal maps by describing a method for obtaining a single distance measure between causal maps using all of the information available within each map. However, there are three things which need to be added to their work for such a technique to be rendered fully usable by management science researchers. (1) Since any systematic comparison of maps crucially depends on the meaning of the (parts of) the maps, which in turn depend on how the maps are elicited, sufficient attention needs to be paid to elicitation techniques. (2) Their distance ratio formula needs to be modified in two directions: One is to make the formula customizable, so it could be used by researchers who elicit maps which are not exactly like the ones used by Langfield-Smith and Wirth (1992); the other is to rethink the interpretation of missing information. (3) Methods for analyzing the distance data resulting from the comparison need to be extended and developed, since tools for analyising distances remain fairly sparse. The goal of the present study is to do all three of the above at a level of sufficient detail to enable researchers to use the method. The elicitiation technique is the subject of section 4, the comparision technique is the subject of section 5, the analysis of distance measures is the subject of section 6, and the other tools of analysing the maps are the subject of section 7. Section 8 briefly reviews an example.

Before proceeding with the elicitation technique, we first provide some working definitions concerning what causal maps are and what they mean.

2 Causal maps

Causal mapping has been used for investigating managers and decision makers since Axelrod (1976) introduced it to management studies. Since then, many varieties and usages have appeared. Huff (1990), Eden (1992), and Laukkanen (1992) provide a good overview of the current state of the field. From these varieties we have chosen one form of causal mapping. The reasons and issues surrounding our choices can only be discussed in terms of the use that we put the maps to; and, therefore, these issues are not discussed in this section, but in the relevant subsections of section 4 as they arise.

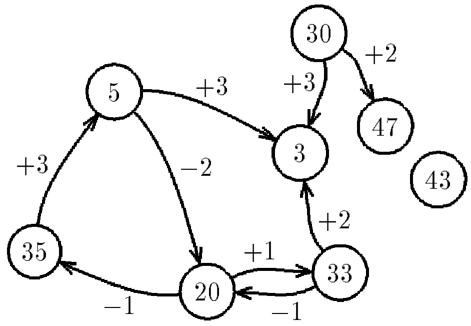
2.1 Terminology and examples

In our definitions below, we restrict the term 'causal map' to those particular objects we elicit or construct from individuals, and as such they are only partial descriptions of belief structures. **Causal maps** are representations of individuals (or groups) beliefs about causal relations. They include elements, with only two kinds of properties. The first property is 'relevance'. The second is the possibility of being in one (of two) 'influence relationships' (positive or negative) with one (of three) strengths (weak, moderate, or strong). Note that the number of strengths (three in our case) and the types of influence relations (positive and negative) are not crucially part of the definition of causal map.

Factors which the individual thinks are relevant correspond to **node**s and each node corresponds to a description called a **construct label**. The map in figures 1 and 2 for example, contains the node 35 (n_{35}) which is accompanied by the construct label 'brand recognition'. Another node in the map is 5 (n_5), labeled by 'market share'. The arrow in figure 1 from n_{35} to n_5 indicates that there is a belief that the level, degree or quantity of 'brand recognition' influences the level degree or quantity of 'market share'. The arrow from one node to another represents an **arc**. In the text we will represent arcs such as the one from n_{35} to n_5 by '$n_{35} \xrightarrow{+3} n_5$'.

The terms 'arc' and 'nodes' apply to causal maps irrespectively of whether the maps are presented as diagrams as in figure 1 or as association matrices as

Figure 1 Diagrammatic presentation of causal map



3	Vision and strategic direction	5	Market share
20	Competition in market	30	Leadership in organization
33	Knowledge of market needs	35	Brand recognition
43	Bank connections	47	Efficiency/productivity

in figure 2. Other terminology is also in use for 'nodes' and 'arcs'. They have been referred to as 'vertices' and 'edges' (e.g., Thulasiraman and Swamy, 1992), as 'elements' and 'beliefs' (e.g., Langfield-Smith and Wirth, 1992), or as 'nodes' and 'links' (e.g., Bougon and Komocar, 1990). Axelrod (1976) uses the terms 'points' and 'arrows' which has led a number of researchers to mistakenly believe that they are only properties of the diagramatic presentation of the maps.

Each arc has a number associated with it. In the case of the arc from n_{35} to n_5 that number is +3. The sign (positive or negative) of this number is the **polarity** of the arc. If it is **positive**, then the influence is positive: an increase in n_{35} leads to an increase in n_5 and a decrease in n_{35} leads to a decrease in n_5. If the polarity were **negative** then an increase in the first would lead to a decrease in the second and a decrease in the first would lead to an increase in the second.

The value of the number associated with an arc indicates the **strength** of the relation: 1 for weak, 2 for moderate, and 3 for strong. In the figures, n_{20} (competition in the market), has a moderate negative effect on n_5 ($n_{20} \xrightarrow{-2} n_5$).

Two ways of displaying causal maps illustrated in figures 1 and 2 are the **diagramtic** and the **association matrix** forms respectively. Where a node in the diagram is a circle with a number in it, the node in the matrix is a column and row, with the column headed by a number. An arrow between two circles corresponds to an arc in the diagram, while a matrix cell with a non-zero value

Figure 2 Assoxciation matrix presentation of causal map

	30	47	3	5	33	43	20	35
30	0	2	3	0	0	0	0	0
47	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
5	0	0	3	0	0	0	0	0
33	0	0	2	0	0	0	−1	0
43	0	0	0	0	0	0	0	0
20	0	0	0	−2	1	0	0	−1
35	0	0	0	3	0	0	0	0

3	Vision and strategic direction	5	Market share
20	Competition in market	30	Leadership in organization
33	Knowledge of market needs	35	Brand recognition
43	Bank connections	47	Efficiency/productivity

corresponds to an arc. In the the diagram the direction of the influence follows the direction of the arrow. In the matrix the direction of the influence is from row to column. For example, a cell in the matrix such as $m_{8,4}$ will contain the label of the arc in this example from n_{35} to n_5. In this example that is +3.

The **expanded association matrix** is a varient of the association matrix, which contains extra rows and columns as needed so that row and column 5 will correspond to n_5 and row and column 35 will correspond to n_{35}. The extra rows and columns which do not have corresponding nodes in the map will be filled with zeros, and some record will be kept (a set of nodes, for example) of which nodes are in the map. We do not display one of these as it is not a very convenient display form and, for our example, it would require a 47 by 47 matrix at least. However, the formula for calculating distances between causal maps (detailed in 5) can most easily be stated in terms of expanded association matrices and sets of included nodes.

2.2 Interpreting causal maps

As made very clear by Axelrod (1976, chap. 6), CMs are not magical entities. If we construct CMs from individual's stated beliefs, then a CM is only about her/his system of stated beliefs. As Axelrod says:

> The cognitive mapping approach is, of course, in no better (or worse) position in this regard than any other procedure that relies on a person's conscious and monitored linguistic behavior to make inferences about his or her beliefs. [p. 252]

This is part of a wider point, which, in its obviousness, is easy to overlook. *The meaning of a causal map is not only a function of the map itself, but of the way in which it is elicited.* The elicitation technique described below provides the real meaning or our particular CMs.

3 Preview of the method

Several aspects of the method we describe may be unfamiliar. Therefore, before jumping into a detailed description of the method we will provide an overview of the method so that the reader can understand 'where we are going' in what follows.

1. Develop a *pool of constructs* by conducting and analyzing interviews with managers and a review of relevant literature. This is done prior to the study so that each subject selects constructs from the same pool.
2. Have each subject select a fixed number of *constructs* by identifying items from a constant pool of constructs.
3. Construct the *causal map* of each individual subject by having her/him assess the influence of each of her/his selected constructs on her/his other selected constructs.
4. Calculate *distance ratios* between causal maps using a generalized version of Langfield-Smith and Wirth's (1992) formula.
5. Perform a variety of statistical tests on the distance ratios to identify what *characteristics* account for similarities in thinking.

Steps 1–3 will be discussed on the corresponding subsection of the next section. Step 4 will be the subject of section section 5. The forms of statistical analysis are the subject sections 6–7.

4 Creating causal maps

4.1 Deriving constructs

We first assemble a **pool of constructs**. These are constructs which managers may find influential or related to some notion of the issues which are at the center of the investigations. For example, these could be factors which influence business success. The pool of constructs could be derived, in part, from interviews with individuals in similar positions and organizations to those used in the final sample. To ensure that the pool of constructs covers a broad domain it is always advisable to complement these constructs with items drawn from the relevant literature. It is always advisable to 'test' the list on a pilot sample, and survey subjects as to whether they felt that any thing was missing. Furthermore, during the course of the study itself, one can ask subjects whether anything is missing as a way of confirming the completeness of the list. Although Ford and Hegarty (1984) and Bougon, Weick and Binkhorst (1977) used fixed lists of constructs, they did not give subjects the opportunity to select from the list. The present method, on the otherhand, preserves many of the benefits of using fixed lists, while still providing subjects with a large range of responses.

As seen in the next section, the number of constructs does not need to be constrained in principle. All that must be done is avoid duplication of constructs, and make sure that the ground is well covered. Because of this, the number of constructs need not be limited to those elicited by interview; constructs can be added to the list as seen fit, excluding no potential source of items. Here unimportant and even absurd items do little harm, as long as each individual surveyed has the same opportunity to select items important to her/him. In practice a constraint on the number of constructs is desirable. If there are too many items, the selection process may become too difficult. In one preliminary trial of selection techniques (with 64 items on a list) we found that most subjects would not notice duplicate items. Another difficulty is that we might not get enough overlap of selected items to meaningfully compare the maps. We suspect that somewhere between 40 and 50 items is correct.

The subjects will select a small number of items from the pool of constructs. By letting the individuals select their constructs, more scope for differences between subjects can be expressed. But by having them select from a predefined list, we avoid the coding problems of lets say one person using *income* and another using *earnings*. If subjects were not presented with a pool of constructs, we would have no way of knowing whether *income* and *earnings* should be coded together or not. This question will not arise since either one of the two is on the list; or, if the researcher puts both on the list, S/he will have reason to believe that the choice of one instead of the other is meaningful.

By creating this finite pool of contructs, we are choosing to deal with the inevitable coding problem prior to elicitation as opposed to after elicitation. The merits are that it makes coding much easier (the list only needs to be created one, and so can be done very carefully), the other is that it ensures that each subject is presented with the same stimuli. The disadvantage is that it percludes the elicitation of novel constructs.

4.2 Selecting the constructs

Each manager is asked to select those constructs from the pool of constructs which they believe are most important to the same issue (e.g., business success) which was used for constructing the pool of constructs. They will be asked to identify the 10 most relevant constructs. The reason for limiting them to only 10 constructs is purely practical: Subjects will later be asked to consider the influence relationship between each pair of constructs. Ten constructs will lead to 90 influence relations being considered.

Because subjects are limited to a small number of constructs, we cannot infer from their non-selection of a construct an 'unbelief'. This is unfortunate, because it makes the interpretation and comparison of CMs more difficult. We do not wish to underrate this problem (and we will come back to it), but we are not alone with this sort of problem with CMs. The meaning given to the non-presence of a node crucially depends on how the nodes are elicited. Clearly the non-presence of a node in the CM cannot be taken to mean that the subject believes the construct to be irrelevant, it only means that it is not in the top ten. We will discuss this at greater length in section 5.

We place each item along with its definition on a card and ask subjects to remove items which are important (positively or negatively) to issue used in constructing the pool of constructs. The important pile will again be subject to this type of sorting 'most important' versus 'least important'. This procedure will continue until the subject has a pile of 10 or fewer cards. S/he will then be asked if necessary to re-select cards from the most recently eliminated pile to make up a total of 10. The cards are initially presented to the managers in a random order, but the same random order is used with each subject.

Using this sort of selection task has a number of advantages over an alternative of asking subjects to select from a written list. First, the ability to physically manipulate the cards means that the subjects have to keep fewer things in mind when selecting. Second, it avoids the tendency of some subjects to make sure to select items from 'every category' and avoid selecting non-independent items. Third, by making the selection process more indirect it is more difficult for the subjects to be insincere.

The researcher can also get an idea of partial rankings of all of the constructs by noting which cards are selected out first. Our experience from pilot studies is that subjects become annoyed or impatient if no record of each selection stage is kept.

4.3 Eliciting the Arcs

For each pair of selected constructs, subjects will be asked if one construct influences the other, whether it does so positively or negatively and whether it does so weakly, moderately, or strongly. All of these questions need to be made more precise, otherwise, different interpretations of the questions will lead to an unacceptable level of experimental error. Roberts (1976) reports that experts demanded fairly full definitions of these questions and terms and our experience confirms this.

We should be neither surprised nor concerned if we find that our maps are 'unbalanced' in Axelrod's (1976) terms. An unbalanced map would contain arcs like n_1 positively influences n_2, which in turn positively influences n_3, while n_1 *negatively* influences n_3. This does not imply any inconsistency in the belief system of the individual whose map it is, since there may be other constructs which are not included in the map. But even if the map were somehow complete, we should not expect balance. Unbalanced maps are not the consequence of inconsistent beliefs, but only of the fact that composed influence relations can be complex.

Subjects feeling that they are being tested may be biased toward providing balanced maps. This bias, if it exists, is countered by telling them not to worry about consistency and by presenting pairs using Ross ordering (Ross, 1934) to maximize the time between questions involving the same elements. In addition subjects are not allowed to review their previous responses. The use of a computer program here to prompt for the questions releases the researcher from the task of identifying which pair should be asked next. It also appears to generate faster (and possibly less rationalized) responses from the subjects.

5 Comparing Maps

5.1 Langfield-Smith and Wirth's formula

Langfield-Smith and Wirth (1992) (henceforth L&W) develop a series of formulas leading up to their most complete formulation in their formula 12 (p. 1149) (henceforth the L&W formula). Their formula provides the basis for what we propose, and therefore merits some discussion, but for a detailed discussion of its meaning and construction the reader is referred to their paper. We quote their formula in figure 3, but have modified some of the variable names, reorganized some clauses to more clearly represent the meaning, and have added the tags '(i)' and '(ii)' to make it easier to discuss the parts of the formula.

Before continuing with the substantive modifications we propose, it is necessary to provide a few brief notes about what different parts of this formula do, without going into detail as to how they do it.

The formula yields a **distance ratio** (DR) which is a number between 0 and 1 between pairs of maps (represented as expanded association matrices, A and B). If the DR is 0, then the maps are identical, if the DR is 1, then the

Figure 3 Langfield-Wirth & Smith's formula 12

$$DR = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p}\left|a^*_{ij} - b^*_{ij}\right|}{6p_c^2 + 2p_c(p_{u_A} + p_{u_B}) + p_{u_A}^2 + p_{u_B}^2 - (6p_c + p_{u_A} + p_{u_B})}$$

$$\text{where } m^*_{ij} = \begin{cases} 1 & \text{(i)} \quad \text{if } m_{ij} \neq 0 \text{ and } i \text{ or } j \notin P_c \\ m_{ij} & \text{(ii)} \quad \text{otherwise.} \end{cases}$$

Where p is total number of possible nodes, P_c is the set of nodes common to both maps, p_c is the number of such nodes, p_{u_A} is the number of nodes unique to map A and p_{u_B} is the number of nodes unique to B. m_{ij} is the value of the ith row jth column in the extended association matrix M.

distance between the maps is maximal.

The general idea of the formula is to sum up (node by node, and arc by arc) all of the little differences between the maps (represented as extended association matrices A and B), and then divide that sum by the greatest difference possible given the number of nodes in each map and the number of nodes common to the maps. Thus the denominator of the formula is that total possible difference (this will ensure that the DR is never greater than 1).

Differences, according to L&W, can be of three sorts: '(a) differences in the strengths of the commonly-held beliefs; plus (b) the differences due to existence or non-existence of beliefs involving common elements; plus (c) the difference due to beliefs consisting of unique elements.' (p 1148). Their use of the terms 'element' and 'belief' corresponds to our use of 'node' and 'arc'.

Clause (ii) of their formula deals with the differences of sort (a), since it will apply when both maps A and B have nodes i and j and so the contribution to the total distance for matrix cell ij will be just the difference between those cells. Clause (ii) also covers the case where neither map has either node i or j; here there is no difference and clause (ii) correctly reflects that since zero minus zero is zero, so nothing is contributed to the total distance.

Clause (i) covers the situation where one map has an arc between nodes i and j, but the other map doesn't because the other map does not include i or j. In this case the total contribution to the distance is independent of the strength of the arc, but is merely 1 if an arc exists.

It is important to note that the L&W formula will treat the absence of an arc between two existing nodes as being a similar to the absence of an arc due to absence of a node. That is, if map A has both node i and node j, but has no arc between them, and B is missing either i or j (or both), the difference will be zero since these will fall under clause (ii) and zero minus zero is zero.

The treatment of missing nodes and missing arcs is one of the most problematic issue in devising a truly meaningful distance ratio. The correct interpretation of these absences depends crucially on how the maps are elicited. One of our extensions of the L&W formula address this issue more directly.

5.2 Modifications and extensions

There are two short-comings of the L&W formula. The first is generalizability and the second is treatment of missing nodes.

L&W's formula does not transparently generalize to maps of slightly different types. For example, it is not immediately obvious what would need to be done to the formula for maps which had four instead of three strength values. Nor is it transparent what should be done if only strength is considered without polarity (whether the influence is positive or negative); or if self-loops should be considered. Once a full understanding of all the details of formula has been acquired, it is not difficult make the necessary modifications in each individual case. But we propose a parameterized version of their formula which allows the researcher to plug-in, so to speak, the values for the particular variant of causal mapping s/he uses.

The formula in figure 4 provides a measure of the DR between two expanded association matrices, A and B, whose elements are denoted by a_{ij} and b_{ij}. We use the first five letters of the Greek alphabet $(\alpha, \beta, \gamma, \delta, \epsilon)$ for the five parameters.

β is the maximum strength, ($\beta = 3$ for L&W and the maps described in section 2), δ is the additional weight given to a polarity change. If $\delta = 0$ then the difference between an arc with strength 3 and with strength 1 is the same as the difference between $+1$ and -1. δ is the amount to be added to the latter difference. L&W assume that the influence a node has on itself will not differ from matrix to matrix. And their formula relies on that assumption. If one wishes to not compare the values for nodes directly influencing themselves (if it never varies) then set $\alpha = 1$. If there are variable amounts of direct self-influence in the matrices, set $\alpha = 0$. γ is by far the most difficult of these, but essentially it is set for how to interpret matrix cells for which one of the maps cannot have an arc because of a mismatch of nodes. This is discussed in more detail below. γ can be set to 0, 1, or 2. ϵ is the number of possible polarities. For the maps that we have been describing with $+$ and $-$ as the two possible polarities, ϵ should be set to 2. For maps like that those used by Laukkanen (1992) where there is no polarity distinction ϵ should be set to 1. With the parameters $\alpha = 1$, $\beta = 3$, $\gamma = 1$, $\delta = 0$, and $\epsilon = 2$, this formula (more or less)

Figure 4 Generalized Distance Ratio Formula

$$DR(A,B) = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p} \mathrm{diff}(i,j)}{(\epsilon\beta+\delta)p_c^2 + \gamma'(2p_c(p_{u_A}+p_{u_B}) + p_{u_A}^2 + p_{u_B}^2) - \alpha((\epsilon\beta+\delta)p_c + \gamma'(p_{u_A}+p_{u_B}))}$$

$$\mathrm{diff}(i,j) = \begin{cases} 0 & \text{(i) if } i=j \text{ and } \alpha=1; \\ \Gamma(a_{ij},b_{ij}) & \text{(ii) if either } i \text{ or } j \notin P_c \text{ and } i,j \in N_A \text{ or } i,j \in N_B; \\ |a_{ij}-b_{ij}|+\delta & \text{(iii) if } a_{ij}b_{ij} < 0; \\ |a_{ij}-b_{ij}| & \text{(iv) otherwise.} \end{cases}$$

$$\Gamma(a_{ij},b_{ij}) = \begin{cases} 0 & \text{(a) if } \gamma = 0; \\ 0 & \text{(b) if } \gamma=1 \text{ and } a_{ij}=b_{ij}=0; \\ 1 & \text{(c) otherwise.} \end{cases} \qquad \gamma' = \begin{cases} 0 & \text{if } \gamma=0; \\ 1 & \text{otherwise.} \end{cases}$$

Where A and B are two extended association matrices, a_{ij} (or b_{ij}) is the value of the ith row and jth column of A (or B), and where p is total number of possible nodes, P_c is the set of nodes common to both maps, p_c is the number of such nodes, p_{u_A} is the number of nodes unique to map A and p_{u_B} is the number of nodes unique to B. N_A and N_B are the sets of nodes in the maps A and B. α, β, γ, δ, and ϵ are parameters described in the text.

reduces to L&W's formula.

Clause (i) of the 'diff' function simply says that if $\alpha = 1$ reflexive arcs ($n_k \rightarrow n_k$) are not looked at, i.e, ignore the diagonal of the association matrix. Note that α plays a role in the denominator because if we don't compare reflexive arcs, the maximum *possible* difference between the CMs is reduced.

Clauses (iii–iv) apply when comparing arcs which could occur in both maps. That is when both maps have nodes n_i and n_j. In this case, we take the difference between the strengths. If one of the values is positive and the other is negative (clause (iii)), we add δ to the difference. The idea behind the use of δ is that if person A believes that n_1 weakly positively influences n_2 ($n_1 \xrightarrow{+1} n_2$) and person B believes that the influence between the same two nodes is weak, but negative ($n_1 \xrightarrow{-1} n_2$) while a third person believes it is strong and positive ($n_1 \xrightarrow{+3} n_2$) we may wish to say that A is nearer to C than A is to B.

Clause (iv) also applies when there is no potential arc in either map. (There is a **potential arc** between two nodes, if only if the map has both nodes.) The difference between two non-potential arcs will always be 0 because the extended association matrix will always use the same number, 0, for rows and columns for which there is no node. Having clause (iv) serve for two very different cases allows us to write the formula more compactly (a choice made by L&W as well), but it communicates its meaning less directly.

Clause (ii) applies when only one of the maps has a potential arc between n_i and n_j because the other map is missing either n_i or n_j (or both). In L&W's formulation ($\gamma = 1$) a difference of 1 may be added. The intuition behind this is that if one person believes in a particular causal relation, while the other person has no such belief (clause (iic)); that is a difference and 1 is added. By the same reasoning it must be a similarity if one person has a potential arc and it is 0 and the other does not have a potential arc (iib). That is, neither believes in an influence relation there.

The problem with this is that it is drawing an inference about individuals' beliefs based on both the absence of a node (if the node is not in the map then it is not believed to be in any causal relations) and based on the absence of an arc (if an arc doesn't exist between two nodes then there is not believed to be any causal relation). The latter assumption is safe if the map was elicited (as we do) by comparing each pair of nodes in the map. The first assumption is safe if subjects are free to select *all* constructs they find relevant. This is not true for the elicitation technique described above where we ask only for the top 10 items.

We have therefore introduced γ into the clause (ii) type cases. One way to follow the logic is to say that nothing can be deduced from the absence of nodes, and therefore only common potential arcs play a role. This would involve setting $\gamma = 0$. In this configuration, maps with two overlapping arcs can be as similar as maps with 90 overlapping arcs. Distances between maps with no overlapping arcs would not be defined (0/0).

We do wish to assign meaning to the fact that one person selects a construct and another person doesn't, but we don't want the values of the potential arc to play any role at all. We therefore will perform our calculations with $\gamma = 2$ (clause (iic)). This way, differences in nodes will always lead to a larger distance ratio, irrespective of the values the potential arcs when compared to non-potential arcs.

Although our formulation may superficially appear more complicated than L&Ws, it usually reduces to something as simple or simplier when appropriate values for α through ϵ are substituted in.

By using an explicit formula our comparison technique achieves total reliability (of the comparison) in that any two invocations of the formula with the same parameter settings on the same two maps will yield the same results. We have also checked the external reliability of the elicitation technique by eliciting maps from the same individual on two different occasions separated by five weeks and from another individual on two occasions separated by eight weeks. The difference between the two maps of the first individual (according to the formula discussed below) is 0.2302 and is 0.2896 between the second individual's two maps. Distance ratios between maps elicited (with the same list of constructs) from individuals within a single organization average around 0.730 with a standard deviation of around 0.140. In fact after looking at more than 950 distance ratios, we have only seen one that is smallar than these two.

5.3 What DRs don't measure

It is important also to note what CMs or differences between them do not measure. First the choice of items selected by the managers is unranked, and the DR is not equipped to take ranking into account. A second measurement failure is more difficult to cope with. Differences between CMs include differences in the choice of constructs, but it does not include the fact that some constructs are more different then others. Imagine, for example, three maps (A, B, and C) that are identical except that where A has *earnings*, B has *income*, and C has *management's understanding the mentality of the host country managers.* Given our measure of CM, these three maps would be equidistant from each other. It may be possible to provide some notion of how similar items are to one another, and use that information in comparing maps, but it is not a simple task. Our list of items will be designed to avoid synonyms, but unequal differences between constructs will remain.

6 Analysis of distances

We can analyze our results in a number of different ways, and we use several different techniques in what may be unfamiliar ways. The purpose of this section (and one which follows) is not only to justify and explain what we do, but to illustrate how rich the data are and the kinds of questions they can answer or raise. In the following subsections, we progress from more familiar forms of analysis to the more novel.

There are some concepts that are common to all distance based data. Given a sample of M causal maps, the results will be a symmetrical dissimilarity matrix with $N = M(M-1)/2$ distinct distances. These distances can be thought of as defining a space with various properties. The space defined by distances of our sort will not be **metrical** because there is no guarantee that the **triangle inequality** will be satisfied. That is, it is possible for the distance between map A and map C to be greater than the distance between map A and map B plus the distance between map B and map C. In other words, in a non-metrical space the shortest distance between two points is not always a straight line. Our distance ratios do, however, satisfy two other important conditions. They satisfy the **symmetry condition** in that the distance between A and B

is the same as the distance between B and A. Although it may seem difficult to imagine such a condition not being satisfied, social networks illustrate such properties (Breiger, Boorman and Arabie, 1975; Tversky, 1977). Finally there is the **minimality condition** which states that distance are never negative, and 0 if and only if the points are the same point in the space. A space which satisfies all three of these conditions is metric (Everitt, 1990; de Leeuw and Heiser, 1982). A space which satisfies minimality and symmetry but not triangle inequality is said to be **semi-metrical** (de Leeuw and Heiser, 1982, p. 292).

Throughout this and subsequent sections we will occasionally refer to the elicited causal map of the subjects as if they were the subjects themselves. So when, lets say, we talk about the distance between managers with property X, we mean the distance as calculated between the causal maps elicited from managers with property X. In most places it should be clear from the context whether the actual subjects or their maps are referred to. Where there would otherwise be confusion, we will use the more complete form to avoid ambiguity.

6.1 Multidimensional scaling

Multidimensional scaling (MDS) along with its applications has been well described by de Leeuw and Heiser (1982) and Wish and Carroll (1982). We use it here more as a presentational technique than as an analysis technique. The general notion of MDS is that it allows one to approximate semi-metrical data (or data which is metrical but in a large number of dimensions) in a conveniently small number of dimensions (usually 2). Thus, our semi-metrical data can be presented in a two dimensional plane, making it possible for the investigator and others to see directly where the entities are with respect to each other. Figures 5 and 6 shows what MDS output can look like in two dimensions.

The extent to which the positions have been approximated by reducing to lower dimension can be expressed in a number of ways. The most common is the **Stress**, the lower the Stress value, the better the fit. A Stress value of 0 means that dissimilarities between entities in the original data are perfectly correlated with the distances between entities in the reduced dimensions. Since it is easier to fit the data into a higher dimensional space, the Stress value will decrease as the dimensionality of the output increases. Exactly what levels of Stress are considered acceptable is a matter of much controversy (de Leeuw and Heiser, 1982; Wish and Carroll, 1982; Füstös and Kovács, 1989), but if MDS will only be used for presentational purposes, the Stress will just be an indicator to the reader of how large a grain of salt the diagram should be taken with. Therefore, unless there is a tremendous improvement in the Stress value from two to three dimensions, we recommend that data be presented in two dimensions, since the accuracy of the higher dimensional representation will be lost in the difficulty of seeing what is going on. Stress and the number of dimensions to use can be a more substantive issue if various attributes are to be correlated with the dimensions (Wish and Carroll, 1982).

6.2 Cluster analysis

Cluster analysis is probably the most common technique for analyzing distance (or similarity) based data. The general idea of cluster analysis is to identify groups or clusters of entities by examining the distances between them. There are a very wide range of clustering techniques. Different techniques are appropriate for different types of data, and it is therefore useful to combine cluster analysis with MDS so that one can see if the clusters produced really make sense (Everitt, 1990; Breiger et al., 1975, p. 103). The large number of techniques may indicate that there is a fair amount of arbitrariness in all clustering techniques. It is not our intent to provide a review of clustering techniques, since our difficulty with cluster analysis and the means by which we get around it, do not depend on the features of any particular technique.

The chief difficulty for us is not the way the clusters are created, but with the end result itself. Our goal, recall, is to see whether some particular characteristic (say nationality) of a subject can be a predictor of how s/he thinks. Imagine then, that one performs a study on 30 managers, half from Freedonia, and the other half from Slobonia. CMs are elicited, distance ratios calculated, and cluster analysis of one sort or another is performed on the distance ratios. Suppose then that three clusters of roughly equal size are found, and that there appears to be no relationship between cluster membership and nationality. Now imagine that although the nationalities are evenly divided between the clusters, all of the Freedonians are very close to the border between the three clusters and are generally close to each other, thus leading to a type II error of incorrectly failing to reject the null hypothesis. Figure 5 contains an example in two dimensions of this sort of situation.

Similarly, it is not difficult to imagine a situation which would lead to a type I error (incorrectly rejecting the null hypothesis). Here many of the Freedonians would appear in one cluster, but those who do, would be near different extremities of that cluster, while those who are not in the cluster are quite distant from it in different directions. Thus, individuals who are not really close to each other, would appear to be if calculated on cluster membership. Figure 6 shows an example of this sort. In one pilot study (Markóczy and Goldberg, 1992) conducted on 49 executive MBA students at Cranfield University, we found that 4 of the 6 women in the sample were in one of the five clusters. However, the women who were in the cluster were 'late' members of

Figure 5 Type II error with cluster analysis

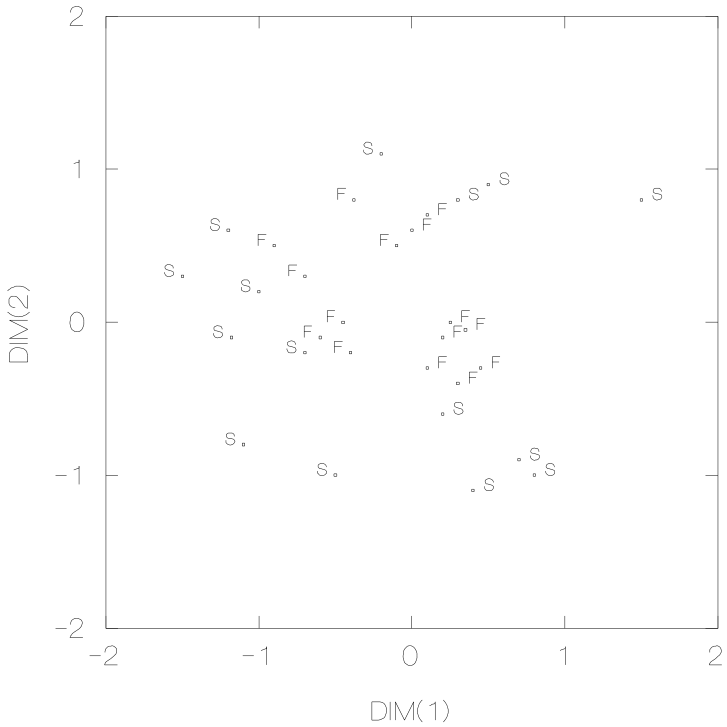
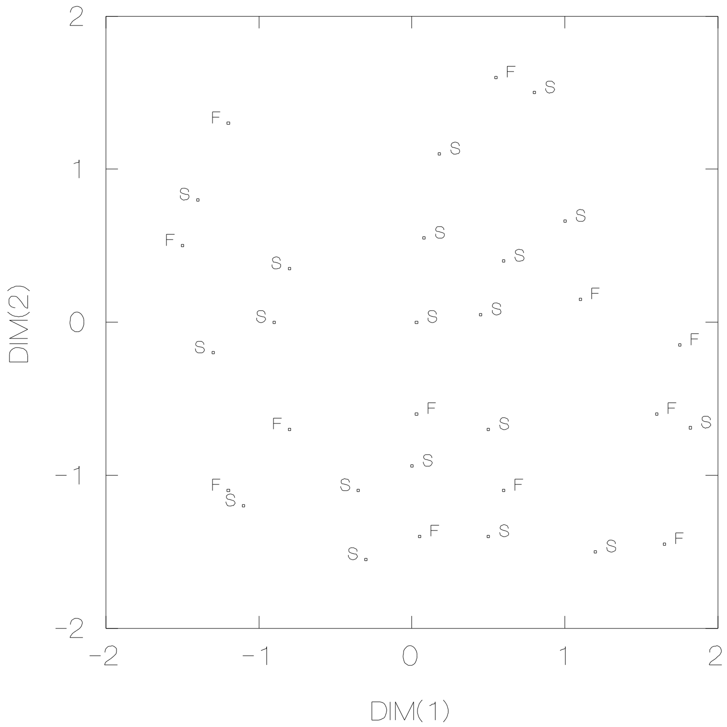


Figure 6 Type I error with cluster analysis



the cluster and so far from its center (and not really near each other). The two women who were not in the cluster were very far from it, and also not near each other. When gender was correlated with distance from cluster centers (see section 6.3), no correlation was found.

MDS can help one notice these errors, but some other modification is needed to identify who is similar to whom within the space defined by the distance measures.

6.3 Fuzzy clusters

The problem with cluster analysis could be avoided if we had some notion of degree to which a particular entity is a member of that cluster. The intuition here is the same as the intuition behind fuzzy set theory (e.g., Bellman and Zadeh, 1970), but it should be noted that we are using an analogy here; in general we reject fuzzy logic and sets because logical operators cannot meaningfully be defined over them (Johnson-Laird, 1983, pp. 198)

With a fuzzy notion of cluster membership we would solve two problems at once. The first is that we would not need to worry about minor variations in the clustering technique since exactly those entities whose cluster membership is 'disputed' would show-up as borderline whether they are included or excluded from the cluster in question. The second problem this would solve is that it would recover information that is lost in the binary decision of something being in a cluster or not.

Our solution is roughly to measure the distance of each map from the 'center' of every cluster. The closer a map is to a cluster center, the more it is 'in' that cluster. For this, we need to establish clusters, identify the center of the clusters, and measure the distance of each map from those centers.

In order to get enough centers which are distant enough from each other to be useful, we need to employ a cluster analysis method which will bias toward small clusters without large distances within them. We use **Ward's method** (Everitt, 1990) for agglomerating clusters since it will tend to give us a number of equally sized coherent clusters.

Once we have such clusters, we create a **central** map for each cluster. The nodes in the central map will be those nodes which appear in more than half of the maps within the cluster. The strengths of the influences between arcs will be the average of the influences from all maps within the cluster that contain both nodes involved. For example if we have a cluster with twelve maps, we first find which nodes appear in at least seven of those maps. Lets say that there are five such nodes (n_8, n_{11}, n_{17}, n_{21}, n_{38}, and n_{45}). To calculate the strength between, say, n_{21} and n_{45} in the central map we need to find out how many of the maps have both of these nodes in them. Let's say that only four of the maps have *both* n_{21} and n_{45}. We sum up the values assigned to those arcs in those four maps and divide the total by four. An example central map is shown later in figure 7 in section 8.7. Once we have the central maps we no longer use the clustering information.

In the next step we calculate the distance ratio between each map in the sample and each central map. Thus, for each map, we will have an indication of how far (or near) it is the 'center' of every cluster. We then correlate characteristics such as nationality not with cluster membership, but with distance from the center of a cluster. Such a correlation table is illustrated in table 2 in section 8.7. This technique will reduce the likelihood of both the type I and type II errors described above.

There are three things to note about this technique. First is that the cluster analysis technique does not have to be very good. Overall results should not differ if somewhat different clusters are identified. Second, the central map does not have to be very close to the center of each cluster, since as long as the central maps of the different clusters are not very close to each other, the technique will work. Third, although the central maps will be different sizes and thus effecting the likelihood of overlapping nodes to occur in comparisons, such effects will modify the relationship between that map and all the maps in the sample evenly so the correlation will come out correctly.

An alternative to creating central maps would be to identify the centroid of each cluster in metrical Euclidean space (created via MDS) and calculate the distance (in the same reduced space) of each map from the centroid. There is no principled reason why one technique should be preferred, but we choose to create central maps instead of identifying centroids because the central maps can be used for other purposes.

6.4 Comparing means of distances

If we wish to know whether some subgroup (let's say Freedonians) are closer to each other than they are to other members of the sample, we should be able to answer that question directly. First we calculate the mean and the standard deviation of all of the distance ratios between Freedonians. In our example of 30 managers, 15 of whom are Freedonian, that would be the average of 105 distances. Let us suppose that that average ($\overline{x}_F$) is 0.65 and σ_F is 0.10. Now we calculate the average distance between Freedonian's and non-Freedonians. (This will be the average of 225 distances. Note that we do not include here the distances between non-Freedonian and non-Freedonian, since the question is whether Freedonians are more like each other than they are like non-Freedonians. The question is not whether Freedonians are more like each other than non-Freedonians are like each other.) Suppose that this average

($\overline{x}_{F\times\tilde{F}}$) is 0.79 and $\sigma_{F\times\tilde{F}}$ is 0.12. We can see then that the average distance between the Freedonians is substantially less than the average distance between Freedonians and others. What we do is compare the average distance within a particular subgroup with the average of the distances between members of that subgroup and outsiders. Section 8.6 contains a 'real' example of such a calculation with the results in table 1.

It is one thing to show the means (and standard deviations) of distances between members of the subgroups, and quite another to determine whether the differences in the means are significant. McKeithen, Reitman, Rueter and Hirtle (1981), in a study on how novice and expert computer programmers organize information, calculated distances between individuals and presented intrasubgroup averages; but they did not attempt to show that the differences in the means were significant. For non-distance based data there are a number of ways of testing the significance of differences of means. The most straight forward is the Student t-test. The t-test does require that the distribution be approximately normal or that the subgroups be sufficiently large. It is therefore recommended to confirm the assumption of normality of the distribution of the distances ratios using standard tools. Unless there are good reasons to support normality, we recommend the use of non-parametric comparisons of means. The Mann-Whitney or Wilcoxon-Mann-Whitney (Siegel, 1956; Sandy, 1989) tests are well suited to the task and should be used.

We only have a 'result' if the members of the subgroup are close to each other. Thus, we are only interested in those intra-subgroup averages which are less then the remaining average. It may be possible to draw conclusions about the opposite occurrence (when members of a subgroup are further then expected from each other), but such effects can come about for a number of reasons and other tests would need to be applied before being able to conclude that members of that subgroup are somehow 'diverse'.

Both the t-test and the Mann-Whitney test require that the individual observations which are being compared (in our case the distance ratios) be independent of each other. Our distance ratios, like all distance data, are not independent of each other. We have found no principled way of modifying these tests to provide meaningful significance levels. Although it is not possible to assign a precise interpretation to the t (or Z for Wilcoxon-Mann-Whitney) value, it is possible to compare them to each other. So if some subgroup, Freedonians lets say, yields a Z of 2.9, and another subgroup, left-handed people for example, yields a Z of 0.3, we can safely say that nationality plays a stronger role than handedness, even if we cannot calculate exact and meaningful probabilities for Z.

If studies are repeated across a number of organizations, then a **rank cor-**

relation (Siegel, 1956) of this relative significance can be used to determine whether there is an overall pattern. Similarly, if enough individuals are used from one organization, it would be possible to take random subsets (of a suitable size) from the full sample, calculate t or Z statistics for the characteristics within each subset of individuals and then perform the identical rank correlation of relative significance between subsets of single sample. This would help mitigate the otherwise large effect that outliers have on this sort of calculation.

6.5 Continuous characteristics

The above techniques have been used to examine whether some particular characteristic such as nationality relates to whether individuals group together. We may also wish to investigate whether some characteristic such as age or rank or years within the organization plays a role. We can always put individuals into various bins or categories based on such numerical data and then treat these characteristics as we do the others. But such binning of data involves an unnecessary loss of information, and we wish to use as much information as is available for our tests whenever possible. We therefore investigate the relationship between the distance ratio and the distance in these characteristics. For example we check the influence of age by correlating for each pair of subjects the distance between their maps that the absolute difference in their ages. A positive correlation indicates the further people are in age, the further apart are their measured belief structures. A negative correlation would tell us nothing interpretable, and so, again, we just conduct a one-tailed test.

As with the comparison of means, there is no principled way to calculate absolute significance. But, again we can calculate relative significance to determine which effects are strongest. Continuous data may also be correlated with the distance from cluster centers (section 6.2) without any binning.

7 Non-distance based analysis

There are a number of forms of analysis which can be conducting on the causal maps which do not depend on the distance ratios. Several excellent tools and methods have been developed by others for exactly these purposes (e.g., Ackermann, Cropper and Eden, 1992; Laukkanen, 1992). We will not discuss those here, but instead focus on forms of analysis which have not to our knowledge been applied to causal maps. Many of these methods will not require that the maps be elicited as specified above, but will be more broadly applicable.

7.1 Factoring constructs

One form of analysis would be to investigate patterns in the choice of constructs. Did people who selected one construct, n_1 tend to select n_j as well? This sort of problem is routinely addressed by both factor and cluster analysis (Kim and Mueller, 1978; Füstös and Kovács, 1989). However, we again wish to use all the available information, so instead of correlating only whether the a construct was selected or not, we can correlate on how 'influential' a construct was taken to be. We sum up the strengths of the arcs leading from each construct. This degree of influence for each construct is calculated for each map. These numbers not only indicate whether a construct was selected, but how important it is taken to be within the system. We calculate the correlation matrix on these influences and continue with factor and cluster analysis from there.

One caveat with this technique is a consequence of forcing subjects to select a small number of constructs. Despite our efforts, subjects may tend to select constructs which don't overlap in meaning. For example, subjects who select 'annual profit' may be less likely to select 'long term profitability'. Our elicitation technique is designed to minimize this problem, but we cannot eliminate it. Note, that this is much less of a problem for maps elicited without the restrictions we require, so this form of analysis should be more productive with maps elicited in a freer format.

7.2 Complementarity

We can turn adversity into a virtue by examining what pairs of constructs were not selected by the same individual. That is, we look for pairs of constructs which are in **complementary distribution** (Harms, 1968). If no individual selected both n_i and n_j and there is an overlap in meaning between the two, we can explore whether there is a pattern to who selected which of the two. For example, we can ask whether only Freedonians selected n_i while Slobonians selected n_j. If enough individuals selected either one or the other, we can perform the Fisher Exact Test (Siegel, 1956) on exactly this sort of data. Section 8.8 contains an example of such a calculation. Note that given the power of contemporary computers, there is no reason to calculate a χ^2 approximation for a two by two contingency when Fisher's Exact Test takes a few milliseconds longer.

7.3 Non-quantitative analysis

Various forms of non-quantitative analysis can be performed. Central maps can be constructed for any group of individuals (not just for those in clusters) and their properties explored and compared. Maps can be inspected by eye to see if recurrent themes emerge. If such patterns are discovered, it would always be possible to design a test to confirm (or refute) such hypotheses

7.4 Filling gaps

In the last few pages we have gone over a number of forms of analyzing CMs, both in terms of their distances form each other and without reference to distance measures. Some techniques we have discussed in great enough detail for someone to blindly follow the instructions. Other forms of analysis we have barely alluded to. We hope, though, that we have given enough hints to enable others to more fully develop those techniques. We have attempted in these sections to help fill a gap pointed out by Ginsberg (1992, p. 147) in discussion of quantitative tools for CM analysis.

8 Example

Markóczy (1994) reports on the study this method was originally designed for. That study involved the investigation of more than 20 managers in five organizations in Hungary for a total of 111 managers. Markóczy (forthcoming) briefly describes (but in greater detail than Markóczy (1994)) the results from one of the five organizations studied. Markóczy and Goldberg (1992) describes a pilot study conducted with a sample of 49 executive MBA students at Cranfield University. We are also aware of others who are using some parts of the method, particular the comparison technique described in section 5, for work unrelated to our own (John Sawyer, personal communication).

In this section we discuss an example from another one of the organizations in Markóczy's (1994) sample, although some examples will be brought in from the works mentioned above. Not every techinque for analysing the distance data will be exemplified, but the two most novel ones will be.

8.1 The investigated organization

The organization, XYZ, is a food industry company which specializes in biscuit, confectionery, and wafer production. The company was acquired by a British company ABC in 1991 when the foreign partner bought 84% of its shares. In 1993 ABC managed to increase its ownership to 96% by buying the shares from the local authorities. ABC bought the company in April 1991 but it did not initiate changes until November 1992 at which time ABC sent five foreign managers and hired some western trained young Hungarian managers who were made responsible for introducing changes. Since most of these people had worked in XYZ only since November 1992, the changes they initiated were very recent and most of them were still in progress at the time of the study (May–June 1993).

At the time of the study the number of employees in the company XYZ was 1500 and has not changed substantially since then. The company was profitable one and made around 470 million HUF in 1992. Most of the products which XYZ produces are sold in the Hungarian market with only 6–7% of the return on sales originating from export (down from 8–10% at the time ABC acquired the company).

8.2 The pool of constructs

Following the guidelines in section 4.1, the pool of constructs was compiled prior to the study. The forty-nine constructs have their basis in an informal analysis of interviews with a sample of managers similar to the sample in the present study (Child, Markóczy and Cheung, 1994), items suggest by the literature, items (or revisions of items) suggested by participants in a pilot study (Markóczy and Goldberg, 1992), and suggestions of colleagues. Once the list, and their definitions were constructed, they were translated into Hungarian by a native speaker of Hungarian, retranslated back to English by a professional translator, and then compared with the original to identify ambiguities. The constructs, without the explanatory definitions, are listed in appendix A.

8.3 The sample of managers

Twenty-two managers from this company were included in the sample. This included all of the five non-Hungarians (who are referred to as 'the foreigners' below). The other 17 were Hungarian. The managers were spread over a number of functional areas, and all of the top managerial team were included.

8.4 Elicitation of the maps

The causal maps were elicited as described in sections 4.2 and 4.3. In addition a half hour unstructured interview was conducted with each subject, after the maps were elicited, so that suggestions that may have come up in the interview would not prime the subjects to be particular sensitive to some issues during map elicitation. The response time of managers varied, but usually the 10 node maps could be elicited in slightly less than one hour.

8.5 Calculating distance

The formula in figure 4 was used to calculate the 231 distance ratios between all pairs of the 22 maps. Given how the maps were elicited (three strengths, two polarities, no consideration of a construct's influence on itself, and specific questioning of all pairs), the appropriate values for the parameters were $\alpha = 1$, $\beta = 3$, $\gamma = 2$, and $\epsilon = 2$. δ was set equal to zero on a more arbitrary basis. Section 5.2 provides a description of what these parameters mean.

8.6 Means of distances

Following the analysis technique discribed in 6.4, the means of the intra and inter subgroup distances were calculated. For example, in the subgroup of the foreign managers, the average of those 10 distance ratios was 0.5142 with a standard deviation of 0.136. The average of the 85 ($5 \times (22-5)$) for the distances between each of the five foreigners and each of the 17 Hungarians) distances between foreigner–Hungarian pairs was 0.7338 with a standard deviation of 0.144. (The averages between members of a group and the members of the complement group are not shown in the table; in practice those averages are not very different from the all sample average.) Thus it appears that the foreigners are more likely to have similar response to each other than to the Hungarians. Although a t-test and a Wilcoxon-Mann-Whitney test can be calculated for these, it must be noted again (as in section 6.4) that *the probability values do not represent true probabilites and are underestimates of the probability of these results occuring by chance.* However, the p values, although not very meaningful in themselves, can be ranked. These p values are shown for various subgroups in the sample in table 1. The groups are listed in the table in order of increasing p values as calculated by the Wilcoxon-Mann-Whitney test.

It is important to note that just because one subgroup (the foreigners in this example) are relatively similar to each other there is no reason to expect that the complement group (the Hungarians) will also be similar to each other. As we can see in table 1, the avereage distance between 136 pairs of 17 Hungarians was 0.7633 (standard deviation 0.126). This average is greater than both the average distance between everyone in the sample, and the average distance between foreigner and Hungarian. That is, for the foreigners to somehow share beliefs in this organization, it is not necessary that the Hungarians do as well.

The alternative is also possible. Although the numbers are less dramatic, we see in table 1 that the male managers are closer to each other than they are two the female managers, and at the same time the female managers are slightly closer to each other than they are to the male managers.

8.7 Clusters and correlations

Following the steps outlined in section 6.3 we run a cluster analysis on the 231 distance ratios. Because we are not so interested in cluster membership itself, we have a fair amount of freedom in chosing a cluster cut-off. Such choices always carry a degree arbitariness, and our goal was to identify small clusters. Given the normal criteria (Everitt, 1990), plus our goal, a cut-off at four clusters was chosen. The clusters, A, B, C, and D, have 7, 6, 3, and 6 members respectively.

Once the clusters were found, the central map for each cluster was created.

Subgroup	M	N	$\overline{x}$	σ	t	Z_{MW}
All	22	231	0.7417	0.1418		
NA-Frgn	5	10	0.5142	0.1363	4.5883	3.7411
DFA-MA	6	15	0.5733	0.1402	4.1138	3.5409
Lang-Multi	10	45	0.6625	0.1340	3.8561	3.5161
FA-MA	5	10	0.5922	0.1302	2.5667	2.4981
SEX-Male	17	136	0.7948	0.1277	1.4395	1.0615
ED-Other	16	120	0.7315	0.1458	0.9352	0.8983
SEX-Fem	5	106	0.8012	0.0966	0.4413	0.7215
Pos-Top	6	15	0.6994	0.1264	0.6936	0.3838
Lang-Mono	12	66	0.7697	0.1331	−0.6509	−0.7230
DFA-PE	10	45	0.7823	0.1296	−0.7600	−0.9165
ED-Tech	6	15	0.7716	0.1330	−0.5692	−0.9359
NA-Hung	17	136	0.7633	0.1262	−1.6017	−1.1807
Pos-NonTop	16	120	0.7590	0.1416	−1.6666	−1.8623
FA-PE	6	15	0.8306	0.1155	−2.3424	−2.2341

M is the number of Maps, N is the number of distances between the maps. $\overline{x}$ is the mean distance between maps and σ is the estimated standard deviation. t is the result of the Student t-test and Z_{MW} is the Z statistic for the variant of the Wilcoxon-Mann-Whitney test.
Only subgroups which have at least four members are shown here. In the subgroup names NA is Nationality (`NA-Hung` is Hungarian, `NA-Frgn` is foreign), `ED` is Education type (`ED-Tech` is technical education, `ED-Other` is non-technical education), `FA` indicates current Functional Area (`FA-MA` is marketing and sales, `FA-PE` is production), `DFA` is Dominant Fuctional Area (for subdivision see functional area), `Lang-multi` is multilingual, `Lang-mono` is monolingual, and `POS-Top` is top managerial position, `POS-NonTop` is non-top managerial position. `SEX-Male` for male or `SEX-Fem` for female.

Table 1: Inter-subgroup averages for XYZ

Figure 7 Central map for cluster A

	3	5	6	21	33	36	41	47	49
3	0.00	3.00	3.00	2.50	1.50	2.60	1.83	2.33	2.25
5	2.33	0.00	2.50	1.00	0.00	1.00	0.50	2.50	1.00
6	1.00	1.50	0.00	1.67	1.00	2.00	1.25	2.67	2.50
21	3.00	3.00	3.00	0.00	2.75	2.75	2.00	2.67	2.50
33	2.00	3.00	2.33	2.00	0.00	2.33	0.50	2.00	2.00
36	1.20	2.25	3.00	1.25	2.00	0.00	0.83	0.33	1.00
41	1.00	0.00	2.75	0.40	0.50	0.50	0.00	2.75	1.50
47	2.00	1.00	3.00	1.67	2.00	0.33	3.00	0.00	1.00
49	1.25	0.00	2.50	1.50	2.00	1.00	2.50	2.00	0.00

3	vision/strategy	**36**	quality of products
5	market share	**41**	cost control
6	long term profit	**47**	efficiency/productivity
21	management flexibility	**49**	relation with suppliers
33	understanding the market		

The first step is to indentify which nodes are in more than half the maps in a cluster. For cluster A that meant finding which nodes were in at least four of the seven maps in the cluster. It turns out that there are nine such nodes, two of them are node 5 (market share) and node 6 (long term profitability). Each of these nodes are in four out of the seven maps, so the just qualify for the central map. However, both of these are only together in two of the maps. So for calculating the value of the of the influence of node 5 (n_5) on node 6 (n_6) (and n_6 on n_5) in the central map, we only consider the values in those two maps. In the one the maps the relavent arcs are $n_5 \xrightarrow{+3} n_6$ and $n_6 \xrightarrow{+2} n_5$. In the other map which has both n_5 and n_6 the arcs are $n_5 \xrightarrow{+2} n_6$ and $n_6 \xrightarrow{+1} n_5$. Thus in the central map for cluster A the arcs $n_5 \xrightarrow{+2.5} n_6$ and $n_6 \xrightarrow{1.5} n_5$ will appear. Figure 7 lists the central map for cluster A.

Once all of the central maps are created in this fashion, a set of distance ratios between all of the maps from the sample with all of the cluster central maps can calculated. In this example, that involves the calculation of 88 distance ratios since there are 22 maps in the sample and 4 central maps.

Table 2 contains the correlations between various characteristics of the managers in the sample, the the distances of their elicited maps from the various central maps. In addition to all of the characteristics tested in table 1 it was possible to add characteristics such as age in years (`Age`) and year of employment at the company (At_comp).

Note because the distance ratio increases with dissimilarity, it is generally the negative correlations we are interested in. So, for example the negative r of $-.54$ ($p \approx .009$) for distance from cluster D for the individuals whose dominant functional area is marketing (DFA_MA) indicates that those individual's maps are close to the central map cluster D. When we see a positive r of .611 ($p \approx .003$) with between being Hungarian (NA_Hung) and distance from the central map of cluster D, we learn nothing about the Hungarians, but we can infer that the *non*-Hungarians are very close to the center of that cluster. If we can conclude that the foreign managers' maps are similar to the central map of a cluster, than we can conclude that they are similar to each other as well. However, if we can conclude that the Hungarian managers' maps are far from a particular cluster, we can neither conclude that those maps are far from each other or near each other.

With Age we find a negative correlation with distance from cluster C ($r = -.525, p \approx .012$). What this indicates is that the older people are the nearer they are to the central map of C. That is, in increase in age correlates with a decrease in distance from the cluster.

Given that the results from the analysis method described in section 6.4 and exemplified in section 8.6 does not provide true probability levels, and that the method exemplified in this section and outlined in section 6.3 involves a great number of intermediate calculations, it is advised that only those tested characteristics which show positive results under both tests be considered as results.

8.8 Node selection

Following the guidelines in section 7.2 we can see whether there are pairs of constructs with some meaning overlap, which are not often both selected by subjects. A pair of seach constructs are 'long-term profitability' and 'yearly profit'. In this sample 13 managers selected exactly one of these two items. The contingency table showing the relationship between the Hungarian and foreign managers in their selection of these cards (table 3) shows no statitical relation ($p = .437$ by Fisher's exact test). It should be noted that in another organization in the study a strong relation was found with the foreigners more concerned with short term profitability and the Hungarians with long term profitability (Markóczy, forthcoming).

FA_PE	-0.33															
DFA_MA	0.89***	-0.15														
DFA_PE	-0.50*	0.47*	-0.56**													
ED_Tech	-0.33	0.31	-0.37†	0.47*												
NA_Hung	-0.48*	-0.15	-0.64***	0.28	0.33											
Age	-0.54**	0.17	-0.62***	0.64***	0.59***	0.63***										
At_Comp	-0.61***	0.00	-0.69***	0.60***	0.54**	0.61***	0.86***									
FuncDiv	-0.15	0.37†	-0.15	0.20	0.05	-0.19	0.12	-0.00								
Topman	-0.09	0.08	0.08	-0.15	-0.15	-0.15	-0.09	-0.31	0.11							
Multling	0.59***	-0.15	0.67***	-0.47*	-0.56**	-0.59***	-0.70***	-0.75***	-0.14	0.47*						
Gender	-0.18	-0.25	-0.25	-0.12	0.17	0.41†	-0.02	0.13	-0.08	-0.04	-0.12					
Dist_A	-0.13	0.12	0.02	-0.11	0.20	-0.00	0.06	0.08	0.11	-0.18	-0.26	-0.01				
Dist_B	-0.33	-0.03	-0.32	0.20	0.15	-0.11	0.25	0.54**	0.06	-0.33	-0.27	0.09	0.04			
Dist_C	0.20	-0.25	0.29	-0.42**	-0.25	-0.30	-0.53*	-0.52*	0.05	0.28	0.35	0.05	-0.01	-0.40†		
Dist_D	-0.36†	0.22	-0.54**	0.57**	0.10	0.61***	0.45*	0.37†	0.07	-0.18	-0.36†	0.07	-0.08	-0.15	-0.61***	
Mean	0.23	0.27	0.27	0.45	0.27	0.77	42.09	15.32	0.80	0.27	0.45	0.36	0.62	0.80	0.77	0.67
σ	0.43	0.46	0.46	0.51	0.46	0.43	7.63	12.78	0.16	0.46	0.51	0.49	0.17	0.19	0.19	0.24
Vars.	FA_MA	FA_PE	DFA_MA	DFA_PE	ED_Tech	NA_Hung	Age	At_Comp	FuncDiv	Topman	Multlin	Gender	Dist_A	Dist_B	Dist_C	Dist_D

$N = 22$, $^{\dagger}p < 0.1$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .005$

Variables are as in table 1, with the addition of `Age` is age in years, `At_Comp` is the time spent at the company, `Gender` is gender (which here marks females) `FuncDiv` is functional diversity, and `Dist_A`–`Dist_D` are the distances from the central maps in cluster A, B, C, and D. In the table those managerial characteristics which are complementary to each other (Hungarian vs. foreign, technical education vs. non-technical education, top vs. non-top, male and female) are only listed under one of their possibilities.

Table 2: Cross Correlations for company XYZ

	Foreign	**Hungarian**	Total
Sort-term profit	2	4	6
Long-term profit	1	6	7
Total	3	10	13

$p \approx .437$

Table 3: Contingency table for nationality and long vs. short term profit selection

9 Conclusions

Causal mapping has long been valued as a tool for exploring individuals' idiosyncratic beliefs. In this paper we have done three things to make these benefits more utilizable by researchers. First, by detailing a replicable elicitation technique, we have laid the ground for systematic and objective comparison of these beliefs: Systematic analysis and comparison require systematic elicitation. Secondly, through a careful analysis of what a causal map means, we have been able to refine a way of comparing maps which uses all of the available information within the maps and is objective in the sense that all subjective aspects of the method are made explicit (e.g., the pool of constructs, the interpretation of missing information, etc). Thirdly, we have discussed and extended various tools for analyzing distance based measures. Each one of these, we feel, is a contribution to the set of tools researchers will use. Researchers may use or further develop any one these three independently of the others, but each was worked out to work with the other two, so that in combination they form a powerful package.

9.1 Computational tools

Several sets of computational tools already exist for the study of causal maps such as CMAP2 (Laukkanen, 1992) and COPE (Eden, Ackermann and Cropper, 1992). In a move counter to the trend toward large integrated interactive systems, Goldberg has been developing a set of tools each one of which performs one of the tasks described above with little duplication of what already exists. Although the programs in the package are independent programs, they are designed so that the output of one can be used as the input for another. To date, only very preliminary versions are available to the academic community in portable C source code and DOS executables for programs to calculate distance measures between maps discussed in section 5.2, to elicit the causal maps (section 4.3), and to create central maps (section 6.3), and other minor tools. These are available from the second author, and also appear on the management archive (`gopher::ursus.jun.alaska.edu`).

References

Ackermann, Fran, Steve Cropper and Colin Eden (1992). 'Graphics COPE users guide', *Report*, Department of Management Science, University of Strathclyde, Glasgow.

Axelrod, Robert (ed.) (1976). ***Structure of decision: The cognitive maps of political elites***, Princeton University Press, Princeton, NJ.

Barr, Pamela S., J. L. Stimpert and Anne S. Huff (1992). 'Cognitive change, strategic action, and organizational renewal', ***Strategic Management Journal*** **13**: 15–36.

Bellman, R. E. and L. A. Zadeh (1970). 'Decision making in a fuzzy environment', ***Management Science***.

Bougon, Michel G. and John M. Komocar (1990). 'Directing strategic change: A dynamic wholistic approach', in Anne S. Huff (ed.), ***Mapping strategic thought***, Wiley, New York, pp. 135–163.

Bougon, Michel G., Karl E. Weick and Din Binkhorst (1977). 'Cognition in organizations: An analysis of the Utrecht jazz orchestra', ***Academy of Management Review*** **22**: 606–639.

Breiger, Ronald L., Scott A. Boorman and Phipps Arabie (1975). 'An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling', ***Journal of Mathematical Psychology*** **12**: 328–383.

Child, John, Lívia Markóczy and Tony Cheung (1994). 'Managerial adaptation in Chinese and Hungarian strategic alliances with culturally distinct partners', in Sally Stewart (ed.), ***Joint ventures in the People's Republic of China***, Vol. 4 of *Advances in Chinese industrial studies*, Jai, Greenwich, CT. (forthcoming).

de Leeuw, Jan and William Heiser (1982). 'Theory of multidimensional scaling', in P. R. Krishnaiah and L. N. Kanal (eds), ***Classification, pattern recognition and reduction of dimensionality***, Vol. 2 of *Handbook of statistics*, North-Holland, Amsterdam, pp. 285–316.

Eden, Colin (1992). 'On the nature of cognitive maps', ***Journal of Management Studies***.

EDEN, COLIN, FRAN ACKERMANN and STEVE CROPPER (1992). 'The analysis of cause maps', ***Journal of Management Studies* 29**(3): 309–324.

EVERITT, BRIAN (1990). ***Cluster Analysis (2nd Edition)***, Halsted Press, Wiley, New York.

FORD, JEFFREY D. and W. HARVEY HEGARTY (1984). 'Decision makers' beliefs about the causes and effects of structure: an exploratory study', ***Academy of Management Journal* 27**: 271–291.

FÜSTÖS, LÁSZLÓ and ERZSÉBET KOVÁCS (1989). ***A számítógépes adatelemzés statisztikai módszerei (Computational methods for statistical data analysis)***, Tankönyvkiadó, Budapest.

GINSBERG, ARI (1992). 'Review of a. s. huff 1990 "mapping strategic thought"', ***Academy of Management Review* 17**: 145–148.

HALL, ROGER I. (1984). 'The natural logic of management policy making: Its implications for the survival of an organization', ***Management Science* 30**(8): 905–927.

HARMS, R. T. (1968). ***Introduction to phonological theory***, Prentice-Hall, Englewood Cliffs, NJ.

HUFF, ANNE S. (ed.) (1990). ***Mapping strategic thought***, John Wiley & Sons, New York.

JOHNSON-LAIRD, P. N. (1983). ***Mental models***, Cambridge University Press, Cambridge, England.

KIM, JAE-ON and CHARLES W. MUELLER (1978). ***Factor analysis: Statistical methods and practical issues***, Vol. 14 of *Sage university paper series on quantivative applications in the social sciences*, Sage, London.

LANGFIELD-SMITH, KIM M. and ANDREW WIRTH (1992). 'Measuring differences between cognitive maps', ***Journal of the Operational Research Society* 43**(12): 1135–1150.

LAUKKANEN, MAURI (1992). ***Comparative cause mapping of management cognitions: A computer database method for natural data***, Helsinki School of Economics and Business Administration Publications, Helsinki.

MARKÓCZY, LÍVIA (1994). 'Barriers to shared belief: The role of managerial characteristics, environmental factors, and strategic interest', Meeting of the Academy of Management, Dallas.

MARKÓCZY, LÍVIA (forthcoming). 'States and belief states', ***International Journal of Human Resource Management***.

MARKÓCZY, LÍVIA and JEFF GOLDBERG (1992). 'On identifying managerial types based of beliefs about business', *Research paper 17/92*, Judge Institute of Management Studies, University of Cambridge, Cambridge, England.

MCKEITHEN, KATHRINE B., JUDITH S. REITMAN, HENRY H. RUETER and STEPHEN C. HIRTLE (1981). 'Knowledge organization and skill differences in computer programmers', ***Cognitive Psychology* 13**: 307–325.

ROBERTS, FRED S. (1976). 'The questionnaire method', in Robert Axelrod (ed.), ***Structure of Decision. The Cognitive Maps of Political Elites***, Princeton University Press, Princeton, NJ.

ROSS, ROBERT T. (1934). 'Optimum orders for the presentation of pairs in the method of paired comparison', ***Journal of Educational Psychology* 25**: 375–382.

SANDY, ROBERT (1989). ***Statistics for business and economics***, McGraw-Hill, New York.

SIEGEL, SIDNEY (1956). ***Nonparametric statistics: for the behavioral sciences***, Series in Psychology, McGraw-Hill, New York.

THULASIRAMAN, K. and M. N. S. SWAMY (1992). ***Graphs: Theory and algorithms***, John Wiley & Sons, New York.

TVERSKY, AMOS (1977). 'Features of similarity', ***Psychology Review* 84**: 327–352.

WISH, MYRON and J. DOUGLAS CARROLL (1982). 'Multidiminsional scaling and its applications', in P. R. Krishnaiah and L. N. Kanal (eds), ***Classification, pattern recognition and reduction of dimensionality***, Vol. 2 of *Handbook of statistics*, North-Holland, Amsterdam, pp. 317–345.

A List of constructs

1 Quality and implementability of hiring and firing policy
2 Satisfied employees
3 Vision and strategic direction of the company's leadership
4 Host country national culture (values, society's attitude towards business)
5 Market share of the company
6 Long term profitability of the company
7 Parent company's support
8 Inflation of host country currency
9 Access to external R&D
10 Lobbying capabilities with the government
11 Satisfying shareholders or owners
12 Employees attitude
13 Debt/Equity ratio
14 The quality of the production technology
15 Growth of the company
16 Responsibility to the community
17 Job security for the employees of the company
18 Investment intensity of the company
19 The degree of unionization of the employees within the company
20 Competition in the market
21 Management flexibility
22 Prices applied by the company
23 Internal R&D capability/Product development
24 Shared organizational culture
25 Economy of scale of the production
26 Size (number of employees)
27 Customer relations
28 Economic and political conditions (excluding inflation and currency exchange rate)
29 Legal expertise of the company
30 Leadership within the organization
31 Yearly profit of the company (net income after tax)
32 Size of the company's market
33 Knowledge of needs of the company's market
34 Returns on sales
35 Brand recognition of the company's products
36 Quality of products and services
37 Survival of the company

38 Management's understanding the mentality of the host country managers
39 Purpose of the parents in establishing the company
40 Relative valuation of host country currency
41 Cost control within the company
42 Incentives given to employees (financial and fringe benefits)
43 Bank connections
44 Familiarity with environmental conditions
45 Individual contribution of employees
46 In house training
47 Efficiency/productivity
48 Quality of the distribution channels
49 Relationships with suppliers