Measuring the Effect of Systems Thinking Interventions on Mental Models

James K. Doyle, Michael J. Radzicki, and W. Scott Trees
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609-2280 USA
email: doyle@wpi.edu

Introduction

The concept of mental models has been central to the practice of system dynamics since its inception. When system dynamicists intervene in educational or corporate settings to help people become better systems thinkers, they typically begin by "surfacing" the participants' mental models to make assumptions explicit and to assist the model building process. And, changing mental models to make them more complete, consistent, and dynamic is typically one of the primary goals of the intervention.

It is the purpose of this paper to describe the limitations of currently available techniques for eliciting, changing, and measuring changes in mental models of complex systems. In particular, when designing its techniques for studying mental models, the system dynamics community has not taken into account recent developments in the fields of cognitive psychology and the psychology of judgment and decision making and has also ignored certain important standard procedures for collecting unbiased data from human subjects. In addition, despite the compelling need for rigorous scientific evaluation of the effectiveness of efforts to improve systems thinking (Cavaleri and Sterman, 1995; Ganter, Doyle, and Radzicki, 1995), current interventions typically fail to include rigorous follow-up measures of the changes in mental models they claim to bring about.

The paper also describes a new methodology for measuring the effect of a systems thinking intervention on the mental models of participants, and presents preliminary results from an intervention based on the simulation game of the economic long wave, or Kondratiev Cycle, developed by Sterman and Meadows (1985).

Limitations of Current Methods for Studying Mental Models of Systems

The following represent serious limitations to the ability of current methods to accurately measure mental models of systems:

1. Current facilitation methods for studying mental models are not primarily designed to measure mental models but to *improve* them. As a result, the mental models of participants are often altered before or during measurement. As soon as a facilitator becomes involved in the measurement process, as soon as one member of the group begins to share his or her ideas and assumptions, the mental models of the group members may begin to change. Beginning with a process that immediately starts to change clients' mental models does not allow the effectiveness of the intervention to be determined, since no pre-intervention data are collected.

2. In a typical intervention, great care is taken to elicit clients' mental models at the beginning of the process. However, comparatively little effort is made later on to gather evidence that mental models have changed in desirable ways. When mental models are measured after an intervention, they are often measured using a completely different procedure than was employed at the beginning of the process. If this is done, *it is not possible* to unambiguously determine if any differences observed are due to changes in mental models or to the way the mental models were measured. For purposes of assessment, mental models must be measured using the exact same procedures and instruments at the beginning and end of an intervention. It is also important that mental models actually be elicited during the post-test rather than simply

asking clients whether or not they have perceived a change in their mental models as, for example, was done by Cavaleri and Sterman (1995). Because clients know the expectations of the intervener, there is a high probability of what psychologists call "subject bias"; that is, clients may report the changes the intervener wants to hear rather than the true state of affairs.

3. Many researchers assume that the improved, more dynamic mental models facilitated by the systems thinking interventions are easily accepted and stable. In fact, cognitive psychological studies have shown that when a new mental model is learned, the older, inferior mental model does not typically disappear, but remains in memory to compete with the new model. And, when it comes time to make a decision, the older model often gets called up because it has been used more often and has formed more connections with other information held in long-term memory. In addition, unless a newly learned mental model is applied frequently, its details can be quickly forgotten. As a result of these features of mental models, assessment of mental models must be conducted periodically to ensure that mental models have not changed over time in undesirable ways. This is rarely if ever done in current practice.

4. There are serious limitations to eliciting mental models in a group setting. First, it cannot simply be assumed, as it often is, that each individual member of a facilitation group that reaches a "shared consensus" has actually adopted the shared group model -- the extent to which models are shared must be measured by eliciting and comparing the mental models of individuals in isolation, when pressures to conform to the group are reduced. Second, by eliciting mental models in a group setting, valuable information is lost. It is a well-known social psychological result that "brainstorming" groups generate fewer ideas and lower quality ideas than the same number of individuals working alone. This is a result of social loafing, evaluation anxiety, and people simply being interrupted and losing their train of thought.

5. The active role taken by the facilitator during measurement of mental models is also problematic. Since the facilitator takes responsibility for summarizing the ideas of participants and controls the direction of the discussion, there is a high probability of what experimental psychologists refer to as "experimenter bias"; that is, the facilitator may inadvertently give the participants clues about what ideas are better than others or lead the discussion in a direction that the participants would not choose on their own. In addition, with current methods the facilitator decides when to end discussion on a topic. If the facilitator stops requesting input too soon, important information in clients' mental models will be left out of its surface representation; if the facilitator goes-on requesting input too long, clients may be forced to go beyond their knowledge and make wild guesses, resulting in a surface representation that is full of unreliable detail.

6. The task given to clients during facilitation sessions is often ill-defined and quite different from the way people naturally go about making decisions. This is a critical point because the most current thinking in the fields of cognitive psychology and judgment and decision making is that what people remember and think about can be highly dependent on subtle characteristics of the situation they are in at the time and even seemingly inconsequential differences in how questions are worded. Changes in what questions clients are asked, and even the order in which they are asked, can potentially lead them to call up entirely different mental models from memory. This means that, in order to ensure that the elicited mental models are identical to those that are used to make decisions and solve problems, the models must be measured in the context of a well-defined decision task that is as similar as possible to how clients actually make business decisions. If this is not done it is possible to accidentally measure transient mental models that were created on the spot during an intervention rather than the more durable models that people have used to make decisions in the past and will use again in the future.

7. In many systems thinking interventions, clients are taught a new way of thinking about their mental models, such as the hexagon method or causal-loop diagramming. However,

introducing a new method for people to describe their thinking may alter the content and structure of the surface representation of clients' mental models; may introduce errors because people do not fully understand the new procedure; and may yield results that do not generalize when people later revert back to their old ways of communicating their ideas. To ensure that mental models are accurately represented, unless the group is already expert in a new technique, mental models should be elicited in a manner that is as close to the way people typically communicate their ideas as possible, i.e., through verbal narrative explanations.

These limitations are sufficiently serious that the mental models elicited by current systems thinking interventions may be substantially different from the mental models that clients actually use to make business decisions.

A New Methodology

With the above limitations in mind, we have developed and applied a new method for measuring the ability of systems thinking interventions to change mental models in desirable ways. Prior to the intervention, a pretest survey is administered to each individual participant which presents reference mode data and gives them a specific question to answer: What caused the pattern in the data? The survey guides participants through the process of "telling the story" behind the pattern in the data and explaining the important causal events, factors, and variables and the relationships between them. The participants decide how much or how little information to include in their responses, and they are asked to express the degree of confidence they have in their explanations. Following the systems thinking intervention, which may incorporate a variety of facilitation, mental model elicitation, group, or simulation techniques, individual mental models are assessed again using the exact same instrument and procedure employed during the pretest. The advantages of this method include separating the processes of measuring and changing mental models; minimizing the potential for experimenter and subject bias; and increasing the likelihood that the measured models are those that are used in real-life decision making.

To avoid errors associated with asking participants to express their ideas in a new and unfamiliar way, participants' narratives are coded into the format of a causal loop diagram by experimenters blind to experimental condition using established psychological techniques for identifying explicit and implicit structures in narrative text. By identifying the number of subjects who include a variable and counting the number of variables, connections between variables, and feedback relationships, pre-post differences in the content, structure, complexity, and dynamics of mental models can be quantified. The extent to which mental models are shared by individuals may also be quantified, allowing assessment of the effectiveness of the intervention in building group consensus.

This new methodology was applied in the context of an intervention designed to teach college undergraduates about the causes of the economic long wave. Half of the students were given a pre-test in which they were shown data of how de-trended GNP in the U. S. has fluctuated from 1800 to the present and were asked to explain what caused the pattern in the data. Students then played the simulation game of the long wave, following the procedures and debriefing strategies developed by Sterman and Meadows (1985) and Sterman (1989). Finally, all of the students were given a post-test identical to the pre-test.

Preliminary Results

Preliminary results show that the content of the mental models of over half of the students were significantly altered by the intervention. Before the intervention, almost all of the students held a mental model in which depressions and expansions in the economy are caused directly by isolated events that are often, but not always, external to the economy, e.g., the onset of war, a drought, revolutions in science and technology, and/or a crash in the stock market. The pretest mental models show almost no evidence that students see the decisions made by the managers of firms as important, that they see depressions and expansions as related in any way, or that they view

depressions or expansions in the economy as causal agents that affect other variables. After the intervention, many of the subjects mention key factors in the expert model of the long wave described by Sterman (1989), e.g., management decisions, excess capital, and depreciation, but these are mixed in with the key variables and events from the pre-test mental models. The new mental model discussed during the debriefing session clearly did not replace the original mental models held by the students but were integrated into them. (See Table I.)

Table I
Most Often Mentioned Factors in Mental Models of the Economic Long Wave ($N = 25$)

Pre-test		Post-test	
Factor	% Ss	Factor	% Ss
War	80%	War	68%
Technological advance	56%	Technological advance	60%
Unemployment rate	48%	Unemployment rate	60%
Consumer spending	40%	Management decisions	52%
Consumer confidence	40%	Overproduction of goods	48%
Government spending	28%	Demand for goods	48%
Stock market crash	28%	Government spending	36%
Overproduction of goods	28%	Excess capital	36%
Wages	20%	Stock market crash	32%
Natural disaster	16%	Depreciation	28%

Although the study demonstrated marked changes in the content of mental models due to the intervention, there is less evidence that students' mental models became more detailed, interconnected, or dynamic. Most of the students incorporated the important, basic points of the expert model into their mental model, but very few developed a detailed, comprehensive understanding of the expert model.

Conclusion

By applying a new, rigorous methodology this study has presented some of the first reliable evidence for the ability of systems thinking interventions to improve mental models. And, some of the problems identified by the study, e.g., the merging of old and new mental models, are likely to be generalizable to many other systems thinking interventions. It is our belief that this type of rigorous assessment of the mental models resulting from interventions cannot be avoided if systems thinking programs are to become more effective in the future.

References

Cavaleri, S., and Sterman, J. D. (1995). Towards evaluation of systems thinking interventions: A case study. *Proceedings of the 1995 International System Dynamics Society (Tokyo, Japan)*, pp. 398-407.

Ganter, S. L., Doyle, J. K., and Radzicki, M. J. (1995). Assessing system dynamics curricula: Past, present, and future. *Proceedings of the 1995 International System Dynamics Society (Tokyo, Japan)*, pp. 484-493.

Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*, 301-335.

Sterman, J. D., and Meadows, D. (1985). Strategem-2: A microcomputer simulation game of the Kondratiev cycle. *Simulation and Games, 16*(2), 174-202.