MENTAL MODELS OF DYNAMIC SYSTEMS

James K. Doyle, Department of Social Science and Policy Studies, Worcester Polytechnic Institute, USA

David N. Ford, Department of Civil Engineering, Texas A&M University, USA

Michael J. Radzicki, Department of Social Science and Policy Studies, Worcester Polytechnic Institute, USA

W. Scott Trees, Department of Economics, Siena College, USA

Key words

system dynamics, mental models, dynamic decision making, learning, causal reasoning, simulation, information processing, memory, cognitive processes, cognitive limitations, bounded rationality, feedback

Contents

Glossary

Bounded rationality: the theory that human decision making is rational only insofar as the rational solution does not require calculations or mental efforts that exceed cognitive limitations

Causal loop diagramming: A diagramming technique for representing the feedback structure of systems in which variables are linked by causal arrows that indicate the direction (positive or negative) of relationships

Chunk: a meaningfully organized grouping of information held in working memory

Cognitive process: any mental activity that acquires, stores, transforms, reduces, elaborates, or uses knowledge

Decision rule: a plan, strategy, or set of procedures that describes how decisions should be made

Dynamic decision making: the process of assessing and choosing among alternatives in the course of managing a system that changes over time

Feedback: the transmission and return of information

Group model building: A methodology for building system dynamics computer models in which a group or team of people participate actively and simultaneously in building the model

Heuristic: A simplifying “rule of thumb” used to make decisions when experience, time pressure, stress, complexity, habit, or other factors do not allow for a more complete analysis

Knowledge elicitation: the act of attempting to extract mental model information from memory and represent it externally (e.g., in a diagram or computer model) so that it can be shared and scrutinized

Long-term memory: a very large capacity memory store in which knowledge and life experiences can be held for long periods of time

Long wave: a pattern of alternating periods of overexpansion and depression in the economy with a period on the order of 50 years

Mental model: a relatively enduring and accessible, but limited, internal conceptual representation of an external system (historical, existing, or projected) whose structure is analogous to the perceived structure of that system

Mental simulation: the act of inferring the dynamic consequences of a mental model in one’s head without computer assistance

Nonlinearity: a relationship between variables or the time series of a single variable that is described by any mathematical function other than a straight line

Operational thinking: An approach to conceptualizing systems that emphasizes the "physics" of how the real world is constructed

Retrieval cue: a piece of information that enhances the ability to recover additional information from long-term memory

Virtual system: a computer model of a real system that can be simulated and explored to aid learning about the real system

Working memory: the “mental workbench” on which information can held in memory in a temporary state and manipulated during reasoning, problem solving, and decision making

Summary

Mental models play a central role in system dynamics efforts to improve learning and decision making in complex systems. In fact, the system dynamics methodology can be generally described as a feedback process in which mental models are used to develop a computer model, which in turn creates new opportunities for learning that improve the accuracy, coherence, and complexity of mental models. This article describes the history of the mental models concept in the fields of system dynamics and psychology and offers a comprehensive definition of the term for use in system dynamics research. The characteristics of mental models of dynamic systems identified by the empirical literature are reviewed, with an emphasis on important flaws and limitations, and their underlying causes, that typically limit the utility of mental models for dynamic decision making. A mental model based theory of dynamic decision making is presented that is consistent with this evidence, and the mechanisms by which system dynamics computer modeling can improve mental models within this theoretical framework are described. The implications of the theory for developing appropriate techniques for studying mental models, as well as specific priorities for future research, are discussed.

Introduction

The idea that human minds create, store, and manipulate internal models of the dynamic systems with which they interact has been central to the theory and practice of system dynamics since its inception. The nature and properties of these mental models in fact provide the primary rationale for the need to employ system dynamics modeling to improve learning and decision making in the face of complexity. According to the system dynamics view, people can (and should) manage systems by constructing a mental model and mentally simulating it to determine the likely outcomes of policy decisions. Learning occurs by comparing expectations with the actual observed consequences of policy decisions and using this outcome feedback to revise or update the mental model. However, both practical experience in the field of system dynamics and controlled laboratory experiments on dynamic decision making have shown that mental models of complex systems are typically subject to a variety of flaws and limitations. For example, mental models often omit feedback loops, time delays, and nonlinear relationships that are important determinants of system behavior. In addition, the limited capacity of working memory makes it impossible for people to mentally simulate the dynamic implications of all but the simplest mental models. According to the system dynamics view, only by adopting the feedback perspective and modeling discipline of system dynamics and taking advantage of the computer's ability to calculate the dynamic consequences of mental models can these flaws and limitations be overcome.

Despite the known flaws and limitations of mental models, they are often tapped as a primary source of information for system dynamics model building. In fact, a system dynamics modeling project typically begins with an effort to elicit or externalize the mental models of policymakers so that they can be shared and submitted to a process of scrutiny and evaluation. There are several advantages to such an emphasis on mental information to supplement the written and numerical databases that are vitally important to all approaches to modeling systems, including system dynamics. First, the mental database that people form through observation and experience is vastly larger than the other databases. Only a fraction of the knowledge people gain during their lives is ever written down, and an even smaller percentage of what is written is expressed numerically. Second, the mental database is more likely to contain the type of information that is needed to build system dynamics models, namely, the details of system structure and the cognitive processes by which managers make decisions. Third, the mental database can be more easily probed, and can provide information that allows modeling to proceed when written and numerical information is absent. Fourth, the elicitation and use of mental data allows managers to develop a sense of ownership of the resulting model, as well as to gradually see their own mental models revised and transformed, both of which aid learning and increase satisfaction with the modeling process. The main disadvantage with the use of mental data in model building, of course, is the increased potential for errors and biases, since the data have not been subject to the editorial or review processes that are typical of written and numerical data. System dynamics practitioners acknowledge this potential for error in data collection and have developed knowledge elicitation procedures that attempt to minimize it. Furthermore, they argue that the mental database is not so flawed that it can't serve as a useful starting point for modeling, and that the iterative nature of the modeling process will eventually uncover the important errors that do exist.

Given their importance for improving dynamic decision making and as an information source for model building, it is not surprising that system dynamics researchers have devoted an increasing amount of their research effort in recent years to the study of mental models. System dynamics practitioners have developed a variety of diagramming techniques for representing mental model information in ways that promote learning as well as substantial practical experience in how to design and implement group facilitation and group model building programs to change mental models and improve decision making. In recent years, efforts have increased to validate this practical knowledge about measuring and changing mental models through rigorous experimentation and assessment.

Definition

Despite its role as one of the most important concepts in system dynamics, the term "mental model" is also one of the least well defined. To some degree this is due simply to the inherent difficulties of defining any mental concept. Mental models are not directly observable, and their character must be inferred from observations of overt human behavior. They are also subject to what one system dynamics research group has termed "the mental model uncertainty principle," that is, the mere act of trying to understand or measure them may itself alter mental models. And, of course, there is the added difficult that researchers themselves must rely on potentially flawed or biased mental models while struggling to identify their character.

A second reason for the difficulty of system dynamics researchers in establishing a mutually agreeable definition of mental models is the checkered history of the concept. Since its first use by the psychologist F. I. M. Craik in his 1943 book The Nature of Explanation, the term mental model has taken on a variety of meanings, all of which are still in current usage. For example, in psychology and related fields, mental models have been variously referred to as mental diagrams or picture-like images, mental representations, intuitive theories, collections of beliefs, schemas, and knowledge networks. In system dynamics, Jay Forrester introduced the term to the field in his seminal work Industrial Dynamics in 1961, stating that mental models are “mental images or verbal descriptions . . . [that] substitute in our thinking for the real system that is represented.” In the intervening decades, system dynamics researchers have at times described mental models as intuitive generalizations, collections of ideas, opinions, and assumptions, networks of facts and concepts, and implicit causal maps of systems.

One recent effort to synthesize existing ideas into a useful form for system dynamics research and practice resulted in the following conceptual definition:

> A mental model of a dynamic system is a relatively enduring and accessible, but limited, internal conceptual representation of an external system (historical, existing, or projected) whose structure is analogous to the perceived structure of that system.

In this definition, the phrase “relatively enduring” implies that a mental model, while subject to changes that occur on the order of minutes or seconds, may persist in long-term memory in some form for years or decades. The word “accessible” suggests that people are generally consciously aware of their mental models and to a large degree can mentally inspect them and communicate them to others. “Limited” means that the term mental model should not refer to all of the knowledge held by an individual but to a precompiled subset of information in long-term memory. The maximum size of a mental model is thought to be determined by the capacity of working memory, the mental workbench on which people store information temporarily while thinking about it. Since the amount of information that can be organized into a meaningful grouping of information, or chunk, in working memory is flexible, the maximum size of a person's mental model may increase to some degree as they gain expertise and learn to organize information more efficiently. However, no amount of experience or expertise can alter the maximum number of chunks. Due to unalterable cognitive limitations such as this, it is not surprising that, as Jay Forrester, the founder of the field of system dynamics, has remarked, most mental models are usually no more complex than "a fourth-order differential equation."

The term "internal" in this definition is meant to indicate that mental models are cognitive phenomena, that is, they exist only in the mind and they should not be confused with the results of efforts to "elicit," "surface," "map," or "measure" mental models due to the strong possibility of measurement error or bias. The word "conceptual" restricts the definition of mental models, for system dynamics purposes, to models composed of symbols, concepts, ideas, or other language-like components rather than mental imagery. The phrase "representation of an external system" implies, first, that mental models are cognitive structures that store information, rather than cognitive processes which transform information, and second, that mental models refer to or represent objects, processes, information, or mental constructs that are outside the boundary of the mental model. The use of the term "structure" implies that mental models are not simply knowledge but knowledge that has been organized and interconnected in some way.
Finally, it should be noted that the word "perceived" in the final phrase is important to make it clear that an individual's mental model of a system may or may not bear a resemblance to the real system, depending on the accuracy of their perception.

It should be noted that the field of system dynamics has yet to converge on a widely agreed upon definition of mental models, which is not surprising given the ethereal nature of mental models and the inherent difficulties of studying them. The above-stated definition, although the most comprehensive available, should not be considered to be correct or complete but merely a useful starting point for further refinement and debate.

Characteristics

The above-mentioned definition tells us, from a theoretical perspective, what characteristics mental phenomena <u>must</u> have in order to be classified as "mental models" for system dynamics purposes. However, it is also important to consider, from a practical point of view, what characteristics mental models <u>are likely to</u> have when people make dynamic decisions in the face of complexity. There is in fact a substantial body of experimental evidence, collected in the fields of system dynamics and experimental psychology, that has identified a number of errors and biases in mental models and the dynamic decisions based on them. Among the most commonly reported shortcomings of mental models of dynamic systems are the following:

1. Mental models are typically vastly oversimplified. While all models must simplify reality in order to be useful, mental models are often so simple that they omit information that is crucial for understanding the dynamics of systems. For an extreme example, in a study of novice mental models of the causes of long-term cycles of expansion and depression in the economy, one subject's model (despite several opportunities to relay more detailed information) seemed to consist only of the following link:

   $$\text{Taxes} \xrightarrow{\qquad\qquad -} \text{GNP}$$

   where the negative sign, consistent with standard causal loop diagramming notation, indicates that, all else being equal, taxes and GNP move in opposite directions. That is, if taxes were to increase, GNP would be expected to decrease more than it otherwise would; if taxes were to decrease, GNP would be expected to increase more than it otherwise would.

2. Mental models can only rarely be mentally simulated accurately, primarily because even relatively simple dynamic models outstrip the mind's ability to calculate or infer their consequences.

3. Mental models are typically "sloppy" or "messy." For example, their causal pathways often have gaps or omissions and sometimes lead to dead ends; signs of relationships between variables are often ambiguous; and variables are rarely described quantitatively (see Fig. 1). This lack of completeness and coherence acts as a further hindrance to mental simulation.

4. Mental models, particularly those of relative novices, are prone to errors and biases that result from biased information processing, unwarranted assumptions, overconfidence in one's knowledge base, and other barriers to learning. For example, studies have shown that many people hold a "pre-Newtonian" mental model of physical systems, despite a lifetime of

experience with objects in motion under the influence of gravity, and they subsequently cannot accurately predict the paths of moving objects such as a bomb dropped from a plane or a ball flung from a roulette wheel.

5. The boundaries of mental models are "fuzzy," that is, they are ill-defined and easily changeable. This is due to the structure of human memory, in which mental models are interconnected with other information in memory in a complex network of associations.

6. Mental models tend to be highly unstable over time, at least in their details. This is due in part to the cue-dependent nature of human memory recall. That is, finding a particular piece of information in memory requires the presence of an appropriate reminder or "cue" to its location. Thus which parts of a mental model are brought to bear on a particular decision or problem depends on what external and internal cues are present at that particular time to "jog" one's memory. Of course, simple forgetting may also alter the content of a mental model over time.

7. Mental models typically fail to account for important time delays that can create instabilities in systems. For example, people who play the well-known Beer Distribution Game often completely ignore delays in the production system when making ordering decisions, even though the delays are made highly salient and visible.

8. People have particular difficulty perceiving and representing relationships between variables or time series data that are nonlinear. Such relationships are almost always simplified and represented mentally in a linear fashion. A well-known example of misguided linear thinking is the "paper-folding task," in which people are asked to estimate how thick a piece of paper would be after 100 folds. The majority of people, assuming linear growth, give answers on the order of inches or feet, whereas the correct answer for this exponential growth problem is in fact 800 trillion times the distance between the earth and the sun!

9. Mental models often fail to incorporate important feedback mechanisms. Rather than adopting a closed-loop perspective in which decisions produce outcomes which feed back to inform and alter decisions, people tend to adopt an "open-loop" perspective in which one event leads to another in a single causal chain that, because it follows a strict timeline, is unidirectional.

10. Causation in mental models is generally represented in an overly simplified way. For example, people tend to seek simple, isolated causes for events rather than create networks of multiple, interrelated causes. They also show a preference for causes that are close in time and space to their effects (and occasionally confuse mere temporal correlation with causation). In addition, they tend to view causes as external to the system in question rather than see the structure of the system itself as an important causal factor.

11. The mental models of people who are not educated in system dynamics often fail to embody operational thinking, that is, people rarely spontaneously construct system dynamics stock/flow models that physically describe how systems work. When they do think operationally people often have great difficulty distinguishing which variables are stocks and

which are flows and also have trouble mentally inferring the behavior over time of stocks from information about flows (and vice versa).

12. There are inherent time delays involved in changing mental models. When mental models are updated with new information, the old information is not instantaneously forgotten, but persists in memory alongside the new information, where it may still be recalled and influence decision making.

Many of these characteristics were apparent in the great majority of the mental models of participants in the study of student understanding of long-term economic cycles or “the long wave” cited previously. Figure 2 shows “composite” diagrams that illustrate the mental models of the novice participants both before and after a management flight simulator based educational intervention, as well as a diagram representing a simplified version of an expert model of the long wave. According to the expert model, which focuses on the capital-producing sector of the economy, the capital investment decisions of managers cause overexpansion of capacity due to time delays in production and the reinforcing nature of capital self-ordering. This reinforcing loop is amplified by the effect of rising employment and wages, which lead to increased efforts to substitute capital for labor that creates additional demand for capital. The long time necessary for this excess capital stock to depreciate leads to a subsequent period of economic depression when there is little demand for capital production. Note that the expert model represents the feedback perspective of system dynamics, in which causality is embodied by an interconnected network of feedback loops, effects are remote in time from their root causes, and system behavior arises not from the impact of external events but from the structure of the system itself. In contrast, the novice models are typically simple, often incomplete, chains of events in which immediate, external events (e.g., wars or technological breakthroughs) are the primary causal factors. Important system dynamics concepts such as feedback, the impact of time delays, and operational thinking are virtually absent. The relationships between variables are not always specified. Note also that, although learning did take place, as evidenced by the inclusion of such expert concepts as managerial decisions and excess production in the post-test but not the pretest diagrams, this new information did not apparently replace or even integrate very well with subjects' pre-intervention notions.

To some degree the above-stated shortcomings are due to the intrinsic properties of complex dynamic systems that are known to inhibit learning, for example, unknown dynamic structure, time delays in receiving feedback, the existence of multiple, confounding causal variables, and nonlinearities. However, the same errors, omissions, and biases tend to be found when people interact with simplified systems in the laboratory in which they have perfect information, receive instantaneous feedback, and can isolate potentially causal variables through controlled experimentation. This strongly suggests that the primary determinants of shortcomings in mental models of dynamic systems are not external barriers but barriers that are internal to the mind, namely, strict and inviolable limits on the mind’s ability to process information.

The most important such cognitive limitation is most likely the limited capacity of short-term or working memory, the "mental workbench" on which our mind holds information temporarily while consciously thinking about it. This limit has been determined to be only 5 to 9 "chunks" or meaningful groupings of information. Although the amount of information that can be packed into each chunk is somewhat flexible and can be increased by grouping information together more efficiently, the number of such chunks can never be more than about 7. This apparently inviolable limit places a severe restriction on both the size of mental models of systems and human ability to compute their dynamic consequences. Other cognitive limitations also play important roles in acquiring and using information about complex systems. For example, knowledge acquisition is constrained by limits in the kinds of information human senses can perceive, by the limited capacity of perceptual channels, by human inability to focus conscious attention on more than one thing at a time (unless processing of one of the stimuli is so well-practiced that it can be done automatically), and by the limited rate at which information can be transferred into long-term memory. Use of knowledge in decision making is also limited by, among other cognitive processes, simple forgetting, interference effects (in which intervening events inhibit recall of an earlier event), proactive inhibition (in which old, outdated information is recalled instead of more recent, updated information), and context effects (in which learned information cannot be retrieved from memory when needed due to the absence of effective retrieval cues).

These sorts of cognitive limitations fall under the umbrella of Herbert Simon's well-known principle of "bounded rationality," which proposes that human decision making is rational only insofar as the rational solution does not require calculations or mental efforts that exceed cognitive limitations. However, there are also many well-documented forms of biased information processing that produce errors in mental models even when cognitive limits have not been exceeded. For example, when evaluating data people often perceive correlations that do not in fact exist, fail to perceive even very strong correlations, perceive patterns in data that in reality are completely random, and see relationships that they expect or want to see whether the data supports them or not. When testing hypotheses, people have a strong bias to focus on confirming rather than disconfirming information and often ignore such crucial information as base-rate probabilities. When assigning causes to events people are prone to confusing correlation and causation and to making the "fundamental attribution error," whereby they mistakenly assume that someone's behavior is due to a stable personality trait rather than the environment or circumstances in which the behavior occurs. Such flawed information processing can lead not only to errors and biases, but also to fixed mental models that are not open to change in the face of new information and evidence.

In summary, the research record that describes the common characteristics of mental models provides a compelling argument for the need for system dynamics to improve learning and decision making in the face of complexity. The next two sections describe a theory of dynamic decision making that is consistent with the evidence and explain how computer simulation modeling using system dynamics principles can improve dynamic decision making.

The role of mental models in dynamic decision-making

System dynamics is one of a growing number of applied disciplines (including human-computer interaction, risk communication, and educational assessment) that view mental models as centrally important to dynamic decision making and problem solving. According to the system dynamics view, human management of a complex system can be thought of in general terms as a continual feedback process in which actions taken by managers change the state of the system, which in turn provides feedback that is used to formulate a new set of actions designed to bring the system closer to the desired state (see Fig. 3). More specifically, the system provides feedback in the form of measurable and perceptible cues. Which cues to the state of the system are selected depends on attention and scanning processes that are informed by mental models, that is, the mental model helps direct attentional resources to the most relevant information (L1). From these cues people form perceptions of the state of the system which may be accurate or inaccurate depending on the quality and complexity of the available cues (e.g., the presence or absence of random error, confounding variables, or nonlinearities). These perceptions may guide the construction of a mental model of the system, which may in turn influence how feedback is perceived (L2).

Selection of a decision rule as a basis for action may proceed along one of two pathways. People may use their perceptions to alter their mental model of the system and then conduct a mental simulation run to predict the future state of the system, given the newly perceived feedback. Any perceived gap between the mentally simulated state of the system and the goal, or desired state of the system, motivates a search for the most potentially effective decision rule. In the mental model based decision making loop this search takes the form of thought experiments that mentally simulate the effects of alternate possible decision rules, and the rule whose simulated effects move the system closest to its desired state is adopted (L3).

Alternatively, people may bypass the mental model loop and use any gap between the goal state and perceived state of the system to motivate a memory search for a potentially effective decision rule. People generally have stored in memory a variety of decision rules, particularly simplifying heuristics, that have been used with some success to make prior decisions. In the heuristic decision making loop people select the decision rule that, based on prior experience in similar circumstances, they expect will best close the gap between the system's perceived and desired state (L4).

For most people, in most situations, with most systems, the simpler best-matching-heuristic loop dominates because the effectiveness of the mental-model-based loop is severely constrained by cognitive limitations and (quite often) poor reasoning skills. Once a decision rule is selected by the dominant decision making loop, people form an intention to act. These behavioral intentions, which are often constrained by a variety of barriers to action (for example, competing agendas, lack of resources, risk aversion), form the basis of action that changes the state of the system, and the cycle repeats.

Time delays can have a marked impact on the effectiveness of both decision making loops at several points. For example, in some cases it can take years or decades for important cues to the state of the system to become readily apparent. It also takes time for individuals and organizations to formulate decision rules and take action, whether they are mental-model-based or not. And, of course, it takes time for actions to achieve their desired (or, in some cases, undesired) effects. The effect of time delays is well illustrated by the case of ozone depletion. For example, CFCs were widely used for decades before their effect on the ozone layer was first detected in the mid-1970s. A period of more than 10 years then elapsed before effective action, in the form of an international treaty, was decided upon. Finally, this action is not expected to have its desired effect (returning the ozone layer to pre-CFC levels) until the middle of this century, due to the long residence time of CFCs in the stratosphere and the catalytic nature of their effect on ozone.

From the system dynamics perspective, the heuristic loop has several disadvantages that limit its effectiveness. Since heuristic thinking takes place in the context of existing mental models, its effectiveness is limited by the quantity and quality of the decision maker's past experience and it is often strongly biased toward organizational and cultural norms. Since heuristic thinking does not improve the existing mental model, it cannot in turn provide better focus to attention and scanning processes or foster the development of more accurate perceptions. In contrast, the mental model loop, while strongly inhibited by cognitive limitations, offers much more potential for effective learning, if means can be developed to overcome these limitations. In addition, mental models, which are informed by goals, form the basis of mental simulations that may provide insights that improve the specificity and appropriateness of the goals (L5). The challenge for system dynamics has been, and will continue to be, to devise effective ways to change the dominant dynamic decision making process from the heuristic loop to the mental model loop.

The role of system dynamics in improving mental models

Given the central role of mental models in the system dynamics theory of dynamic decision making, it is not surprising that one of the primary goals of system dynamics interventions is to change mental models to make them more accurate, more complex, and more dynamic. In system dynamics this goal is accomplished by the construction of a system dynamics computer model that provides a "virtual system" that people can simulate to observe the effects of alternate decision rules on the system (see Fig. 4). The virtual system can be thought of as being embedded within the real system, providing a virtual learning loop (L6) that has several advantages over learning processes that rely solely on mental simulation:

1. Since the computer's ability to manipulate information vastly exceeds the capabilities of human working memory, people can create and simulate models of far greater complexity.

2. The process of building a working computer model encourages operational thinking as well as completeness, coherence, and quantification in formulating mental models.

3. While of course a computer model may contain errors and biases, the computer simulation, in contrast to a mental simulation, is consistent and reliable.

4. The computer's ability to efficiently and accurately store and retrieve information exceeds the capabilities of human long-term memory, which is subject to forgetting, retrieval failure, and distortion.

5. As Figure 4 suggests, in comparison to the real system, the virtual system is not subject to significant time delays. Time can be accelerated to the point that actions can be implemented and feedback can be received almost instantaneously. Hundreds or thousands of simulations

can be run in the virtual system in the time it would take to cycle through the real system just once. Alternatively, compared to some real systems, time can be slowed down to offer more opportunity for reflection and planning before taking action.

6. The virtual system allows for the type of systematic, scientific experimentation that is rarely possible with real systems. Variables can be changed one at a time to observe their effects in isolation. Actions that would be avoided in the real system due to irreversible consequences can be taken without fear in the virtual system.

7. Unlike real feedback, virtual feedback is complete, unambiguous, and always perceptible.

8. In the virtual system, there are no barriers to action and decision rules are implemented perfectly.

9. Unlike most real systems, the structure of the virtual system is completely open and available for inspection.

The virtual system can effectively interact with the real system at several points to aid learning and decision making. For example, at first the virtual system is based solely on the existing mental model of the real system, but then lessons learned from the virtual world can begin to change the mental model of the real system (loops L7 and L9; note that the interaction between the computer model and the mental model of the real system is mediated by a mental model of the virtual system, which represents the user's current understanding of the computer model). In turn, lessons learned from applying virtual learning to the real world feed back to improve the quality of the virtual system, and the cycle repeats. In a similar fashion, decision rules and actions taken in the real world suggest decisions and actions to try in the virtual world, which then suggest new decisions and actions to attempt in the real world (loops L11 and L12). Perceptions of the real world affect perceptions of the virtual world, and vice versa (L13). Mental simulation at first guides computer simulation, which feeds back to improve mental simulation ability (L10).

Of course, the increased potential for effective learning offered by virtual systems presupposes that people can and will use the simulation tools systematically and scientifically and will apply good reasoning skills to simulation results. However, the available research suggests that these skills don't come naturally. For example, beginners interacting with simulations tend to rush into trying the first strategy that comes to mind rather than thinking carefully. Also, even when people do stop to think carefully during simulation, they often fail to change variables systematically. Therefore efforts to teach system dynamics-based decision making should emphasize the need for a scientific approach to learning through simulation.

Does the available evidence suggest that the development and use of system dynamics computer models improves mental models and dynamic decisions? Many successful case studies have appeared in the literature, but it is not always clear that the successes described arise from improved mental models or other factors or even whether or not the decisions were arrived at by a mental model based or heuristic based approach. There has also been comparatively little controlled experimental research that would allow generalizations to be drawn beyond particular cases. The limited controlled research that is available, however, does offer qualified support to the ability of system dynamics-based interventions to improve mental models.

In one study of the effectiveness of an educational intervention based on a management flight simulator, the intervention reliably increased the size of mental models, changed the content of mental models toward an expert model, increased the number of mental models that included feedback, and increased subjects' confidence in their mental models. However, the intervention did not produce mental models that were more intricately interconnected or that included longer chains of causal reasoning in which ultimate causes and effects were more widely separated in time and space. In addition, the post-intervention mental models still had many of the undesirable characteristics commonly observed in mental models, e.g., lack of completeness and coherence, ambiguity, and failure to acknowledge time delays. This research is only preliminary and much work remains to be done before these kinds of questions can be answered definitively. The next two sections describe the general requirements for research on mental models of dynamic systems as well as more specific research needs.

Principles for mental models research

Accurate, unbiased elicitation and mapping of mental model structure and content is important in system dynamics for several reasons: (1) to provide accurate information for model building; (2) to most effectively design learning interventions (interventions should account for the impact existing mental models will have on knowledge acquisition); and (3) to rigorously assess the effectiveness of educational interventions based on system dynamics. However, their elusive, unstable, nature, combined with the overwhelming complexity of the human mind (Scientific American once referred to the human mind/brain as "the most complex object in the known universe") make accurate measurement of mental models an exceedingly difficult task. In addition to the usual concerns of experimental research on human subjects such as experimental control, external validity, and the potential for experimenter and subject bias, there are several special concerns that should be addressed when conducting research on mental models. For example:

1. Researchers should strive to capture the messiness, sloppiness, lack of completeness, and fuzziness that exists in the mental models they study rather than adopt commonly employed methods that invariably produce neater, cleaner, and more complete representations of mental models. Glossing over the true, unrefined character of most mental models makes it more difficult to design learning interventions and to assess their effectiveness.

2. Since mental models exist only in human memory, they should not be confused with the external products (for example, causal loop diagrams and computer models) that are based on them. Diagrams and models that are the result of hours, days, or weeks of cognitive effort are likely to be much more complete and complex than the internalized mental model that drives dynamic decision making. External models may also include information that was only held in a temporary state in working memory and was never fully learned. And, internal mental models, of course, are subject to processes that externalized models are not, such as forgetting, interference, retrieval failure, and distortion. For these reasons researchers should assess mental models by testing what is in memory at a particular point in time.

3. Since mental models tend to be unstable, particularly in their details, researchers should routinely verify that any identified changes due to a system dynamics intervention persist for a reasonable amount of time.

4. The relationship between mental models and behavioral intentions is not necessarily straightforward and predictable. This is because decisions may be made on the basis of heuristics rather than mental models, or alternatively, people simply may not understand the behavioral implications of their mental models. It is even possible for a wildly incorrect mental model to be more likely than the correct model to result in desired behavior. For example, one study of public understanding of how thermostats work found that people typically hold one of two mental models: (a) the correct feedback model in which the furnace turns on only when the temperature drops below the thermostat setting and stays on only until the setting is reestablished or (b) an incorrect model in which the furnace is assumed to be operating all the time and turning the thermostat up or down increases or decreases the output of this continual operation. Results showed that people who held the incorrect model conserved much more energy than those with a better understanding of the system. Thus, to guard against mistaken assumptions about how mental models relate to behavior, researchers should carefully document both.

5. Mental models, due to the cue-dependent nature of memory recall and the reliance on inference to fill in memory gaps, are easily influenced by even seemingly subtle changes in the ways knowledge is elicited. Even small changes in the choice of words or the order in which questions are asked can bring different information to mind and result in the "elicitation" of a somewhat different mental model than the one that truly drives dynamic decision making. It is therefore important for researchers to adopt methods that are as naturalistic as possible, that is, that correspond to the settings, tasks, and question formats that people normally deal with when they think about dynamic systems.

6. Mental models tend to be highly variable and idiosyncratic. For example, in the study of mental models of economic cycles cited previously, more than 100 different causal variables were mentioned by at least one person. This high level of variation in mental model content means that large numbers of subjects are generally necessary to obtain sufficient statistical power.

7. When people are probed for knowledge, they often feel compelled to provide an answer even when they don't know the answer or have very little confidence in their knowledge. To avoid mischaracterizing mental models, researchers should routinely assess the level of confidence with which elicited knowledge is held.

8. The true nature of the organization of mental model information in memory is not fully known. Also, the nature of mental models most likely changes qualitatively as people develop expertise (for example, moving from an event-oriented to a more abstract variable-oriented perspective). It is therefore important to use elicitation methods that do not impose a particular structure on the elicited information, but allow the structure to arise from subjects' responses.

Mental models research in the field of system dynamics has not yet addressed all of these concerns and it remains a challenge for the field to develop and apply new methods that fully acknowledge the difficulties and complexities of conducting research on mental models.

Priorities for future research

In recent years the field of system dynamics has increasingly come to the realization that system dynamics models cannot effectively improve systems without careful attention being paid to psychological processes such as learning, knowledge elicitation, changing mental models, and group dynamics that determine whether or not modeling insights are translated into effective action. This growing emphasis on "modeling for learning" has resulted in a large number of promising system dynamics based techniques and strategies (e.g., group model building and management flight simulators) for improving mental models and dynamic decision making. These techniques have been widely applied in business, governmental, and educational settings, resulting in many satisfied clients and students and credible, albeit anecdotal claims of success. However, the rapid growth of the field has far outpaced the ability of assessment research to validate the apparent successes. And, the increasing demand for practical application of the techniques has inhibited the growth of more basic research.

As a result, many important questions about system dynamics theory and methodology relating to mental models remain unanswered. The following represent some of the most pressing needs for future research on mental models of dynamic systems:

1. Theoretical work is needed to further develop and specify the mental model based theory of dynamic decision making, accompanied by empirical work that tests the basic assumptions of the theory. For example, what are the specific cognitive processes by which mental models are updated or changed, perceptions of the state of the system are formed, and intended actions are translated into actual behaviors?

2. The ability of system dynamics programs to make mental models more complete, coherent, complex, dynamic, and feedback-oriented currently has only very limited empirical support. More basic research is needed to document the specific qualitative as well as quantitative effects of interaction with system dynamics models on mental models. Rigorous studies of long-term learning programs based on system dynamics (that is, longer than a week or so) are particularly rare.

3. There are several well-documented cases in which improved mental models of systems failed to lead to better decisions. Is this phenomenon relatively common or rare? Research needs to be done to identify the circumstances under which improving mental models does and does not lead to better management of systems. For example, can people generally accurately infer the implications of a simulation model for decision making? What are the factors in an organization that can prevent people from acting in a manner that is consistent with their mental models?

4. Despite the well-known sensitivity of mental models to differences in the way questions are framed or asked, a wide variety of distinctly different knowledge elicitation techniques are currently used to study mental models. Controlled comparisons of these techniques need to be conducted to identify which are the most accurate and unbiased.

5. A wide variety of methods and practices are currently used for teaching and facilitating the use of diagramming techniques, modeling, group model building, and management flight simulators. Controlled comparisons need to be conducted to identify the best existing practices for particular circumstances as well as to generate hypotheses for improving existing practices.

6. Nearly all of the reported investigations of mental models of dynamic systems study people who are beginners or novices in system dynamics. Research is needed to document how mental models change qualitatively and quantitatively as people gain experience and expertise in system dynamics.

7. The heavy reliance on mental models as a source of information for building system dynamics models, while often necessary, is troubling given the many documented shortcomings of mental models. Research is needed that clearly delineates the relative advantages and disadvantages of mental models versus other sources of information for model building. In addition, guidelines need to be developed for practitioners to help them determine which people under what circumstances can be expected to hold accurate and unbiased mental models.

8. Although computer simulation goes a long way toward reducing the effects of cognitive limitations on mental models, there are undoubtedly still limits on the size and complexity of simulation-aided mental models. What are these limits? Are there important problems or sets of problems that exceed these limits?

Addressing these questions about mental models of dynamic systems will be difficult and time-consuming, but their answers ultimately hold the promise of greatly improving human ability to learn about and manage complex systems.

Acknowledgments

The authors would like to thank David Lane for his helpful comments on defining mental models. James Doyle gratefully acknowledges the support of the System Dynamics Group, Sloan School of Management, MIT, where this article was completed during a sabbatical leave.

Annotated Bibliography

Axelrod, R., ed. (1976). *The Structure of Decision: The Cognitive Maps of Political Elites*. Princeton, NJ: Princeton University Press. [A classic work that employs content analysis techniques to understand the mental models underlying political decision making.]

Baddeley, A. (1998). *Human Memory: Theory and Practice*. Boston, MA: Allyn & Bacon. [A good overview of current knowledge about the workings of human memory.]

Craik, K. (1943). *The Nature of Explanation*. Cambridge: Cambridge University Press. [This book contains the first description of mental models in the psychology literature.]

Doyle, J. K., and Ford, D. N. (1998). Mental models concepts for system dynamics research. *System Dynamics Review*, *14*(1), 3-29. [This paper contains the most comprehensive definition of mental models of dynamic systems currently available.]

Doyle, J. K. (1997). The cognitive psychology of systems thinking. *System Dynamics Review*, *13*(3), 253-265. [This paper describes the need for rigorous controlled research on the effect of system dynamics modeling on mental models and cognitive processes.]

Forrester, J. (1961). *Industrial Dynamics*. Portland, OR: Productivity Press. [Forrester's seminal work contains the first description of mental models in the system dynamics literature.]

Gentner, D., and Stevens, A. L., eds. (1983). *Mental Models*. Hillsdale, NJ: Elrbaum. [An important collection of empirical papers on mental models of simple physical systems.]

Johnson-Laird, P. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, MA: Cambridge University Press. [This classic book presents a mental-model based theory of logical reasoning.]

Kahneman, D., Slovic, P., and Tversky, A., eds. (1982). *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press. [This edited work is an excellent introduction to the large body of empirical research documenting the use of heuristics in decision making.]

Richardson, G. P., Anderson, D. F., Maxwell, T. A., and Stewart, T. R. (1994). Foundations of mental models research. Proceedings of the 12th International System Dynamics Conference, Stirling, Scotland, 11-15 July. [This paper offers perhaps the most detailed theoretical statement on the role of mental models in dynamic decision making currently available.]

Rips, L. (1986). Mental muddles. In *The Representation of Knowledge and Belief* (M. Brand and R. M. Harnish, eds.), pp. 258-286. Tucson, AZ: University of Arizona Press.
[An important critique of the mental models approach to decision making and problem solving.]

Rouse, W. B., and Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, *100*(3), 349-363. [A useful review article of mental models research in human-computer interaction with lasting relevance.]

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129-138. [The classic paper that originated the concept of "bounded rationality."]

Sterman, J. D. (1994) Learning in and about complex systems. *System Dynamics Review*, *10*(2/3), 291-330. [This paper presents a compelling and detailed explanation of the need for and benefits of system dynamics modeling to improve mental models and dynamic decision making.]

Vennix, J. A. C. (1990). Mental Models and Computer Models: Design and Evaluation of a Computer-Based Learning Environment for Policy-Making. Den Haag: CIP-Gegevens KoninklijkeBibliotheek. [This book describes what is perhaps the most rigorous, detailed empirical study of changes in mental models of dynamic systems ever attempted.]

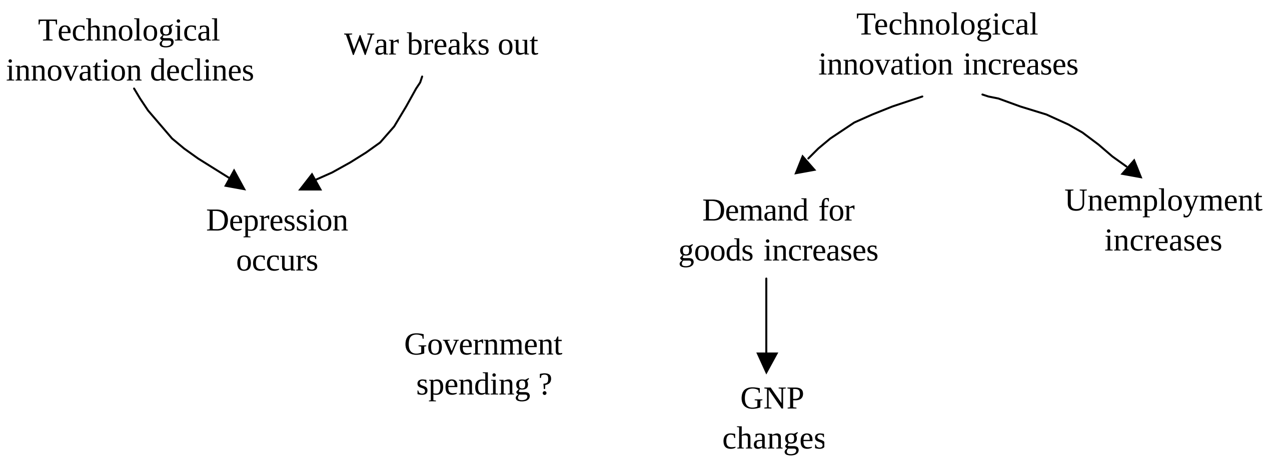
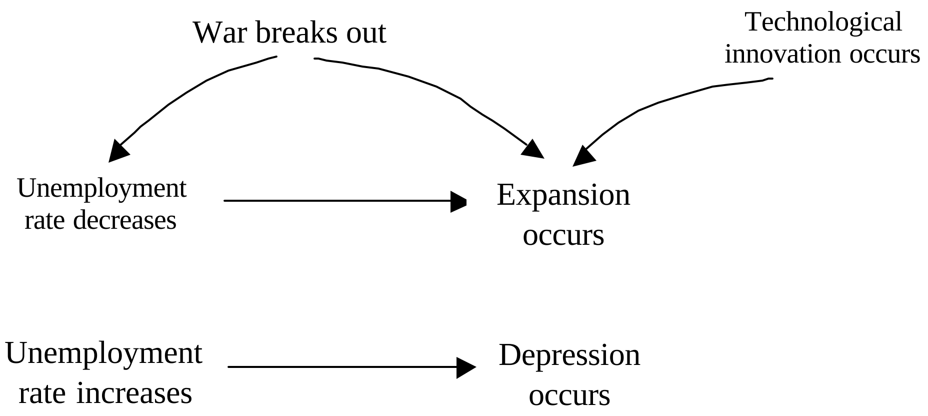
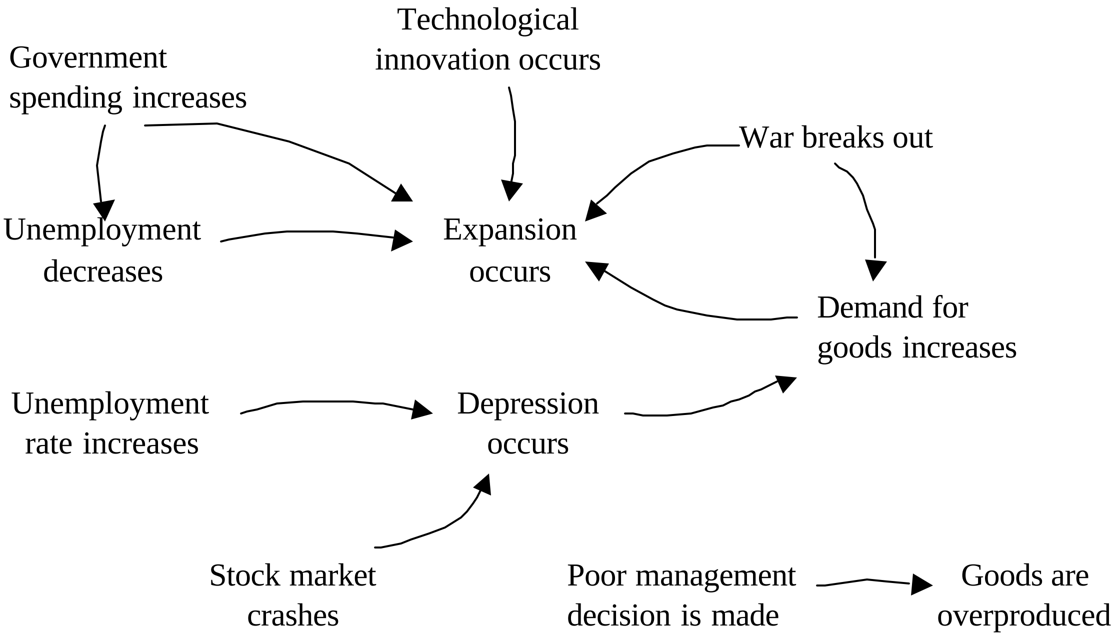


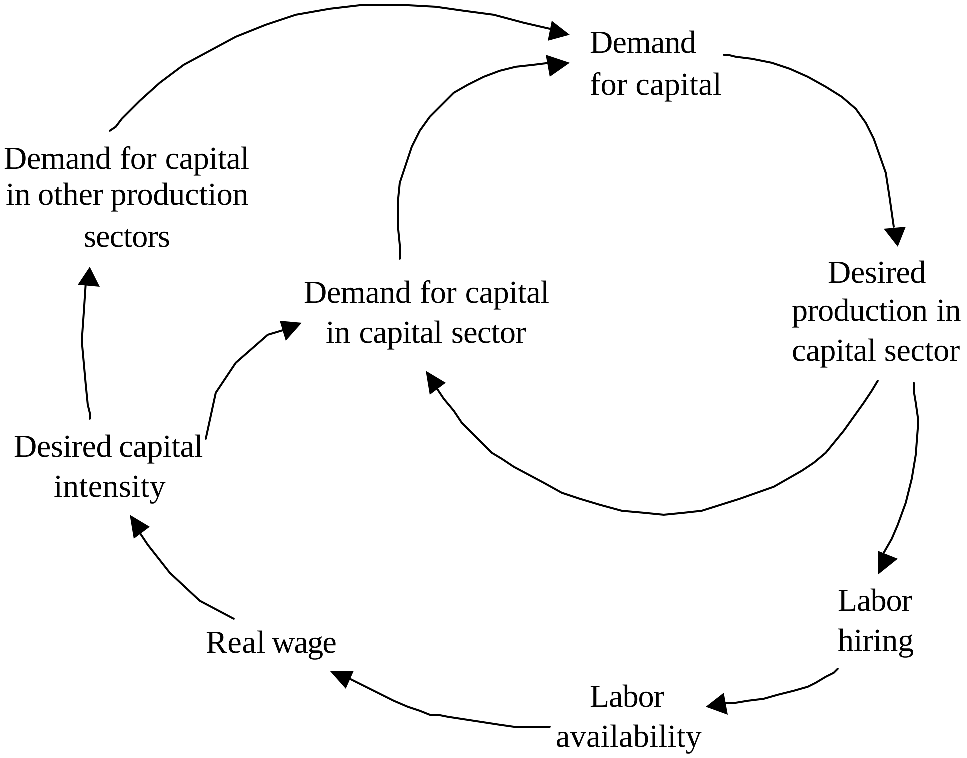
Figure 1. Representation of a typical novice mental model of the causes of economic expansions and depressions. In addition to being event-based, overly simple, incomplete, and quite at odds with expert models, it is also quite sloppy. For example, an increase in technological innovation is thought to increase demand for goods . . . which changes GNP in an unknown direction with unknown consequences. Also, government spending is thought to be relevant in some fashion that is not explicitly described.


War breaks out
Technological innovation occurs
Unemployment rate decreases
Expansion occurs
Unemployment rate increases
Depression occurs

(a)


Technological innovation occurs
Government spending increases
War breaks out
Unemployment decreases
Expansion occurs
Demand for goods increases
Unemployment rate increases
Depression occurs
Stock market crashes
Poor management decision is made
Goods are overproduced

(b)



(c)

Figure 2. Representation of novice and expert mental models of the causes of a long-term cycle of expansion and depression in the economy. (a) Composite diagram representing the most common concepts in pre-intervention novice mental models. (b) Composite diagram representing the most common concepts in post-intervention novice mental models. (c) Diagram representing part of a simplified version of the expert model. Fig. 2c from Sterman, J. D. (1986). The economic long wave: theory and evidence. *System Dynamics Review*, *2*(2), 87-125.

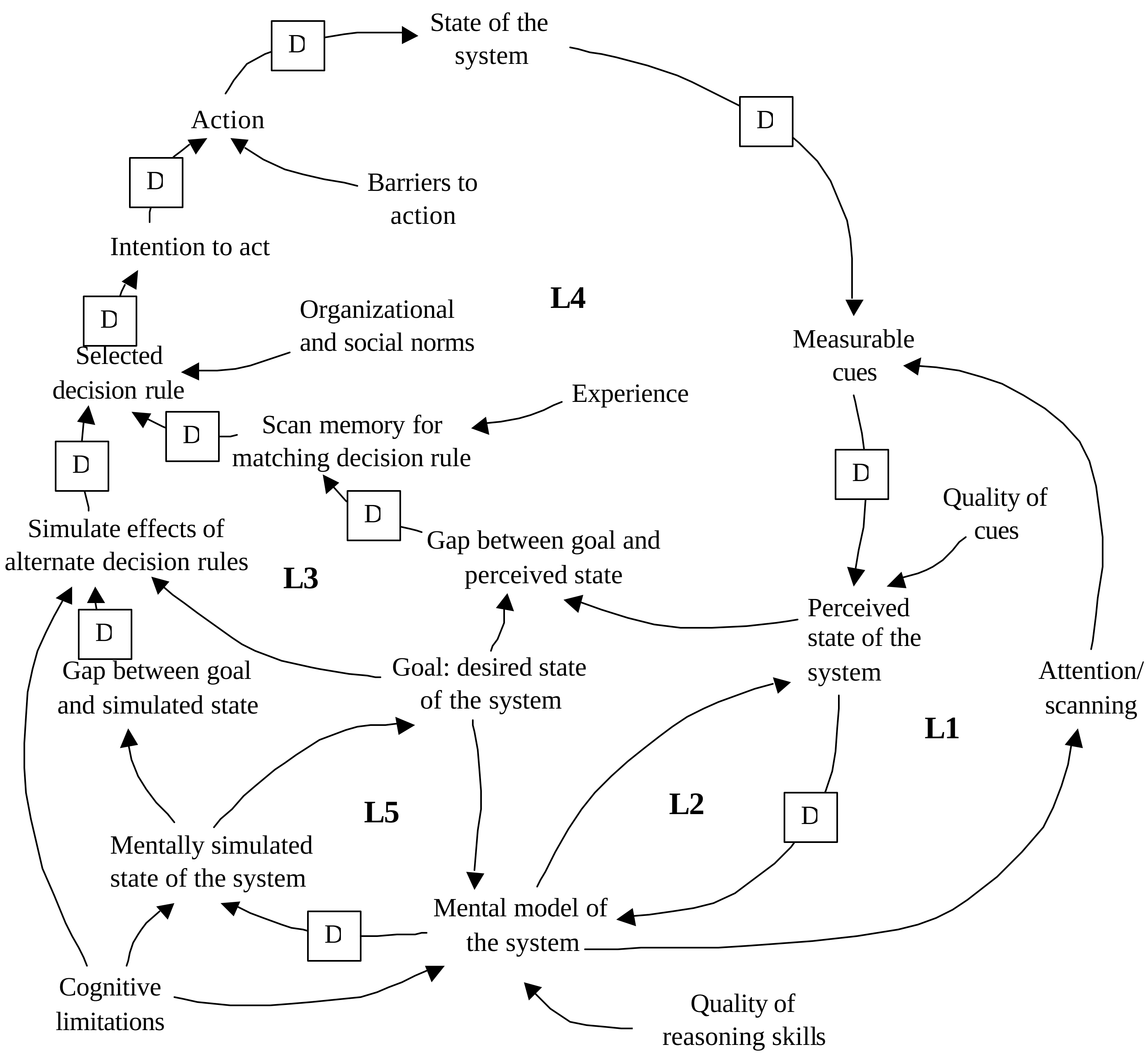


Figure 3. Overview of the system dynamics theory of dynamic decision making. L1: attention/scanning loop. L2: mental model/perception loop. L3: mental model based decision making loop. L4: heuristic decision making loop. L5: goal-changing loop. D = important time delay.

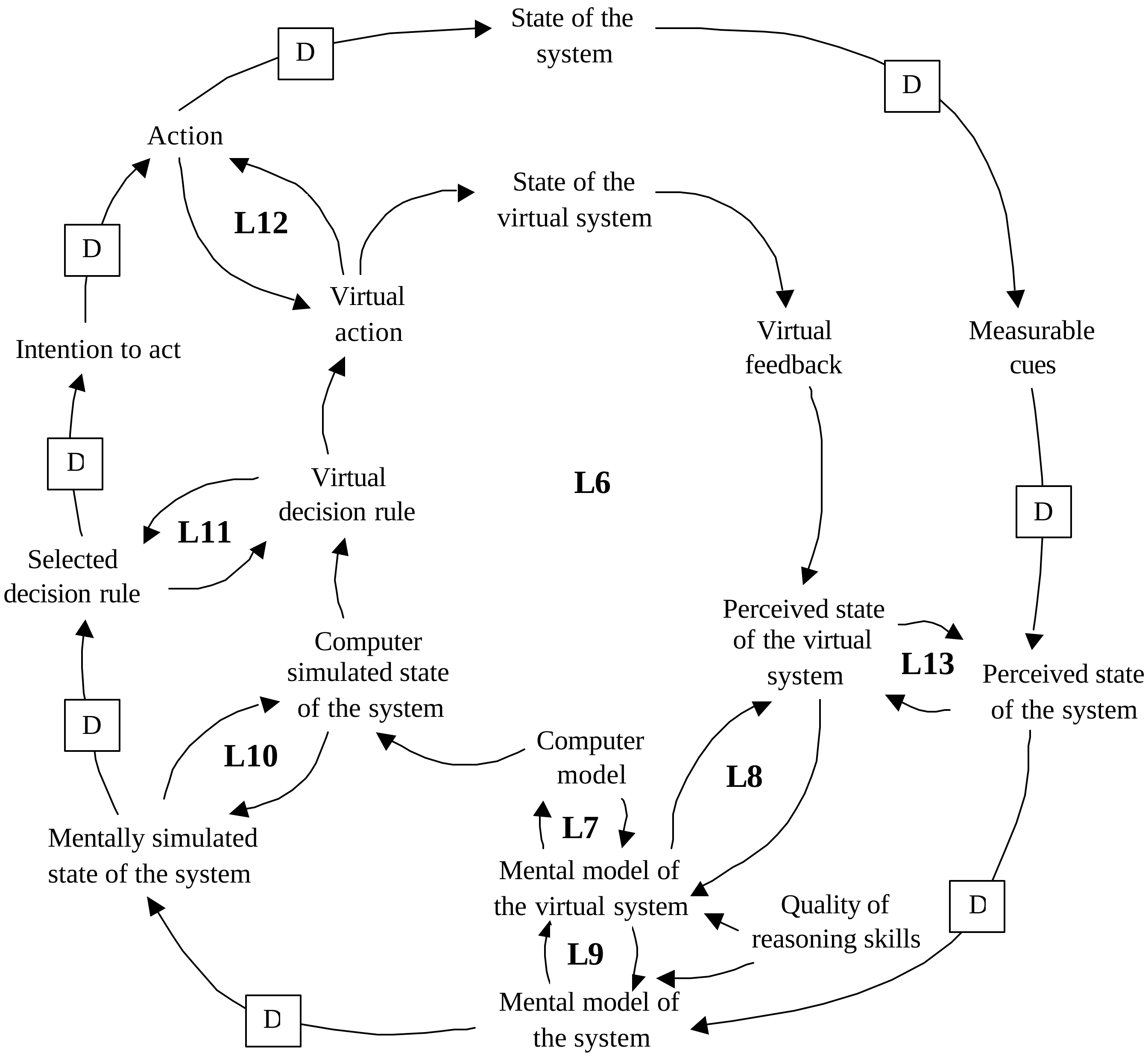


Figure 4. Overview of the idealized system dynamics approach to computer simulation aided dynamic decision making. For the sake of clarity, loop L3 has been simplified and loops L1, L2, L4, and L5 have been omitted (see Fig. 3). L6: virtual decision making loop. L7: mental model of computer model loop. L8: virtual mental model/perception loop. L9-L13: virtual system/real system loops.