

Report No. 6

January 7, 1998

## **Mental Models Concepts for System Dynamics Research**

James K. Doyle<sup>1</sup>  
David N. Ford<sup>2</sup>

System Dynamics Review, in press.

---

<sup>1</sup> Department of Social Science and Policy Studies, 100 Institute Rd., Worcester Polytechnic Institute, Worcester, MA 01609. Email: [doyle@wpi.edu](mailto:doyle@wpi.edu).

<sup>2</sup> Department of Information Science, University of Bergen, N-5020 Bergen, Norway. E-mail: [david.ford@ifi.uib.no](mailto:david.ford@ifi.uib.no)

## ABSTRACT

Although "mental models" are of central importance to system dynamics research and practice, the field has yet to develop an unambiguous and agreed upon definition of them. To begin to address this problem, existing definitions and descriptions of mental models in system dynamics and several literatures related to cognitive science were reviewed and compared. Available definitions were found to be overly brief, general, and vague, and different authors were found to markedly disagree on the basic characteristics of mental models. Based on this review, we concluded that in order to reduce the amount of confusion in the literature, the mental models concept should be "unbundled" and the term "mental models" should be used more narrowly. To initiate a dialogue through which the system dynamics community might achieve a shared understanding of mental models, we proposed a new definition of "mental models of dynamic systems" accompanied by an extended annotation that explains the definitional choices made and suggests terms for other cognitive structures left undefined by narrowing the mental model concept. Suggestions for future research that could improve the field's ability to further define mental models are discussed.

A difficulty for those who want to understand or to appraise mental models is that their proponents seem to have somewhat different views.

Rips (1986, p. 259)

Although the phrase "mental models" is ubiquitous in the literature, there are surprisingly few explicit definitions of them.

Rouse and Morris (1986, p. 349)

## **Introduction**

The concept of "mental models" has been vitally important to the field of system dynamics since its inception. Information about the structure and relationships in dynamic systems gleaned from mental models, for example, are what allow system dynamic computer models to be constructed in the absence of written and numerical data (Forrester, 1961). System dynamics researchers have in fact devoted a substantial portion of their research effort to developing a wide variety of techniques and procedures for eliciting, representing, and mapping mental models to aid model building (see Hall et al., 1994; Vennix et al., 1994). And, the goal of most educational interventions based on systems thinking, management flight simulators, or system dynamics model building is to change or improve mental models in order to improve the quality of dynamic decisions.

Mental models are thus the stock in trade of research and practice in system dynamics: they are the "product" that modelers take from students and clients, disassemble, reconfigure, add to, subtract from, and return with value added. An understanding of exactly what mental models are, what properties and characteristics they have, and how they influence and are influenced by learning and decision making is essential for such an enterprise to succeed. Given their importance to the field, one might expect mental model concepts to be as clearly defined and universally understood as such

other centrally important system dynamics concepts as stocks, flows, and feedback. But as we will show, this is not the case: explicit definitions of mental models are in fact quite rare. Those definitions that are available are typically general and vague, and the definitions offered by different authors often markedly disagree.

The ambiguity and confusion resulting from the lack of a clear, specific, and mutually agreed upon conceptual definition of the term "mental models" has several important consequences for the field of system dynamics. In the absence of consensus different researchers and practitioners develop and apply idiosyncratic conceptions of mental models. This hinders communication among researchers since marked differences of opinion hidden under the same generic name go unnoticed and unexamined. Since research groups employ, to some degree, different techniques for eliciting and mapping mental models based on their unique definitions, it is difficult for research results to cumulate across research programs. In addition, the various definitions of mental models used in the field of system dynamics diverge from the way in which the term is used outside the field, interfering with the ability of system dynamicists to share their insights, techniques, and research results with researchers from other disciplines. Finally, the field's willingness to accept the current level of ambiguity has likely discouraged researchers from developing more sophisticated definitions and descriptions of mental models, which in turn makes the process of incorporating mental models into computer simulation models less reliable.

The primary goals of the present paper are (a) to reduce the amount of confusion about mental models by organizing the available literature and "unbundling" the mental models concept into a set of more specific interrelated concepts and (b) to initiate a dialogue within the field of system dynamics that will lead to a clearer, more sophisticated, and ultimately shared understanding of the concept of mental models. Toward these ends, we will compare and analyze existing definitions of mental models within the system dynamics and related systems thinking literatures and attempt to identify areas of consensus; compare these definitions with those employed by other fields interested in the study of mental models; identify the shortcomings of existing definitions; and, following established criteria, propose a more specific and comprehensive conceptual definition of "mental models" for use in system dynamics

research. The present work will not attempt to present a comprehensive theory of the role of mental models in dynamic decision making (see Richardson et al. (1994) for a recent example of such an effort) but will instead focus on defining mental models concepts in order to facilitate future theory-building, research, and applications.

## **Defining Mental Models: A Review of the Literature**

The literature that touches on the concept of mental models, both within and particularly outside the field of system dynamics, is truly vast. We have therefore limited our review to sources for which mental models serve as the main focus of the work and sources which, in the course of addressing other topics, explicitly attempt to define mental models or describe their characteristics. In addition, we focus on literature in which the term "mental model," rather than related terms such as "cognitive map" or "schema," is commonly and widely used. This limitation in scope was necessary since available definitions are rarely explicit enough to allow reviewers to determine with confidence whether or not authors who use different terms are referring to the same or different concepts.

### *System Dynamics*

In system dynamics the mental model concept dates back to Industrial Dynamics, where Forrester (1961) offers the following introduction to mental models:

A mental image or a verbal description in English can form a model of corporate organization and its processes. The manager deals continuously with these mental and verbal models of the corporation. They are not the real corporation. They are not necessarily correct. They are models to substitute in our thinking for the real system that is represented (p. 49).

Forrester (1971) elaborates on these ideas, providing the following description of mental models:

The mental image of the world around us that we carry in our heads is a model. One does not have a city or a government, or a country in his head. He has only selected concepts and relationships, which he uses to represent the real system. (p. 213)

He goes on to describe some of the characteristics of mental models:

The mental model is fuzzy. It is incomplete. It is imprecisely stated. Furthermore, within one individual, a mental model changes with time and even during the flow of a single conversation. (p. 213)

The main shortcoming of mental models, in Forrester's opinion, is that their dynamic consequences cannot be simulated mentally, providing the primary rationale for using system dynamics modeling to support dynamic decision making. Forrester (1994) describes the limitations of mental models as follows:

The number of variables [people] can in fact properly relate to one another is very limited. The intuitive judgment of even a skilled investigator is quite unreliable in anticipating the dynamic behavior of a simple information-feedback system of perhaps five or six variables. (p. 60)

Several system dynamics researchers have subsequently added to this list of characteristics of mental models. Richardson and Pugh (1981), for example, describe mental models as "fuzzy and implicit," containing "rich, intuitive detail," highly adaptable, and "unable to handle complexity." Meadows et al. (1992) suggest that mental models are "extremely simple compared to reality" and "mostly wrong." According to Vazquez et al. (1996), mental models are "not fixed," "not simple," and "contain rich information." Sterman (1994) describes mental models of systems as being "vastly simplified compared to the complexity of the systems themselves" and "dynamically deficient" in the following ways:

. . . people generally adopt an event-based, open-loop view of causality, ignore feedback processes, fail to appreciate time delays between action and response and in the reporting of information, and are insensitive to nonlinearities that may alter the strengths of different feedback loops as a system evolves . . . (p. 305)

Many of the above-described limitations of mental models have been confirmed by controlled experimental research (see, e.g., Dörner, 1980; Sterman, 1989a, 1989b; Brehmer, 1992; Kleinmutz, 1993).

Thus, system dynamics researchers generally agree about the primary shortcomings of mental models. However, there is much less agreement on precisely what mental models are. <sup>1</sup> In fact few system dynamics authors attempt to provide an explicit definition of what they mean by the term "mental models." The following are some representative examples of available definitions:

Each person carries in his head a mental model, an abstraction of all his perceptions and experiences in the world, which he uses to guide his decisions . . . [mental models are] intuitive generalizations from observations of real-world events. (Meadows et al., 1974, pp. 4-5)

mental models . . . contain the ideas, opinions, assumptions, etc. with respect to a policy problem and related issues (Vennix, 1990, p. 16)

"Mental models" are deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action. Very often, we are not consciously aware of our mental models or the effects they have on our behavior (Senge, 1990, p. 8)

It is useful to think of mental models as a dynamic pattern of connections comprising a core network of "familiar" facts and concepts, and a vast

matrix of potential connections that are stimulated by thinking and by the flow of conversation. (Morecroft, 1994, p. 7)

. . . mental models are multifaceted, including distinguishable submodels focused on ends (goals), means (strategies, tactics, policy levers) and connections between them (the means/end model). (Richardson et al., 1994)

In system dynamics, the term mental model stresses the implicit causal maps of a system we hold, our beliefs about the network of causes and effects that describe how a system operates, the boundary of the model (the exogenous variables) and the time horizon we consider relevant -- our framing or articulation of a problem. (Sterman, 1994, p. 294)

. . . mental models are some sort of psychological construction with an intended representational content. Mental models . . . are usually expressed by a set of sentences in ordinary language, describing both the interactions among the elements within the system and their external influences. (Vazquez et al., 1996, p. 25)

These statements and others available in the system dynamics literature show evidence of substantial disagreement on some centrally important questions about how the term "mental models" should be used. For example, are mental models deeply ingrained and relatively stable (Senge, 1990) or fleeting and unstable (Forrester, 1971)? Are they "extremely simple" (Meadows et al., 1992), "not simple" (Vazquez et al., 1996), or "ranging from simple . . . to complex" (Senge, 1990)? Are they "images" (Forrester, 1971; Senge, 1990), "facts and concepts" (Morecroft, 1994), or "beliefs about . . . causes and effects" (Sterman, 1994)? Should a single belief be considered a mental model (Schley and Laur, 1996) or should the term instead refer to "sets of interacting beliefs" (Ford et al., 1993)? Are mental models an "abstraction of all . . . perceptions and experiences in the world" (Meadows et al., 1974) or some subset of these abstractions that are applied to a particular problem (Vennix, 1990)? Should the term "mental



models" refer to one particular type of cognitive structure (e.g., Morecroft, 1994) or to a set of different types of cognitive structures (Richardson et al., 1994)? Although it is often tempting for researchers to conclude that mental models are "all of the above," such an approach renders the term so all inclusive as to be meaningless.

The contents of available definitions in system dynamics vary widely; however, their character generally does not. First, all of the definitions lack coverage of issues critical for defining mental models: none of the definitions answer all or even most of the above-stated questions. Second, the majority of the available definitions are brief and somewhat vague. Brevity is not necessarily bad, particularly given how much is still unknown about the structure, content, and function of mental models. Vagueness, on the other hand, should be avoided: definitions that describe mental models using terms that are equally as ill-defined (e.g., images, assumptions, generalizations, perceptions) hinder the process of achieving consensus. Third, in most of the definitions the term "mental models" is used very generally to indicate any among a wide variety of quite different and distinct mental constructs. Very general terms, of course, can be quite useful if they serve to organize a set of more specific, subordinate terms. However, that is not the case here: of the references reviewed, for example, only Richardson et al. (1994) attempts to define subcategories of mental models.<sup>2</sup> Finally, references to cognitive psychology and other fields with long histories of research on mental models are relatively rare.

In short, in system dynamics the term "mental models" is currently ill-defined and means too many different things to different people to be useful in research and practice. To inform the development of a more specific definition, we turn now to a brief review of the long and varied history of the mental models concept in psychology, cognitive science, and related fields.

### *Psychology, Cognitive Science, and Related Fields*

In psychology the mental model concept can be traced back to Craik's (1943) book The Nature of Explanation. Craik proposed that people construct internal symbolic representations or models of external events. He defined the term "model" as:

. . . any physical or chemical system which has a similar relation-structure to that of the process it imitates . . . it is a physical working model which works in the same way as the process it parallels (p. 51)

Thus for Craik human reasoning involves mental simulation of dynamic internal representations of the external world.

Over time, the view that human judgment, reasoning, and problem solving is based on the manipulation of complex mental representations that intervene between stimuli and behavioral responses has become the dominant view in psychology and is the very foundation of the disciplines of cognitive psychology and cognitive science (Gardner, 1985; Hunt, 1989). However, the study of mental models of systems of the sort described by Craik and of interest to system dynamics researchers has, for a variety of reasons, not been a major focus of research in psychology. For example:

1. A substantial portion of the research effort in cognitive psychology has been devoted to developing and testing theories of how knowledge is represented in the mind in long-term memory (Atkinson and Shiffrin, 1968). The cognitive structures proposed, for example, conceptual networks (Collins and Loftus, 1975), propositional networks (Anderson, 1983), and connectionist networks (McClelland and Rumelhart, 1986), are general, very large, and universal, and their form does not model the structure of the external world. This view is quite different from the mental models approach, which supposes the existence of different, specialized cognitive structures in long-term memory for different tasks and situations.
2. Other researchers in cognitive psychology have proposed and studied a large number of specialized cognitive structures that differ widely in size, form, function, and character. They also vary in duration: some are thought to be stored permanently in long-term memory, whereas others are thought to be constructed and stored only temporarily in short-term or working memory (Baddeley, 1986). According to this view, people have different cognitive structures that serve different purposes: e.g., scripts (Schank and Abelson, 1977; Bower and Morrow, 1990) for understanding routine

activities, situation models for understanding text (Van Dijk and Kintsch, 1983), causal scenarios (Read, 1987; Tversky and Kahneman, 1973; Kahneman and Tversky, 1982) or stories (Pennington and Hastie, 1991) to aid in making causal attributions or judging likelihood, scenarios (Jungermann and Thuring, 1987) to enable judgmental forecasting, schemas (Fiske and Taylor, 1991) for perceiving and remembering information about people, imagery (Kosslyn, 1990) that allows objects not physically present to be scanned and mentally manipulated, and problem representations (Greeno, 1977) to help structure and manipulate information during problem solving. From this perspective mental models of systems are just one among a large family of cognitive structures, and thus they are not considered to be so centrally important to human cognition as they are described in the system dynamics literature.

3. Psychologists generally view the detailed study of mental models as a difficult and complex, if not impossible, task. According to this view, mental models are continually changing and efforts to elicit, measure, or map them can themselves induce changes in mental models.<sup>3</sup> When people are asked to report their mental models, they may fail to report them accurately for any of several reasons: e.g., they may simply not be aware of the contents of their mental models; they may feel compelled to invent explanations and answers on the spot that did not exist until the question was asked; or they may deliberately or unconsciously change their answers to correspond to the answers they think the researcher wants to hear (Norman, 1983). The methodologies that cognitive psychologists believe are necessary to address these problems and to minimize measurement error (see, e.g., Doyle et al. (1997)) are labor-intensive, time-consuming, and expensive, and are therefore only rarely applied.

4. Mental models have a substantial degree of support in the cognitive psychology literature, but the topic still generates controversy. For example, Johnson-Laird's (1983, 1994) widely cited work on mental models in conditional reasoning and probabilistic thinking has been challenged by other researchers who suggest that human ability to reason deductively and inductively is best explained by supposing that people use abstract reasoning rules or principles either in place of (Rips, 1986; 1990; O'Brien et al.,

1994) or at least in addition to (Roberts, 1993) constructing and manipulating mental models. Galotti et al. (1986) have suggested that the extent to which people use mental models versus abstract rules varies with experience, with experts relying more on abstract rules than novices. Thus, in contrast to the field of system dynamics, in cognitive psychology the jury is still out on whether or not human reasoning is primarily based on the manipulation of mental models.

5. System dynamics researchers have examined the evidence suggesting that people cannot mentally simulate any but the simplest of mental models without error and have responded by developing computer simulation software designed to improve dynamic decision making. Cognitive psychologists have drawn the same conclusions from the evidence, but have generally responded in a quite different way -- by assuming that manipulation of mental models is not a plausible explanation for dynamic decision making in complex environments and developing descriptive models of what people do instead. This has led researchers to focus less on cognitive structures and more on the cognitive processes, particularly simplifying heuristic rules (Newell and Simon, 1972; Kahneman et al., 1982), through which information is mentally reduced and transformed in the face of complexity.

For these reasons the mental models concept and the term "mental models" has historically been (and is currently still) widely used in only a handful of distinct research domains in and related to cognitive psychology and cognitive science. Most research on mental models has in fact taken place in interdisciplinary fields on the fringes of cognitive science and has been applied rather than theoretical in nature. Most prominent among these fields are deductive reasoning, human-machine and human-computer interaction, cognition and instruction, and risk perception and communication, so it is these fields which have been reviewed for statements related to characterizing and defining mental models. Virtually all of these fields agree with each other and with the system dynamics literature that the most noteworthy characteristics of mental models are various deficiencies that arise from bounded rationality (Simon, 1956) and the limitations of experience. Johnson-Laird's (1983) mental models theory was in fact proposed in

order to explain the errors people typically make when trying to answer even fairly simple logical syllogisms. The errors are thought to arise from difficulties in constructing multiple mental models due to cognitive limitations (see Johnson-Laird et al., 1989).

In the human-machine interaction field, Norman (1983) has described mental models as "incomplete," "unstable," and "unscientific." Like Forrester, Norman believes that people's ability to mentally simulate their mental models is "severely limited." Norman goes on to conclude that

. . . most people's understanding of the devices they interact with is surprisingly meager, imprecisely specified, and full of inconsistencies, gaps, and idiosyncratic quirks. (p. 8)

Other authors in the field of human-machine interaction have similarly observed that mental models of mechanical devices are typically oversimplified, inaccurate, and incomplete (see, e.g., Borgman, 1986; Moray, 1987; Williams et al., 1983).

Not surprisingly, since they generally study novices attempting to understand technical and scientific subjects, researchers in the area of cognition and instruction also report finding a wide variety of errors and omissions in people's mental models. Studies of mental models in the domain of physics have found, for example, that most people hold incorrect, "pre-Newtonian" mental models concerning the laws of motion (McCloskey, 1983a, 1983b) and draw incorrect analogies between how water flows in a river and how electricity flows in a wire when forming mental models of electrical circuits (Gentner and Gentner, 1983). Similarly, studies of novices learning to program and use computer software (see, e.g., Bayman and Mayer, 1983; Staggers and Norcio, 1993; and Janosky et al., 1986) have documented serious flaws in people's mental models of how computers work that hinder learning.

In the field of risk perception and communication, mental models are reported to be, if anything, more error-prone than in other fields. Studies in this field have shown, for example, that people's mental models of global warming tend to confuse ozone depletion with the greenhouse effect and weather with climate (Kempton, 1991; Bostrom et al., 1994; Read et al., 1994); their mental models of radon risk often include health

effects that have no basis in fact (Bostrom et al., 1993); their mental models of toxicology are typically insensitive to dose (Kraus et al., 1992); and their mental models of electric fields fail to take into account how quickly the strength of fields decreases over distance (Morgan et al., 1990)

Thus, as in the system dynamics literature, almost all of the researchers who study mental models in cognitive science and related fields agree on what the shortcomings of mental models are. However, again in parallel with the situation in system dynamics, explicit definitions are rare, and a review of representative definitions uncovers very little agreement either within or between literatures.<sup>4</sup>

Johnson-Laird's work on the role of mental models in deductive reasoning is probably the most widely cited work on mental models in any field, and is therefore an appropriate place to begin a review of available definitions. According to Johnson-Laird (1989),

a mental model can be defined as a representation of a body of knowledge -- either long-term or short-term that meets the following conditions: 1. its structure corresponds to the structure of the situation that it represents. 2. It can consist of elements corresponding only to perceptible [capable of being perceived by the senses] entities, in which case it may be realized as an image, perceptual or imaginary. 3. Unlike other proposed forms of representation, it does not contain variables . . . In place of a variable . . . a model employs tokens [symbols that are fixed rather than capable of assuming alternate values or states] (p. 488).

Thus, for Johnson-Laird, a mental model is a sort of "mental diagram" that contains "mental images" which are similar to the images formed during perception and are spatially arrayed in a manner corresponding to their real-life counterparts. During deductive reasoning, Johnson-Laird postulates that inferences are made by "reading" information from the mental diagram that was not stated in the original premises used to construct it.

In the field of human-computer interaction, mental models have been variously described as "knowledge about the system, external influences, and control strategies" (Veldhuyzen and Stassen, 1977), special types of "schema" (Jagacinski and Miller, 1978), "mental representations of a system" (Young, 1983), knowledge about "how a device works in terms of its internal structures and processes" (Kieras and Bovair, 1984, p. 255), "organized structures consisting of objects and their relationships" (Staggers and Norcio, 1993, p. 590), and "abstract concepts that . . . represent a person's knowledge of a decision problem" (Coury et al., 1992, p. 673). In summarizing this literature more than a decade ago, Rouse and Morris (1986) stated that

definitions within the cognitive science community range from broad and intentionally amorphous generalizations to specific and occasionally esoteric constructs. (p. 350)

In response to this problem, they demurred from developing a conceptual definition and offered the following "functional definition" instead:

Mental models are the mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states. (Rouse and Morris, 1986, p. 351)

Available definitions in the field of cognition and instruction are also often functional definitions that largely avoid details. For example, Halford (1993) defines mental models as "representations that are active while solving a particular problem and that provide the workspace for inference and mental operations." Shih and Alessi (1993, p. 157) state that "by a mental model we mean a person's understanding of the environment. It can represent different states of the problem and the causal relationships among states." Vosniadou and Brewer (1994, p. 125) get somewhat more specific when they explain that they "use the term mental model to denote a particular kind of mental representation . . . [whose] structure is an analog to the states of the world that it

represents." However, perhaps because many researchers in this field are new to the study of mental models, the preferred definitions are typically more like Wild's (1996, p. 10) general statement that "a mental model is a mediating intervention between perception and action."

Because it is a relatively new field of study, the definitions of mental models found in the risk perception and communication literature tend to be simple statements that are often indistinguishable from the concept of knowledge in general. According to Fischhoff et al. (1993), for example, "The term mental model is often applied to intuitive theories that are elaborated well enough to generate predictions in diverse circumstances." Maharik and Fischhoff (1992) define mental models as "people's collection of beliefs (both true and false) about a certain topic." Jungermann et al. (1988) suggest that "a mental model is a mapping from a domain into a mental representation that contains the main characteristics of the domain." Occasionally, the shortcomings of mental models of risks become entangled with their definition, as, for example, when Atman et al. (1994) refer to mental models as "the pattern of knowledge gaps, overly general understandings, and outright misconceptions that can frustrate learning."

In characterizing the definitions of mental models offered in the above examples from fields related to cognitive science, it is clear that they suffer from the same problems as the definitions available in system dynamics. With few exceptions, the definitions are brief and vague. And, no single definition addresses more than a small subset of the issues raised by the entire set of definitions. There is one important point on which virtually all of the definitions offered in cognitive science fields agree, namely, the idea that the structure of mental models "mirrors" the perceived structure of the external system being modeled. Johnson-Laird (1983) refers to this feature of mental models as "the principle of structural identity." However, apart from agreement on this basic principle, disagreements on the nature of central, basic features of mental models are, as in system dynamics, easy to find. For example, are mental models composed of picture-like images (e.g., Johnson-Laird, 1983; Rouse and Morris, 1986; Jih and Reeves, 1992), declarative knowledge (e.g., Veldhuyzen and Stassen, 1977), concepts (Coury et al., 1992), or intuitive theories (Fischhoff et al., 1993)? Are they stable (Seel, 1995), or are they fleeting, being constructed and discarded as needed to solve problems (Johnson-



Laird, 1983; Vosniadou and Brewer, 1992)? Do mental models include only information about a system (Young, 1983) or do they also include information about external influences and decision strategies (Veldhuyzen and Stassen, 1977)? Are mental models accessible (i.e., are people aware of their mental models and able to report on their contents), or are they outside of conscious awareness and inaccessible as Van Heusden (1980), Whitfield and Jackson (1982), and Rouse and Morris (1986) have suggested? Do people have a single mental model of a system as assumed, for example, by researchers in risk perception, or can they have multiple alternate mental models of the same system as suggested by McCloskey (1983a), Clement (1983), DeKleer and Brown (1983), Williams et al. (1983) and Moray (1987)?

In short, as in system dynamics, these literatures have not been able to achieve consensus on a specific, unambiguous definition of mental models. System dynamics researchers looking to the cognitive science literature for answers on precisely what mental models are are likely to come away with more questions instead. However, only by addressing these questions can a more specific and useful definition of mental models for system dynamics be constructed.

### **Towards a Shared Definition of "Mental Models" in System Dynamics**

Like other research disciplines that have adopted the mental models concept, the field of system dynamics has developed its own definitions and methodologies largely in isolation from past or current work on mental models in other disciplines. Important distinctions described by other literatures have been largely ignored, resulting in definitions of mental models that are so general they serve to replace such overarching concepts as "psychology" or "cognition." At the same time, the shortcomings of these other literatures have also been ignored, and they are too often described as authoritative when they do not in actuality contain ideas about mental models that are more specific or detailed than those found in system dynamics. The problem is not that existing definitions in system dynamics are wrong; the problem is that, since they are overly general and incomplete, they are all partly right and at the same time they are all different.

We believe that in order to solve this problem, the mental models concept must be "unbundled." That is, its distinct, separable components must be identified and given separate names. The term "mental model" should be used more specifically to refer to only a small subset of the wide variety of mental phenomena to which it is currently associated. Which cognitive structures are given the name "mental model" and which are given alternate names is not critically important at this point; what is important is that a more precise and useful glossary of mental models concepts be developed and shared so that it is amenable to future review and revision by the system dynamics research community.

To begin this process, we propose that system dynamics researchers are primarily interested in specialized cognitive structures that are best described as "mental models of dynamic systems" (MMODS). We suggest that the term "mental model" in system dynamics should be understood to be an abbreviation of this longer term, and that the term "mental representations" (Gardner, 1985) rather than "mental models" should be used to indicate the variegated set of all types of proposed cognitive structures. In the following section of the paper, we propose and elaborate on a new definition of mental models of dynamic systems.

The definition offered below is a conceptual definition, that is, a definition that describes a concept using other concepts.<sup>5</sup> In constructing this definition, we have been mindful of the following features that improve the ability of a conceptual definition to enhance communication (Frankfort-Nachmias and Nachmias, 1992):

1. A definition must point out the unique attributes or qualities of whatever is defined. It must be inclusive of all cases it covers and exclusive of all cases not covered.
2. A definition should not be circular; that is, it must not contain any part of the thing being defined.
3. A definition should be stated positively.
4. A definition should use clear terms [on whose meaning different people agree]. (p. 32)

Our goal is to create a definition that, insofar as possible, corresponds with how the term is most typically used at present in system dynamics. For example, the reliance in system dynamics on stock-flow diagrams, causal loop diagrams, and other conceptually based representations to describe and communicate mental models makes clear the preference in the field for thinking of mental models as concept-based rather than image-based, and this preference is preserved in our proposed definition. Similarly, the central role assigned to mental models in system dynamics descriptions of dynamic decision making suggests that the majority of researchers in the field conceive of mental models as relatively enduring structures rather than temporary structures, and this preference is also reflected in the definition below.

We have attempted to create a definition that is as specific as possible, but not more: issues which have not yet been addressed by the research literature are omitted or are treated more generally than issues which have been more thoroughly studied. For example, we do not attempt to define the role of mental models in dynamic decision making since much of the empirical work necessary to confirm this role has yet to be conducted. Nor do we attempt to render judgment on the issue of whether people have a single mental model or (potentially) multiple mental models of a system. The preference in the system dynamics literature is clearly on the side of unitary models, but the alternative has not yet been considered by most researchers. Finally, since the literature is crowded with too many alternate terms already, we have whenever possible suggested the use of existing terms rather than creating new ones.

### **Mental Models of Dynamic Systems: An Annotated Definition**

We propose that a mental model of a dynamic system is a relatively enduring and accessible, but limited, internal conceptual representation of an external system whose structure maintains the perceived structure of that system. The following annotated version of this definition explains the definitional choices we have made and suggests the use of terms for other cognitive structures not covered by the definition of MMODS:

*A mental model of a dynamic system . . .*

Although some authors prefer to use "theory" in place of "model," we believe that the term "theory" suggests a degree of completeness and coherence that is often lacking in mental models. "Theory" might more appropriately be used to describe only the subset of mental models that have these additional properties.

The terms "beliefs," "set of beliefs," or "belief system" are also often used in place of "mental model." However, we agree with Norman (1983) that mental models can include "knowledge or beliefs that are thought to be of doubtful validity." From this perspective, the term "belief," which implies a fairly high a degree of confidence in one's knowledge, describes some but not all mental models.

*. . . is a relatively enduring . . .*

The stability of mental models is difficult to describe. Parts of mental models may be altered, deleted, or added on a time scale of minutes or seconds. Yet, a mental model considered as a whole, while continually changing in detail, may endure in memory in some form for years or decades. The phrase "relatively enduring" means that the term "mental model" should be reserved for cognitive structures that are stored in a potentially permanent state in long-term memory rather than structures that are stored only temporarily (on the order of seconds or minutes) in short-term or working memory. It also implies that mental models are structures that are "precompiled," that is, they are stored as a unit in memory rather than being constructed from smaller components during decision making.

Reserving the term "mental models" for relatively enduring structures by no means implies that less enduring structures are unimportant: in fact, recent research suggests that constructive processes play a crucial role in human decision making (Payne et al., 1992). For cognitive structures that are actively constructed (from existing mental representations in combination with new information) during decision making and held

only temporarily in working memory, we suggest the term "dynamic problem representation," after Greeno (1977).

*. . . and accessible, . . .*

Virtually all psychological theorists acknowledge that conscious thought is supported by a tremendous amount of mental information and activity that takes place "unconsciously" and cannot be described by people with any degree of reliability. In some cases unconscious mental representations and processes can have a significant effect on people's judgments without their being aware of it (see, e.g., Nisbett and Wilson, 1977; Begg et al., 1992). We suggest the term "implicit model" (see Rouse and Morris, 1986) for mental model-like structures that are outside conscious awareness, reserving the term "mental models" for cognitive structures that are relatively available to conscious introspection.

*. . . but limited, . . .*

We believe the term "mental model" should not be used to refer to knowledge in general or even to all knowledge that can be recalled from memory about a given system, but, as stated above, to a "precompiled" subunit of information held in memory. Such a mental model can vary significantly in size and complexity, just as a real system may vary in size and complexity. However, it is possible to specify upper and lower limits to this variation.

We propose that the upper limit on the size and complexity of a mental model is determined by limits on "bounded rationality" (Simon, 1956). To aid decision making, a mental model must be small enough to be implemented in short term memory, the capacity of which is generally considered to be seven plus or minus two "chunks" of information. This limit is flexible in the sense that the amount of information that can be organized meaningfully into a chunk can increase with experience and expertise; the maximum number of different chunks in short term memory, however, is apparently

unalterable (Miller, 1956; Simon, 1974). Along similar lines, Forrester (1994) suggests that mental models are typically no more complex than "a fourth-order differential equation."

A reasonable lower bound on the size of a mental model for system dynamics purposes is the minimum requirement for a "closed" system, that is, two variables and two causal relationships. Thus a single causal assertion such as "an increase in X causes an increase in Y" is not sufficiently complex to be called a "model" and should instead be referred to as an "assumption" or "belief," depending on the degree of confidence that is associated with it.

... *internal* ...

The term "internal" implies that a mental model is a cognitive phenomenon, that is, something that exists only in the mind. The products of efforts to "surface," "elicit," or "map" mental models, which are typically causal-loop or stock-flow diagrams on paper or in a computer program, should not be confused with the mental model itself due to the strong likelihood of measurement error and the possibility of bias. As discussed by Norman (1983), these externalized models are really the researcher's conceptualization of a subjects' mental model, and may be influenced by the researcher's own mental models of human cognition and behavior. To refer to external representations of mental models, we suggest the commonly used term "cognitive map" (see Axelrod, 1976; Eden, 1994).<sup>6</sup>

... *conceptual* ...

By conceptual we mean that mental models are based on concepts, ideas, or other language-like components. There is substantial evidence for the existence and use of mental imagery (see, e.g., Finke, 1989; Kosslyn, 1990), and some authors who study simple mental models (e.g., Johnson-Laird, 1983) or mental models of physical devices (e.g., Rouse and Morris, 1986) have proposed that such models are primarily image-based. However, the mental models of interest in system dynamics are typically more

complex and abstract social and/or economic models for which the creation of images is more difficult and therefore more conceptual representations likely predominate. We suggest that the term "mental image," not "mental model," be used to refer to picture-like representations.

*. . . representation . . .*

The term "representation" means that mental models are cognitive structures, which should be distinguished from cognitive processes. Cognitive structures store information; cognitive processes are the mental operations that transform, elaborate, and reduce this information during decision making or problem solving.

*. . . of an external system . . .*

This phrase suggests that for a mental representation to be called a mental model, it must have an external referent, that is, it must, in whole or in part, represent something external to an individual's mind. Thus, we are proposing that such important mental constructs as attitudes (Eagly and Chaiken, 1993) and goals (Schank and Abelson, 1977), which exist only in minds, should not be called "mental models" but should simply be referred to as "attitudes" and "goals." Of course, mental models can and often do contain conceptual "nodes" or "tokens" that represent the attitudes and goals of the self or others – however, these nodes should not be confused with the attitudes and goals themselves, which can exist apart from MMODS or which may be represented in multiple different mental models.

In contrast to Sterman (1994), we do not propose that MMODS include exogenous variables and the relevant time horizon. We suggest that these concepts, which may vary from one problem solving situation to another, are features of dynamic problem representations rather than MMODS.

It should be noted that it is possible to have a mental model that refers only to one's own internal cognitive structures. We suggest that this exception to the requirement for an external referent be referred to as a "metamodel," after the cognitive psychology term "metacognition" (see Nelson, 1992).

*. . . whose structure . . .*

The word "structure" implies that mental models include not merely knowledge but detailed information about how knowledge is organized and inter-connected. We suggest that the term "knowledge" be used in place of "mental model" when no assumptions about how information is organized are implied.

The precise nature of the conceptual nodes and links that form a mental model of a dynamic system has yet to be confirmed by empirical research. For example, are the nodes highly abstract concepts that summarize many experiences, or are they more concrete tokens that represent specific examples of systems? Is there only one type of link or are there several types of links? Do the links vary in strength or are they uniform? It is in fact likely that the nature of the nodes and links in a mental model is variable and changes as people gain experience and develop expertise in a certain field. Several authors, for example, have suggested that the degree to which mental models are abstract versus representational varies with expertise (see, e.g., Larkin, 1983; DiSessa, 1983; Greeno, 1983; De Jong and Ferguson-Hessler, 1986).

*. . . maintains the perceived structure of that system.*

The word "perceived" is very important. The evidence presented in the literature review above from a variety of research disciplines that people's mental models are prone to errors and omissions is overwhelming. Mental models attempt to preserve the structure in the external system, but only partially succeed.



## Conclusion

A review of existing descriptions and definitions of "mental models" in a variety of literatures, including system dynamics and several research disciplines associated with cognitive science, has uncovered a common set of problems: available definitions are typically brief, overly general, and vague, and different authors offer definitions that markedly disagree on centrally important features of mental models. We believe that the lack of a clearly specified, comprehensive, and agreed upon definition of mental models has hindered communication between researchers and dramatically slowed the progress of research aimed at describing, understanding, and improving mental models of dynamic systems.

As a step toward the resolution of this problem, we have suggested that the term "mental models" be used much more narrowly than it is at present, and have offered a new conceptual definition of "mental models of dynamic systems." We have also provided an extended annotated version of this definition to clarify its meaning and to suggest a glossary of terms to refer to cognitive structures that have been excluded from our definition of mental models. According to Frankfort-Nachmias and Nachmias (1992), the test of a definition's quality is its utility:

Conceptual definitions are neither true nor false . . . Conceptual definitions are either useful for communication and research, or they are not. (p. 31)

The annotated definition we have described is more specific and detailed than previously existing definitions, and it should therefore be less likely to be misinterpreted or miscommunicated. It is also more comprehensive, identifying and taking a stand on the entire set of important questions raised by the literature on mental models, for example: Are mental models stable or unstable cognitive structures? Are they accessible to introspection or outside conscious awareness? Are they limited in size and complexity? Are they conceptual or image-based? Are they structurally analogous to their external referents? Are they abstract or representational? The definition should therefore help

researchers identify gaps in the definitions they use and provide a mutually agreeable basis for the comparison of alternate definitions.

We make no claim that the conceptual definition we have described is correct or complete, only that it can serve as a more useful starting point for further review and debate. We also believe that it would be premature to make claims of correctness or completeness for any definition of mental models at this time. The majority of research on mental models in all of the disciplines that employ the concept has been applied, and surprisingly little attention has been paid to theoretical concerns or to basic research. In particular, basic research is needed to establish the stability of mental models and how they interact with less enduring cognitive structures, the limitations to people's access to their own mental models, the precise nature of the conceptual nodes and links which form mental models, how the nature and role of mental models changes with experience and expertise, and the extent to which mental models (as opposed heuristic rules) are employed naturally and spontaneously in dynamic decision making. Theoretical work that places mental models in the context of a more elaborate cognitive system, such as that described by Richardson et al. (1994), should be encouraged and followed up by empirical research that puts assumptions to the test. Collaboration between system dynamics researchers, who have developed appropriate tools and techniques for examining and describing complexity, and cognitive psychologists, who have developed a large body of knowledge relating to human cognition and established rigorous empirical techniques for studying the mind, may be particularly important for such a research effort to succeed (see Doyle, 1997).

Conducting the basic and theoretical research necessary to further define mental models will not be easy. Describing mental models is inherently difficult, since they are not directly observable and can change during procedures designed to assess them. Compounding matters is the sheer complexity of the human mind and brain, which has been described as "the most complex structure in the known universe" (Staff, 1992). However, we believe that this difficult work is necessary for the field of system dynamics to fulfill its goals of improving mental models and the quality of the dynamic decisions based upon them. The system dynamics community must achieve the same level of

sophistication in how it thinks and talks about human cognitive systems that it has reached in its research into a wide variety of other complex systems.

### **Biographical Information**

James K. Doyle is an assistant professor of psychology in the Department of Social Science and Policy Studies at Worcester Polytechnic Institute. He holds a Ph. D. in Social Psychology from the University of Colorado at Boulder, where he conducted research at the Center for Research on Judgment and Policy, Institute of Cognitive Science. His research interests include mental models theory and methodology, cognitive processes in dynamic decision making, and risk perception and communication.

David N. Ford is an associate professor in the system dynamics program at the University of Bergen in Norway. He earned his Ph. D. from the Massachusetts Institute of Technology, where he conducted research on the dynamics of development processes. His current research interests include product development management, coordination and policy development..

### **References**

- Alexander, C. (1964). Notes on the Synthesis of Form. Cambridge, MA: Harvard Univ. Press.
- Anderson, J. R. (1983). The Architecture of Cognition. Cambridge, MA: Harvard Univ. Press.
- Argyris, C. (1982). Reasoning, Learning, and Action: Individual and Organizational. San Francisco: Jossey-Bass.
- Atkinson, R. C., and Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds.), The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2. New York: Academic Press.

- Atman, C. J., Bostrom, A., Fischhoff, B., and Morgan, M. G. (1994). Designing risk communications: Completing and correcting mental models of hazardous processes, Part I. Risk Analysis, 14(5), 779-788.
- Axelrod, R., Ed. (1976). The Structure of Decision: The Cognitive Maps of Political Elites. Princeton, NJ: Princeton Univ. Press.
- Baddeley, A. D. (1986). Working Memory. Oxford: Clarendon Press.
- Bayman, P., and Mayer, R. E. (1983). Diagnosis of beginning programmers' misconceptions of BASIC programming statements. Communications of the ACM, 26, 519-521.
- Begg, I., Anas, A., and Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. Journal of Experimental Psychology: General, 121, 446-458.
- Borgman, C. L. (1986). The user's mental model of an information retrieval system: An experiment on a prototype online catalog. International Journal of Man-Machine Studies, 24, 47-64.
- Bostrom, A., Fischhoff, B., and Morgan, M. G. (1993). Characterizing mental models of hazardous processes: A methodology and an application to radon. Journal of Social Issues, 48(4), 85-100.
- Bostrom, A., Morgan, M. G., and Read, D. (1994). What do people know about global climate change? Part 1: Mental models. Risk Analysis, 14(6), 959-970.
- Bower, G. H., and Morrow, D. G. (1990). Mental models in narrative comprehension. Science, 247, 44-48.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. Acta Psychologica, 81, 211-241.
- Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In D. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 325-340. Hillsdale, NJ: Erlbaum.
- Collins, A. M., and Loftus, E. M. (1975). A spreading-activation theory of semantic memory. Psychological Review, 82, 407-428.

- Coury, B. G., Weiland, M. Z., and Cuolock-Knopp, V. G. (1992). Probing the mental models of system state categories with multidimensional scaling. International Journal of Man-Machine Studies, 36, 673-696.
- Craik, K. (1943). The Nature of Explanation. Cambridge: Cambridge Univ. Press.
- De Jong, T., and Ferguson-Hessler, M. G. M. (1986). Cognitive structures of good and poor novice problem solvers in physics. Journal of Educational Psychology, 78, 279-288.
- de Kleer, J., and Brown, J. S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 155-190. Hillsdale, NJ: Erlbaum.
- DiSessa, A. A. (1983). Phenomenology and the evolution of intuition. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 15-34. Hillsdale, NJ: Erlbaum.
- Dörner, D. (1980). On the difficulties people have in dealing with complexity. Simulation and Games, 11(1), 87-106.
- Doyle, J. K. (1997). The cognitive psychology of systems thinking. System Dynamics Review, 13(3), 253-265.
- Doyle, J. K., Radzicki, M. J., and Trees, W. S. (1997). Measuring change in mental models of dynamic systems: An exploratory study. Unpublished manuscript, Dept. of Social Science and Policy Studies, Worcester Polytechnic Institute, Worcester, MA.
- Eagly, A. H., and Chaiken, S. (1993). The Psychology of Attitudes. Fort Worth, TX: Harcourt, Brace, Jovanovich.
- Eden, C. (1994). Cognitive mapping and problem structuring for system dynamics model building. System Dynamics Review, 10(2/3), 257-276.
- Finke, R. A. (1989). Principles of Mental Imagery. Cambridge, MA: MIT Press.
- Fischhoff, B., Bostrom, A., and Quadrel, M. J. (1993). Risk perception and communication. Annual Review of Public Health, 14, 183-203.
- Fiske, S. T., and Taylor, S. E. (1991). Social Cognition, 2nd ed. Reading, MA: Addison-Wesley.
- Ford, D. N., Hou, A., and Seville, D. (1993). An Exploration of Systems Product Development at Gadget, Inc. Technical Report D-4460, System Dynamics

- Group, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA.
- Forrester, J. (1961). Industrial Dynamics. Cambridge, MA: Productivity Press.
- Forrester, J. (1971). Counterintuitive behavior of social systems. In Collected Papers of J. W. Forrester, pp. 211-244. Cambridge, MA: Wright-Allen Press, Inc.
- Forrester, J. (1994). Policies, decisions, and information sources for modeling. In Morecroft, J. D. W., and Sterman, J. D. (Eds.), Modeling for Learning Organizations, pp. 51-84 Portland, OR: Productivity Press.
- Frankfort-Nachmias, C., and Nachmias, D. (1992). Research Methods in the Social Sciences, 4th ed. London: Edward Arnold.
- Galotti, K. M., Baron, J., and Sabini, J. P. (1986). Individual differences in syllogistic reasoning: Deduction rules or mental models? Journal of Experimental Psychology: General, 115, 16-25.
- Gardner, H. (1985). The Mind's New Science: A History of the Cognitive Revolution. New York: Basic Books.
- Gentner, D., and Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 99-129. Hillsdale, NJ: Erlbaum.
- Gentner, D., and Stevens, A. L., Eds. (1983). Mental models. Hillsdale, NJ: Erlbaum.
- Greeno, J. G. (1977). Process of understanding in problem solving. In N. J. Castellan, Jr., D. B. Pisoni, and G. R. Potts (Eds.), Cognitive Theory, Vol. 2, pp. 43-84. Hillsdale, NJ: Erlbaum.
- Greeno, J. G. (1983). Conceptual entities. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 227-252. Hillsdale, NJ: Erlbaum.
- Halford, G. S. (1993). Children's Understanding: The Development of Mental Models. Hillsdale, NJ: Erlbaum.
- Hall, R. I., Aitchison, P. W., and Kocay, W. L. (1994). Causal policy maps of managers: Formal methods for elicitation and analysis. System Dynamics Review, 10(4), 337-360.
- Hoogerwerf, A. (1984). Beleid berust op veronderstellingen: De beleidstheorie. Acta Politica, 4, 493-531.

- Howard, R. A. (1989). Knowledge maps. Management Science, 35(8), 903-922.
- Hunt, E. (1989). Cognitive science: Definition, status, and questions. Annual Review of Psychology, 40, 603-629.
- Jagacinski, R. J., and Miller, R. A. (1978). Describing the human operator's internal model of a dynamic system. Human Factors, 20, 425-433.
- Janosky, B., Smith, P. J., and Hildreth, C. (1986). Online library catalog systems: An analysis of user errors. International Journal of Man-Machine Studies, 28, 643-670.
- Jih, H. J., and Reeves, T. (1992). Mental models: A research focus for interactive learning systems. Educational Technology Research and Development, 40(3), 39-53.
- Johnson-Laird, P. (1983). Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness. Cambridge, MA: Harvard Univ. Press.
- Johnson-Laird, P. (1989). Mental models. In M. I. Posner (Ed.), Foundations of Cognitive Science, pp. 469-499. Cambridge, MA: MIT Press.
- Johnson-Laird, P., Byrne, R. M. J., and Tabossi, P. (1989). Reasoning by model: The case of multiple quantification. Psychological Review, 96, 658-673.
- Johnson-Laird, P. (1994). Mental models and probabilistic thinking. Cognition, 50, 189-209.
- Jungermann, H., and Thuring, M. (1987). The use of mental models for generating scenarios. In G. Wright and P. Ayton (Eds.), Judgmental Forecasting. New York: Wiley.
- Jungermann, H., Schutz, H., and Thuring, M. (1988). Mental models in risk assessment: Informing people about drugs. Risk Analysis, 8(1), 147-155.
- Kahneman, D., and Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), Judgment under Uncertainty: Heuristics and Biases, pp. 201-208. New York: Cambridge Univ. Press.
- Kahneman, D., Slovic, P., and Tversky, A., Eds. (1982). Judgment under Uncertainty: Heuristics and Biases. New York: Cambridge Univ. Press.
- Kempton, W. (1986). Two theories of home heat control. Cognitive Science, 10, 75-90.

- Kempton, W. (1991). Public understanding of global warming. Society and Natural Resources, 4, 331-345.
- Kieras, D. E., and Bovair, S. (1984). The role of a mental model in learning to operate a device. Cognitive Science, 8, 255-273.
- Kleinmutz, D. N. (1993). Information processing and misperceptions of the implications of feedback in dynamic decision making. System Dynamics Review, 9(3), 223-237.
- Kosslyn, S. M. (1990). Mental imagery. In D. N. Osherson, S. M. Kosslyn, and J. M. Hollerback (Eds.), Visual Cognition and Action: An Invitation to Cognitive Science, Vol. 2, pp. 73-97. Cambridge, MA: MIT Press.
- Kraus, N., Malmfors, T., and Slovic, P. (1992). Intuitive toxicology: Expert and lay judgements of chemical risks. Risk Analysis, 12(2), 215-232.
- Lakoff, G. (1987). Cognitive models and prototype theory. In U. Neisser (Ed.), Concepts and Conceptual Development. Cambridge: Cambridge Univ. Press.
- Larkin, J. H. (1983). The role of problem representation in physics. In D. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 75-98. Hillsdale, NJ: Erlbaum.
- Maharik, M., and Fischhoff, B. (1992). The risks of using nuclear energy sources in space: Some lay activists' perceptions. Risk Analysis, 12(3), 383-392.
- McClelland, J. L., and Rumelhart, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland and D. E. Rumelhart (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 2, Psychological and Biological Models. Cambridge, MA: MIT Press.
- McCloskey, M. (1983a). Naive theories of motion. In D. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 299-324. Hillsdale, NJ: Erlbaum.
- McCloskey, M. (1983b). Intuitive physics. Scientific American, 248(4), 122-130.
- Meadows, D. L., Behrens, W. W., III, Meadows, D. H., Naill, R. F., and Zahn, E. K. O. (1974). Dynamics of Growth in a Finite World. Cambridge, MA: Wright-Allen Press.
- Meadows, D. H., Meadows, D. L., and Randers, J. (1992). Beyond the Limits: Confronting Global Collapse, Envisioning a Sustainable Future. Post Mills, VT: Chelsea Green Publishing Co.



- Means, M. L., and Voss, J. F. (1985). Star Wars: A developmental study of expert and novice knowledge structures. Journal of Memory and Language, 24, 746-757.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.
- Moray, N. (1987). Intelligent aids, mental models and the theory of machines. International Journal of Man-Machine Studies, 27, 619-629.
- Morecroft, J. (1994). Executive knowledge, models, and learning. In J. Morecroft and J. Serman (Eds.), Modeling for Learning Organizations, pp. 3-28. Portland, OR: Productivity Press.
- Morgan, M. G., Florig, H. K., Nair, I., Cortes, C., and Marsh, K. (1990). Lay understanding of low-frequency electric and magnetic fields. Bioelectromagnetics, 11(4), 313.
- Morris, N. M., and Rouse, W. B. (1985). The effects of type of knowledge upon human problem solving in a process control task. IEEE Transactions on Systems, Man, and Cybernetics, 15, 694-707.
- Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), Concepts and Conceptual Development. Cambridge: Cambridge Univ. Press.
- Nelson, T. O., Ed. (1992). Metacognition: Core Readings. Boston, MA: Allyn & Bacon.
- Newell, A., and Simon, H. A. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R., and Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.
- Norman, D. A. (1983). Some observations on mental models. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 7-14. Hillsdale, NJ: Erlbaum.
- O'Brien, D. P., Braine, M. D. S., and Yang, Y. (1994). Propositional reasoning by mental models? Simple to refute in principle and practice. Psychological Bulletin, 101(4), 711-724.
- Ormrod, J. E., Ormrod, R. K., Wagner, E. D., and McCallin, R. C. (1988). Reconceptualizing map learning. American Journal of Psychology, 101, 425-433.
- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1992). Behavioral decision research:

- A constructive processing perspective. Annual Review of Psychology, 43, 87-131.
- Pennington, N. (1985). Stimulus Structures and Mental Representations in Expert Comprehension of Computer Programs (Tech. Rep. No. 2-ONR). Chicago, IL: Univ. of Chicago, Graduate School of Business.
- Pennington, N., and Hastie, R. (1991). A cognitive theory of juror decision making: The story model. Cardozo Law Review, 13(2/3), 519-558.
- Read, S. J. (1987). Constructing causal scenarios: A knowledge structure approach to causal reasoning. Journal of Personality and Social Psychology, 52(2), 288-302.
- Read, D., Bostrom, A., and Smuts, T. (1994). What do people know about global climate change? Part 2: Survey studies of educated laypeople. Risk Analysis, 14(6), 971-982.
- Richardson, G. P., and Pugh, A., III (1981). Introduction to System Dynamics Modeling with DYNAMO. Cambridge, MA: The MIT Press.
- Richardson, G. P., Andersen, D. F., Maxwell, T. A., and Stewart, T. R. (1994). Foundations of mental model research. Proceedings of the 12th International System Dynamics Conference, Stirling, Scotland, July 11-15.
- Rips, L. (1986). Mental muddles. In M. Brand and R. M. Harnish (Eds.), The Representation of Knowledge and Belief, pp. 258-286. Tucson, AZ: Univ. of Arizona Press.
- Rips, L. (1990). Reasoning. In M. R. Rosenzweig and L. W. Porter (Eds.), The Psychology of Human Thought. Cambridge: Cambridge Univ. Press.
- Roberts, M. J. (1993). Human reasoning: Deduction rules or mental models, or both? Quarterly Journal of Experimental Psychology, 46A(4), 569-589.
- 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.
- Schank, R. C., and Abelson, R. P. (1977). Scripts, Plans, Goals, and Understanding. Hillsdale, NJ: Erlbaum.
- Schley, S., and Laur, J. (1996). The power of mental models. The Systems Thinker. Cambridge, MA: Pegasus Communications, Inc.

- Seel, N. M. (1995). Mental models, knowledge transfer, and teaching strategies. Journal of Structural Learning, 12(3), 197-213.
- Senge, P. (1990). The Fifth Discipline: The Art and Practice of the Learning Organization. New York: Doubleday.
- Shavelson, R. J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. Journal of Educational Psychology, 63(3), 225-234.
- Shih, Y. F., and Alessi, S. M. (1993). Mental models and transfer or learning in computer programming. Journal of Research in Computing Education, 26(2), 154-175.
- Simon, H. A. (1956). Rational choice and the structure of the environment. Psychological Review, 63, 129-138.
- Simon, H. A. (1974). How big is a chunk? Science, 183, 482-488.
- Staff. (1992). The most complex structure in the known universe. Scientific American, 267(3), 4.
- Staggers, N., and Norcio, A. F. (1993). Mental models: Concepts for human-computer interaction research. International Journal of Man-Machine Studies, 38, 587-605.
- Sterman, J. D. (1989a). Misperceptions of feedback in dynamic decision making. Organizational Behavior and Human Decision Processes, 43(3), 301-335.
- Sterman, J. D. (1989b). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision-making experiment. Management Science, 35(3), 321-339.
- Sterman, J. D. (1994). Learning in and about complex systems. System Dynamics Review, 10(2/3), 291-330.
- Tolman, E. C. (1948). Cognitive maps in rats and men. Psychological Review, 55, 189-208.
- Tonn, B. E., Travis, C. B., Goeltz, R. T., and Phillippi, R. H. (1990). Knowledge-based representations of risk beliefs. Risk Analysis, 10(1), 169-184.
- Tversky, A., and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.
- Van Dijk, T. A., and Kintsch, W. (1983). Strategy of Discourse Comprehension. New York: Academic Press.

- Van Heusden, A. R. (1980). Human prediction of third-order autoregressive time series. IEEE Transactions on Systems, Man, and Cybernetics, 10, 38-43.
- Vazquez, M., Liz, M., and Aracil, J. (1996). Knowledge and reality: Some conceptual issues in system dynamics modeling. System Dynamics Review, 12(1), 21-37.
- Veldhuyzen, W., and Stassen, H. G. (1977). The internal model concept: An application to modeling human control of large ships. Human Factors, 19, 367-380.
- Vennix, J. A. M. (1990). Mental Models and Computer Models: Design and Evaluation of a Computer-Based Learning Environment for Policy-Making. Den Haag: CIP-Gegevens KoninklijkeBibliotheek.
- Vennix, J. A. M., Andersen, D. F., Richardson, G. P., and Rohrbaugh, J. (1994). Model building for group decision support: Issues and alternatives in knowledge elicitation. In Morecroft, J. D. W., and Serman, J. D. (Eds.), Modeling for Learning Organizations, pp. 29-49 Portland, OR: Productivity Press.
- Vosniadou, S., and Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. Cognitive Psychology, 24, 535-585.
- Vosniadou, S., and Brewer, W. F. (1994). Mental models of the day/night cycle. Cognitive Science, 18, 123-183.
- Whitfield, D., and Jackson, A. (1982). The air traffic controller's "picture" as an example of a mental model. In G. Johannsen and J. E. Rijnisdorp (Eds.), Analysis, Design, and Evaluation of Man-Machine Systems, pp. 45-52. London: Pergamon.
- Wild, M. (1996). Mental models and computer modeling. Journal of Computer Assisted Learning, 12, 10-21.
- Williams, M. D., Hollan, J. D., and Stevens, A. L. (1983). Human reasoning about a simple physical system. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 131-154. Hillsdale, NJ: Erlbaum.
- Young, R. M. (1983). Surrogates and mappings: Two kinds of conceptual models for interactive devices. In D. R. Gentner and A. L. Stevens (Eds.), Mental Models, pp. 32-52. Hillsdale, NJ: Erlbaum.

---

<sup>1</sup>The lack of consensus over the precise nature of mental models is reflected in the existence of a wide variety of different terms that seem to be used more or less interchangeably in the system dynamics and related systems thinking literature with the term "mental models," for example, causal policy maps (Hall et al., 1994), cognitive map (Eden, 1994; see also Axelrod, 1976), cognitive model, mental map, mental policy model (Vennix, 1990), implicit model (Vennix, 1990), theories-in-use (Argyris, 1982), and policy theory (Hoogerwerf, 1984).

<sup>2</sup>The framework for describing mental models provided by Richardson et al. (1994) focuses on a more general level of abstraction than the present work. This paper is aimed at defining subcategories of what Richardson et al. (1994) refer to as "cognitive models of system function" or "means/ends models" and we refer to as "mental models of dynamic systems." Although we suggest different terminology, the two approaches to defining mental models are not contradictory, but complementary.

<sup>3</sup>Richardson et al. (1994) have termed this inability to measure mental models without error "the mental model uncertainty principle."

<sup>4</sup>The level of confusion over mental models in cognitive science is again reflected in the use of a wide variety of different terms apparently used interchangeably with the term "mental models," for example: mental picture (Alexander, 1964), mental representation (Pennington, 1985), folk theory (Kempton, 1986), naive theory (McCloskey, 1983a), naive problem representation (Larkin, 1983), intuitive theory (McCloskey, 1983b), implicit theory (Neisser, 1987), knowledge map (Howard, 1989), idealized cognitive model (Lakoff, 1987), conceptual net (Tonn et al., 1990), conceptual model (Young, 1983), internal model (Veldhuyzen and

---

Stassen, 1977), cognitive structure (Shavelson, 1972), and knowledge structure (Means and Voss, 1985).

<sup>5</sup> We do not attempt to offer an operational definition because empirical research on eliciting and mapping mental models in system dynamics and other literatures has not yet progressed to the point where a standard, agreed upon set of measurement procedures can be identified (Vennix, 1990).

<sup>6</sup>It should be noted that in cognitive psychology the term "cognitive map" is often used in a very different and much more restrictive sense to indicate internal representations of geographical maps (see Tolman, 1948; Ormrod et al., 1988).