In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences.* Kidlington, UK: Elsevier Science.

Mental Models, Psychology of

Abstract

Dedre Gentner

A mental model is a representation of some domain or situation that supports understanding, reasoning and prediction. Mental models permit reasoning about situations not directly experienced. They allow people to mentally simulate the behavior of a system. Many mental models are based on experiential generalization and analogies. These generalizations are not always accurate; researchers have identified striking cases of widespread erroneous mental models. An understanding of typical mental models – both correct and incorrect -- is important both in order to design devices that people will use correctly and in order to create effective instructional materials.

In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences*. Kidlington, UK: Elsevier Science.

Mental Models, Psychology of

A mental model is a representation of some domain or situation that supports understanding, reasoning and prediction. There are two main approaches to the study of mental models. One approach seeks to characterize the knowledge and processes that support understanding and reasoning in knowledge-rich domains. The other approach focuses on mental models as working-memory constructs that support logical reasoning. (See *Reasoning by mental models*.) This article focuses chiefly on the knowledge-based approach.

Mental models are used in everyday reasoning. For example, if a glass of water is spilled on the table, people can rapidly mentally simulate the ensuing events, tracing the water through its course of falling downward and spreading across the table, and inferring with reasonable accuracy whether the water will go over the table's edge onto the floor. People's ability to infer and predict events goes well beyond their direct experience. For example, if asked 'Which can you throw further, a potato or a potato chip?' most people can give an answer immediately (the potato) even if they have never actually tossed either item.

However, mental models are not always accurate. Mental models researchers aim to capture human knowledge, including incorrect beliefs. The study of incorrect models is important for two reasons. First, the errors that a learner makes can help reveal what the learning processes must be. Second, if typical incorrect models are understood, then instructors and designers can create materials that minimize the changes of triggering errors.

A striking example of an incorrect mental model is the *curvilinear momentum* error (Clement 1983; McCloskey 1983) When college students are asked: "If a ball on a string is spun in a circle and then let it go, what path will it take?," many of them correctly say that the ball will travel at a tangent to the circle. However, a fair proportion states that the ball will move in a curved path, retaining some of the "curvilinear momentum" gained from being spun in a circle. The usual intuition is that the ball will gradually lose this curvilinear momentum, so that the path will straighten out over time. This erroneous intuition is fairly general; for example, the same error turns up when people are asked about the path of a ball blown through a circular tube. Further, the error does not yield immediately to training; it is found even in students with a few years of physics. However, it does diminish with increasing expertise.

Another striking error is seen when people are asked what trajectory a ball will follow if it rolls off the edge of a table (McCloskey, 1983). Instead of the correct answer, that the ball will fall in a parabolic path (Figure 1a), many people believe the ball will continue traveling straight, and begin falling (either straight down or in a curved path) only when its forward momentum begins to flag (Figures 1c and 1b). People seem to believe that sufficient forward momentum will overcome the tendency to fall. This error is sometimes called 'Roadrunner physics' because it resembles the event in which a cartoon character runs off a cliff but does not fall until some distance over the edge. However, McCloskey noted that the same error occurs in the writings of Jean Buridan and other 14th century Aristotelian philosophers. It appears that cartoon events were created to match a mental model that arises naturally from experience, possibly by overgeneralizing from experiences with linear momentum.

Mental models can facilitate learning, particularly when the structure of the new learning is consistent with the model. For example, Kieras and Bovair (1984) showed that subjects could operate a simulated device more accurately and could diagnose malfunctions better when they had a causal mental model of its functioning, rather than a merely procedural grasp of how to operate it. Similarly, Gentner and Schumacher (1986) showed that subjects were better able to transfer an operating procedure from one device to another when they had a causal mental model of the operation of the first device, rather than just a set of procedures. The degree of facilitation depended greatly on the match between the original model and the new material.

Mental models are used to explain human reasoning about physical systems: devices and mechanisms (de Kleer & Brown 1983; Hegarty & Just 1993; Kieras & Bovair 1984; Williams, Hollan & Stevens 1983); electricity (Gentner & Gentner 1983); the interactions of people with computers and other devices (Norman 1988), and knowledge of home heating systems (Kempton 1986). They have also been applied to spatial representation and navigation (Forbus 1995; Hutchins 1983; Tversky 1991); ecology (Kempton, Boster & Hartley 1994), human population growth (Gentner & Whitley 1997), and the development of astronomical knowledge (Vosniadou & Brewer 1992).

Mental models are related to several other kinds of representational structures (See Markman 1999 for a comprehensive discussion). *Schemas* (or *schemata*) are general belief structures. *Scripts* are schemas summarizing event sequences, characterized by a chiefly linear temporal order, with limited inferential flexibility. *Naive theories* or *folk theories* are global systems of belief, typically encompassing larger domains such as biology. The terms *mental models* and *naïve* or *folk theories* overlap in their application, though mental models are typically more specific than theories.

Characteristics of mental models

Mental models reasoning relies on qualitative relations, rather than on quantitative relations. People can reason well about the fact that one quantity is less than another without invoking the precise values of the quantities. This principle forms the basis for qualitative process theory, discussed below (Forbus 1984)

Mental models often permit *mental simulation*: the sense of being able to run a mental model internally, so that one can observe how it will behave and what the outcome of the process will be. The processes that underlie mental simulation are still under study. However, there is good evidence that people are able, within limits, to mentally simulate the behavior of a device, even if they are simply shown a static display (Hegarty & Just 1993). There is an apparent tradeoff between on-line simulation and retrieval of stored outcomes (Schwartz & Black 1996). As people become familiar with a system, they no longer carry out full simulations of behavior in all cases, but instead simply access their stored knowledge of the outcome.

Another finding of mental models research is that people are capable of holding two or more inconsistent models within the same domain, a pattern referred to as *pastiche models* (Collins & Gentner 1987) or *knowledge in pieces* (diSessa 1988). For example, Collins & Gentner (1987) found that many novice subjects had "pastiche" models of evaporation. A novice learner may give one explanation of what causes a towel to dry in the sun and a completely different explanation of what causes a puddle of water to evaporate, failing to see any connection between the two phenomena. Novices often use locally coherent but globally inconsistent accounts, often quite closely tied to the details of the particular example. This pattern emphasizes the tendency of novices to learn conservatively, with knowledge cached in highly specific, context-bound categories. So long as each model is narrowly accessed in contexts specific to it, the inconsistencies may never come to the learner's attention.

Mental models in everyday life

Kempton, Boster, and Hartley (1994) note that mental models "give an underlying structure to environmental beliefs and a critical underpinning to environmental values." For example, Kempton (1986) proposed on the basis of interviews that people used two distinct models of home heating systems. In the (incorrect) valve model, the thermostat is thought to regulate the rate at which the furnace produces heat; setting higher makes the furnace work harder. In the threshold model, the thermostat is viewed as setting the goal temperature, but not as controlling the rate of heating; the furnace runs at a constant rate. (This is the correct model for most current household systems.)

Having derived these two models from interviews, Kempton asked whether these models could explain people's real behavior in running their household furnaces. He examined thermostat records collected by Socolow (1976) from real households and found that the patterns of thermostat settings fitted nicely with the two models he had found. In particular, some families simply set their thermostat twice a day -- low at night, higher by day, consistent with the threshold model -- while others constantly adjusted their thermostats and used a range from extremely high to much lower temperatures. This is an extremely expensive strategy, in terms of fuel consumption, but it is follows from the valve model. In this model, the thermostat setting controls how hard the furnace works, so the higher the setting, the faster the house will warm up. This reasoning can be seen in the analogies produced by Kempton's interviewees. Those with the valve model often compared the furnace to other valve devices, such as a gas pedal or a faucet and suggested that you need to "turn 'er up high" to make the house warm up quickly.

Three significant generalizations can be made so far. First, people use mental models to reason with; they are not merely a convenient way of talking. Second, mental models can facilitate problem solving and reasoning in a domain. Third, mental models can yield incorrect results as well as correct ones. The next issues are where mental models come from and how they are used in learning. and instruction. There is evidence that mental models can influence real-life environmental decision making.

Analogies and mental models

Mental models are often based on implicit or explicit analogies with other knowledge. The incorrect valve models used by Kempton's informants, discussed above, were apparently drawn from experiential analogies. However, analogical models can also be a useful way to extend knowledge from well-understood domains to less familiar domains. For example, Gentner and Gentner (1983) identified two common mental models of electricity, the *flowing water* model and the *moving crowd* model. In the flowing water model, current flows through a wire the way water flows through a pipe, and a resistor is a narrow pipe. In the moving crowd model, current is viewed as the rate of movement of a crowd through a hall, and a resistor as a gate through to the next hall. Although both analogies can account for many simple facts about D.C. circuits, they each have drawbacks. Voltage is easy to map in the flowing water model (The number of batteries corresponds to the number of pumps pushing the water forward), but it is awkward to map in the moving crowd model (unless perhaps to a loud noise impelling the crowd forward). In contrast, the behavior of resistors is easier to predict if they are seen as gates (as in the moving crowd model) than if they are seen as constrictions (as in the flowing water model). Thus, if these analogical models are really used in reasoning, people with the water model should reason more accurately about combinations of batteries than people with the crowd model, and the reverse for resistors. Indeed, that was what was found. When people filled out a questionnaire about their mental model of electricity, and then made simple predictions about combination circuits, people who held the flowing water model were more accurate about combinations of batteries, and those with the moving crowd model were more accurate about combinations of resistors.

Methods of studying mental models.

The initial elicitation of mental models is often done by the direct method of interviews or questionnaires that explicitly ask people about their beliefs (for example, Collins & Gentner 1987; Kempton 1986) or by analyzing think-aloud protocols collected during reasoning (Ericksson & Simon 1984). (See *Protocol analysis in psychology.*) However, directly asking people about their mental models is not enough, for people are often unable to fully articulate their knowledge. Therefore, many researchers follow this direct interview with other methods of validating the proposed mental models. Once the mental models in a domain are roughly known or guessed, materials can be designed to bear down on the details. For example, problems are designed such that subjects' mental models can be inferred from patterns of correct and incorrect answers, response times, eye movements, or particular errors made (Gentner & Gentner 1983; Hegarty & Just 1993; Schwartz & Black 1996) or patterns of retention for new materials in the domain (Bostrum, Atman, Fischhoff, & Morgan 1994).

Representing mental models

Mental models research often includes an explicit representation of the knowledge. For example, in Patrick Hayes' (1985) classic paper on the naive physics of liquids, roughly 80 axioms are used to represent the knowledge involved in understanding the possible states a liquid can take and the possible transitions that can occur between states. These axioms capture knowledge about when a liquid will flow, stand still, or spread into a thin sheet on a surface.

A useful formalism for representing mental models is *qualitative process (QP) theory* (Forbus 1984). This theory, originating in artificial intelligence, aims to capture the representations and reasoning that underlie human reasoning about physical processes in a manner sufficiently precise to permit computer simulation. A central intuition is that human reasoning relies on qualitative relations, such as whether one quantity or greater or less than another, rather than on quantitative relations. For example, in QP theory, a mental model is represented in terms of (1) the entities in the domain – e.g., water in a pan, (2) qualitative relations between quantities in the domain -- e.g., that the temperature of water is above freezing and below boiling; (3) the processes that create change -- e.g., heat flow or liquid flow; and (4) the preconditions that must hold for processes to operate. An important feature of QP theory is that it uses *ordinal relationships* between quantities, such as that one quantity is greater than another, rather than representing quantities as numerical values. The idea is to match human patterns of reliance on qualitative relations rather than on exact values. A second important feature is that instead of using exact equations, QP theory uses a *qualitative mathematics* to provide a causal language that expresses partial knowledge about relationships between quantities. For instance, *qualitative proportionalities* express simple causal relations between two quantities. The idea is that people may know, for example, that greater force leads to greater acceleration, without knowing the exact numerical nature of the function (linear, exponential, etc.). An interesting aspect of QP theory is that, in addition to representing novice models, it can also capture an important aspect of expert knowledge: namely, that experts typically parse a situation into qualitatively distinct subsystems before applying more exact equations. QP theory allows researchers to describe people's knowledge about what is happening in a situation at a particular time, how the system is changing, and what further changes will occur.

Implications for instruction and design

Mental models developed from experience can be resistant to instruction. In the case of curvilinear momentum cited above, even students who had learned Newton's laws in physics classes often maintained their belief in curvilinear momentum. One technique that has been used to induce model revision is that of *bridging analogies* (Clement 1991). Learners are given a series of analogs. The first analog is a close match to the learner's existing model (and therefore easy to map). The final step exemplifies the desired new model. The progression of analogs in small steps helps the learner to move gradually to another way of conceptualizing the domain.

Mental models have been used in intelligent learning environments (See *Intelligent tutoring systems.*) For example, White and Frederiksen's (1990) system for teaching physical reasoning begins with a simple mental model and gradually builds up a more complex causal model. Early in learning, they suggest, learners may have only rudimentary knowledge, such as whether a particular quantity is present or absent at a particular location. By adding knowledge of how changes in one quantity affect others, and then progressing to more complex relationships among quantities, learners can acquire a robust model.

Another implication of mental models research is that the pervasiveness and persistence of mental models needs to be taken into account in designing systems for human use. Norman (1988) argues that designers' ignorance of human mental models leads to design errors that plague their intended users. Sometimes these are merely annoying – e.g., a door that looks as though it should be pulled, but that need to be pushed instead. However, failure to take mental models into account can lead to serious

An example of such a failure of mental models occurred in the Three-mile Island nuclear disaster. Early in the events that led to the melt-down, operators noted that the reactor's coolant water was registering at a high pressure level. They interpreted this to mean that there was too much coolant and accordingly they pumped off large amounts of coolant. In fact, the level was dangerously low, so much so that the coolant was turning into steam -- which, of course, led to a sharp increase in pressure. Had this alternate model been at hand, the operators might have taken different action.

Mental models as temporary aids to logical reasoning

Another approach to mental models is taken by Johnson-Laird (1983) and his colleagues. (See *Reasoning by mental models.*) This approach differs from the research cited in the remainder of this entry in that it views mental models as temporary working memory sketches set up for the purposes of immediate reasoning tasks such as propositional inference (Johnson-Laird 1983). The focus on immediate working-memory tasks in this approach has led to a relative lack of emphasis on long-term knowledge and causal relations. However, there may be value in bringing together the working-memory approach with the knowledge-intensive approach. There is evidence that long-term causal mental models can influence the working-memory representations that are set up in speeded tasks (Hegarty & Just 1993; Schwartz & Black 1996).

Bibliography

Bostrum, A., Atman, C. J., Fischhoff, B., Morgan, M. G. (1994). Evaluating risk communications: Completing and correcting mental models of hazardous processes, Part II. *Risk Analysis, 14*, No. 5.

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In D. Gentner, A. L. Stevens (Eds.), *Mental models* (pp. 325-340). Hillsdale, NJ: Erlbaum.

Clement, J. (1991). Nonformal reasoning in experts and in science students: The use of analogies, extreme cases, and physical intuition. In Voss, J., Perkins, D., & Siegal, J. (Eds.), *Informal reasoning and education* (pp. 345-362). Hillsdale, NJ: Erlbaum.

Collins, A., Gentner, D. (1987). How people construct mental models. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 243-265). Cambridge, England: Cambridge University Press.

de Kleer, J., Brown, J. S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 155-190). Hillsdale, NJ: Erlbaum.

diSessa, A. A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science, 6,* 37-75.

Ericksson, K. A., Simon, H. A. (1984). *Protocol analysis.* Cambridge, Mass: M.I.T. Press.

Forbus, K. D. (1984). Qualitative process theory. *Journal of Artificial Intelligence, 24,* 85-168.

Forbus, K. 1995. Qualitative Spatial Reasoning: Framework and Frontiers. In Glasgow, J., Narayanan, N., and Chandrasekaran, B. *Diagrammatic Reasoning: Cognitive and Computational Perspectives.* MIT Press, pp. 183-202.

In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences.* Kidlington, UK: Elsevier Science.

Gentner, D., Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99-129). Hillsdale, NJ: Erlbaum.

Gentner, D., Stevens, A. L. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Erlbaum.

Gentner, D., Whitley, E. W. (1997). Mental models of population growth: A preliminary investigation. In M. Bazerman, D. M. Messick, A. E. Tenbrunsel, & K. Wade-Benzoni (Eds.), *Environment, ethics, and behavior: The psychology of environmental valuation and degradation* (pp. 209-233). San Francisco, CA: New Lexington Press.

Hayes, P. J. (1985). Naive Physics I: Ontology for Liquids. In J.R. Hobbs and R.C. Moore (Eds.), *Formal Theories of the Commonsense World.* Norwood, NJ: Ablex Publishing Corporation.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language. 32*, 717-742.

Hutchins, E. (1983). Understanding micronesian navigation. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 191-225). Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kempton, W. (1986). Two theories of home heat control. *Cognitive Science, 10,* 75-90.

Kempton, W., Boster, J.S., and Hartley, J. *Environmental values in American culture.* MIT Press 1994.

Kieras, D. E., Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8,* 255-273.

In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences.* Kidlington, UK: Elsevier Science.

Markman, A. B. (1999). *Knowledge representation.* Mahwah, NJ: Erlbaum.

McCloskey, M. (1983). Intuitive Physics. *Scientific American, 248*(4), 122-130.

Norman, D. A. (1988). *The psychology of everyday things.* New York: Basic Books.

Schwartz, D. L., Black, J. B., (1996). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology, 30*, 154-219.

Socolow, R. H. (Ed.). (1978). *Saving energy in the home: Princeton's experiments at Twin Rivers.* Cambridge, Mass: Ballinger.

Stevens, A., Collins, A. (1980). Multiple conceptual models of a complex system. In R. Snow, P. Federico. & W. Montague (Eds.), *Aptitude, learning and instruction: Cognitive process analysis* (Vol.2, pp.177-197). Hillsdale, NJ: Lawrence Erlbaum.

Tversky, B. (1991). Distortions in memory for visual displays. In S. R. Ellis, M. Kaiser, & A. Grunewald (Eds.), *Spatial instruments and spatial displays* (pp. 61-75). Hillsdale, NJ: Erlbaum.

Vosniadou, S., Brewer, W. F. (1992). Mental models of the Earth: A study of conceptual change in childhood. *Cognitive Psychology, 24* (4), 535-585.

Williams, M. D., Hollan, J. D., Stevens, A. L. (1983). Human reasoning about a simple physical system. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 131-153). Hillsdale, NJ: Erlbaum.

White, B.Y., Fredricksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence, 42* (1), 99-157.

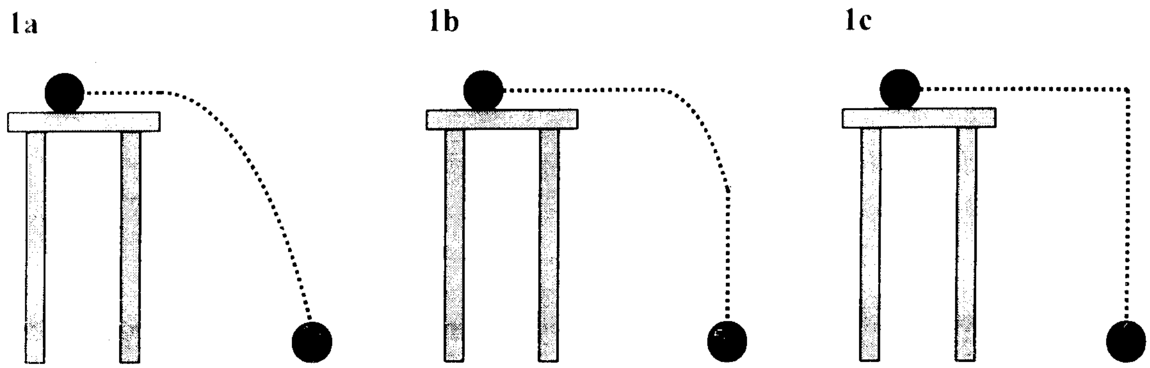
Dedre Gentner, Northwestern University

In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences*. Kidlington, UK: Elsevier Science.

Figure Captions

Figure 1

Responses to the question "What path will the ball take after it rolls off the table?" (adapted from McCloskey, 1983)


1a
1b
1c