Proceedings of the Annual Meeting of the Cognitive Science Society

Component Models Of Physical Systems

https://escholarship.org/uc/item/3w28g542

Proceedings of the Annual Meeting of the Cognitive Science Society, 7(0)

Collins, Allan

1985

COMPONENT MODELS OF PHYSICAL SYSTEMS

Allan Collins
Bolt Beranek & Newman, Inc.

In order to get around in the world people have to make sense of Coke machines and computers, home heating systems and electric circuits, and even evaporation processes and bouncing balls. They must build their own folk models of how these things behave e.g., what will happen if they put different amounts of money in a Coke machine, what to do if the machine doesn't behave as they expect. "Mental Models" is the term that has evolved for a new view of how people conceptualize physical systems. Mental models are meant to imply a conceptual representation that is qualitative, and that you can run in your mind's eye and see what happens.

In order to make the discussion concrete, I will display some mental models for your inspection. I've picked three domains - electricity, home heating systems, and evaporation, - because they cover the diversity of different mental models, and because it is possible to illustrate a variety of hypotheses about mental models in terms of these domains.

Mental Models of Electric Circuits

A number of investigators, have studied people's naive models of electric circuits (Fredette & Lochhead, 1980; Gentner & Gentner, 1983; Osbourne 1981; Osbourne & Wittrock, 1983; Steinberg, 1983). Since my taxonomy is more differentiated then the others, I will describe the naive models I found and point out how these relate to the mental models of electricity found by other researchers.

In the study I analyzed people's naive models of a simple battery, switch, and lightbulb circuit. For the circuit, I administered a questionnaire to twenty-four subjects who had no special knowledge of circuits. The questionnaire asked them (1) to draw a schematic circuit and explain how it worked, (2) to answer a variety of questions about the working of the circuit that they drew, such as which way the current flows in each component and wire, and (3) to evaluate different circuits in order to decide whether they would work properly.

I was able to categorize the types of models people were using both at the circuit level and at the individual component level. Seven of the subjects had a more or less correct model of the way a battery circuit functions. That is, one pole of the battery acts as a source of electrons and the other as a sink. The current is maintained by an ion flow in the battery between the two poles. I explain each of the incorrect models below:

(1) Converging flow model (3 subjects). This model posits that there are two kinds of electricity (i.e., positive and negative) flowing out from the battery, and that you need both kinds to make the bulb light. I have called this model the "epoxy glue" model of circuit flow, while Steinberg (1983) refers to it as the "sink" model and Osbourne (1981) the "bipolar" model. Like me, Steinberg finds this model held by approximately one-seventh of Smith College undergraduates first taking physics, but Osbourne and Wittrock (1983) report finding this model in approximately one-third of young children.

(2) Circular flow model (4 subjects). In this view electrons flow around and

around the circuit, and each time that they pass through the battery, they are given additional momentum. This view is incorrect in that it assumes that individual particles go through the battery much like water particles go through a pump in a circulatory water system. I have called this the "racetrack" model, and it is a version of the "moving crowd" model discussed by Gentner & Gentner (1983).

(3) Impulse-signal model (3 subjects). This model posits that the switch sends an impulse to the battery to trigger current flow from the battery to the light bulb. It views a switch like the trigger of a gun. This is a variation on the "consumer" model described by Steinberg (1983) and Fredette & Lochhead (1980), and the unipolar model described by Osbourne (1981).

(4) Gate-switch model (3 subjects). This model posits that current flows only from the battery to the light bulb. For the switch to work properly, it must be inserted between the battery and the bulb, where it acts as a gate regulating current flow. This too is a version of the unipolar model found by Osbourne, and the consumer model of Steinberg (1983) and Fredette & Lochhead (1980). It seems to derive from the water-flow analogy described by Gentner and Gentner (1983), where the switch is viewed as a kind of valve regulating flow.

(5) Gate-switch model with circuit (2 subjects). This is a variant of the above model. It posits that current flows in a circuit back to the battery, but you need the switch between the battery and the bulb to keep the current from reaching the bulb when the switch is off. Thus, it is like water flowing in a circulating system; if the water can reach a small hole in the pipe (analogous to the light), then water will trickle out (i.e., the bulb will light). This is a partial consumer model of the bulb, where most of the current is assumed to flow back to the battery.

(6) Controller-switch model (2 subjects). This model views the switch as a control device that is hooked up to the battery-bulb circuit by its own control circuit. It was the least coherent of the six incorrect views I encountered and the two subjects who proposed it seemed not to believe it very strongly. None of the previous studies talk about any such view.

I also attempted to analyze the mental models subjects had of the three major components in the circuit. Here I made my best guess on the basis of their answers to the questions as to each subject's view of each component.

Based on the analysis there were four different views of the battery. Ten subjects had a source-sink view of the battery, with electrons flowing from one pole to the other in the circuit. Seven subjects viewed the battery as simply a source of electricity which flowed out to the bulb. Four subjects considered the battery to be an energizer, like a pump in a water circulation system, and three subjects considered it a source of both positive and negative electricity.

There were three prevailing views of the light bulb that I could distinguish. Twelve subjects viewed it as a narrow passage or resistor through which current must pass. Ten subjects viewed it as a consumer of electric current, three of these requiring both positive and negative electricity to come together to be consumed. The last two subjects had what I call the partial consumer view where a little bit of the electricity is burned up as it goes through the bulb.

We distinguished four views of the switch. Most subjects had the correct contact-gate view, where the switch is viewed as a swinging gate attached to one wire and making contact (or not) with the other wire. The incorrect views included the impulse- signal view that I likened to a gun's trigger, a blocking notion like a valve in a pipe, and a control-device notion like the handle on a faucet that is mechanically linked to the valve in the pipe (though both subjects had the control device linked electrically to the circuit).

What these models illustrate is that naive subjects have a diverse set of views on circuits, which are more or less coherent. These views can be used generatively to guess how circuits they have not encountered before will behave, as we will discuss later. Furthermore, these views are structured: as we have shown here, they can be regarded as models of the entire circuit or more locally as views of specific components of the circuit. Different component views can in fact go with different system views. For example, one could have an energizer view of the battery with either a contact gate, control device, or blocking view of a switch and either a resistor or partial consumer view of a light. Thus, mental models of components can be regarded as building blocks for a global view.

Mental Models of Home Heating Systems

Kempton (in press) carried out a series of forty-two, in depth interviews with average Americans to determine how they set their thermostats and why. Based on his interviews he identified two common folk theories of how thermostats work, which lead to quite different patterns of setting a thermostat. The two theories he identified are the <u>Feedback Theory</u> and the <u>Valve Theory</u>.

<u>The Feedback Theory</u>. This theory views the thermostat as a device that senses the temperature in the room, and if it falls below the temperature setting, the furnace turns on, and if it rises above the temperature setting, the furnace turns off. Actually, a thermostat has two set points equidistant from the thermostat setting; the lower of which turns on the furnace and the higher of which turns it off. Kempton quotes one of his subjects who held this view.

"You just turn the thermostat up, and once she gets up there (to the desired temperature) she'll kick off automatically. And then she'll kick on and off to keep it at that temperature."

<u>The Valve Theory</u>. This view holds that the setting on the thermostat controls the rate of heat flow, so that the higher you set the thermostat, the harder the furnace works to produce heat. He quotes a subject who held this view:

"Um, I assume, um, that there is some kind of linear relationship between where the lever is and the way some kind of heat generating system functions. And, um, that it's like stepping on the gas pedal; that there I have a notion of hydraulics, you know, the harder you push there is, the more fluid gets pushed into the engine, and the more explosions there are, and the faster it goes."

There are hints in the data Kempton quotes to indicate that the dichotomy he makes between the two theories really should be thought of as two points in a larger space of possible models. We can see this space of models best in two examples he gives of people who do not quite fit the valve/feedback dichotomy.

One informant quite clearly enunciated the feedback theory at first:

"I guess, what I always thought was when you turn the, the temperature, you turn the thermostat to 65, the furnace works to keep the room at 65 and then as soon as it's 65, the furnace stops working and then when it starts to get a little bit cold again the furnace will work again."

But when Kempton questioned her about heating up the house, she verbalized the valve theory:

Q: Let's say it's very cold ... you come into the house and it's <u>very</u> cold, and you want to heat the house up. Let's say you want to heat the house up to 65. What would you do ...

R: If it's very, very cold?

Q: Uh-huh.

R: I might turn it up to 70, for maybe 20 minutes, half an hour and then turn it back down to 65 to see if I can get it warmer faster

Let me consider how such a hybrid model might work. The essence of the valve theory is that the amount of heat flow is proportional to the temperature: this is a proposition about a factor I will be able <u>rate of heat flow</u>. The essence of the feedback theory is that the heat turns on when the temperature falls below a set point, and turns off when the temperature rises above another set point: this is a proposition about a factor I will call <u>type of control</u>. Because these two theories address different factors, it is perfectly possible to believe in variable heat flow, as in the valve theory, and indirect control, as in the feedback theory. The hybrid theory is simply that the temperature setting controls both factors: the higher the setting, the greater the heat flow, and the higher the set points at which the furnace turns on and off.

Another of Kempton's protocols points up a third factor, <u>heat loss</u>, that affects the way people set their thermostats. In particular, a woman respondent believed in setting the thermostat back at night, but her husband reasoned from his feedback theory that it did not pay to do so:

"Now, my husband disagrees with me. He, he feels, and he will argue with me long enough, that I do not save any fuel by turning the thermostat up and down.... Because he, he feels that by the time you turn it down to 55 and all the objects in the house drop to 55°, you're going to use more fuel than if you would have left it at 65 and it just kicks in now and then."

The essence of the husband's argument is that the time the furnace is on when you move the thermostat up in the morning is as great or greater than the time it is off when you set the thermostat down at night. The two times are in fact roughly equivalent. The savings from turning the thermostat down at night come because the furnace turns on less frequently during the night. But it is easy to see how her husband might not think of this savings, if he thinks the only effect of lowering the setting is the two transients which offset each other. He is implicitly assuming that the furnace is turning on and off at the same rate at a lower setting as a higher setting, once steady-state is achieved. Such a view makes sense if you think heat loss from a house is constant, or simply depends on the temperature outside. In order to really "understand" why there is a savings, you have to have a notion that heat loss is proportional to the difference between the temperature indoors and outdoors, or at least that it depends on the indoor temperature.

We have thus identified three factors that appear in Kempton's protocols, for which people may have different models. These factors are equivalent to the component models I discussed with respect to electricity. I would argue that most combinations of the different models for these three components are possible.

Let me explain each of the possible models and their combinations. With respect to the "heat flow" component, the graded heat flow model is what the valve theory implies. The constant heat flow model, which is implied by the feedback theory, is the correct model (ignoring transients). With respect to the "type of control" component, direct control, which the valve theory implies, means that the rate of heat flow changes as you move the thermostat up or down. Indirect control by temperature, as the feedback theory implies, refers to the model where changes in temperature cause the furnace to turn on and off. A third possible model is an adjustable ratio model: this model assumes that the furnace turns itself on and off with a particular frequency, and changing the temperature setting causes the furnace to turn on and off at a different ratio. This might be called an open-loop model as opposed to a feedback model. Finally, with respect to heat loss, people might have any of the three models I described in the discussion of the husband who wouldn't turn the heat back at night.

In summary, I see there are two dominant views of how a thermostat works and that these influence people's behavior in substantive ways. But in addition to these two dominant views, at a finer-grain level of analysis, there are a host of possible models. These derive from combining component models in various ways.

Mental Models of Evaporation

In a study where they posed eight difficult questions about evaporation to four novice subjects, Collins and Gentner (in press) found that subjects' models of evaporation consisted of five component subprocesses:

1) How molecules behave in the water
2) How molecules escape from the water to the air
3) How molecules behave in the air
4) How molecules return to the water from the air
5) How molecules go from liquid to vapor, and vice versa

They enumerated a number of different views for each of these component processes, that were suggested in subjects' protocols.

Behavior in the water. Figure 1 shows four models for how molecules behave in water. The first view is called the sand-grain model – the molecules just sit there like grains of sand, moving and slipping when something pushes on them. The temperature of the water is the average temperature of the individual molecules. This is a very primitive model. The next two views assume that the molecules are bouncing around in the water like billiard balls in random directions. In both these views, the speed of the molecules reflects the temperature of the water. The difference is that in one version -- the equal speed model -- all the molecules are moving at the same speed. The other version is a random speed model, that allows for differences in speed for different particles. On this view temperature reflects the average speed of a collection of molecules. The fourth view, called the molecular attraction model, incorporates attraction between molecules into the random speed model. In it molecules move around randomly, but their paths are constrained by the attractive (and repulsive) electrical forces between molecules. This view is essentially correct.

Escape from the water. Figure 2 shows three possible component models for escape (pictorially two are the same). The heat-threshold model is a threshold view of escape: The molecules have to reach some temperature, such as the boiling point of the liquid, and then they pop out of the liquid, the way popcorn pops out of the pan when it is hot enough. The remaining two models focus on molecular velocity, rather than the incorrect notion of molecular temperature. The rocketship model is based on the assumption that the molecules in the water are moving in random directions. In order to escape from the water (like a rocketship from the earth), a molecule must have an initial velocity in the vertical direction sufficient to escape from gravity. The third view, the molecular escape model, posits that the initial velocity must be great enough to escape from the molecular attraction of the other molecules. Both these latter models are in part correct, but the major effect is due to the molecular attraction of the water.

Behavior in the air. There are three component models of how the water molecules behave in the air are depicted in Figure 3. The container model posits that the air holds water molecules and air molecules mixed together until it is filled up (at 100% humidity). The variable-size-room model is a refinement of the container model to account for the fact that warm air holds more moisture than cold air. In this model, molecules in warm air are further apart, and so are less dense than molecules in cold air.

That leaves more space to put water molecules in warm air than in cold air. In the exchange-of-energy model, the chief reason that cold air holds less moisture than warm air is that its air molecules are less energetic. When water molecules in the air collide with air molecules, they are more likely to give up energy if the air is cold (and hence less energetic) than if it is warm. If the water molecules become less energetic, they are more easily captured by the molecular attraction of other water molecules (or a nucleus particle). When enough water molecules collect around a particle, they will precipitate. This latter view is essentially correct.

Return to the water. Figure 4 shows three models of how water molecules return to the water. The crowded room model assumes that when all the space in the air is filled, no more water molecules can get in. The aggregation model assumes that water molecules move around in the air until they encounter a nucleus or particle (which could be another water molecule) around which water accumulates. The less energetic the molecule, the more likely it is to be caught by the molecular attraction of the particle. As these particles accumulate water, gravitational forces overcome the random movement of the particles and they precipitate. The recapture model assumes that particles are attracted by the surface of the water (or other surfaces). The less energy they have, the more likely they are to be recaptured. The action in this view takes place near the surface, unlike the aggregation view. Both the aggregation and the recapture models are essentially correct, but the aggregation model takes place over a long time period with relatively high humidities, whereas the recapture model is applicable in any situation where evaporation is occurring.

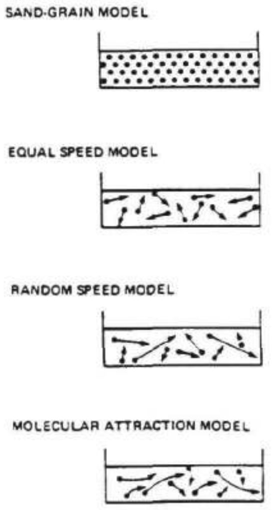
Liquid-vapor transition. Figure 5 shows four different views for the transition from liquid to vapor and from vapor to liquid. One view, the coterminus model, is that the transition occurs when the molecules leave the water and escape to the air, and vice versa. On this view the two transitions, between water and air, and between liquid and vapor, are the same transition. In other words, whether a molecule is in the vapor or liquid state depends solely on location. all molecules beneath the surface of the water are liquid, and all molecules above the surface of the water are vapor. A second view, the intrinsic state model, treats the liquid or gas state as an intrinsic property of the molecule. If the molecule becomes hot enough, it changes from liquid to vapor, and if it becomes cold enough it changes from vapor to liquid. Location is correlated with state, in that molecules in the vapor state tend to move into the air, while molecules in the liquid state remain in the water. A third view, the disassembly model, is based on a little chemistry: in it liquid water is thought of as made up of molecules of H_2O, whereas the hydrogen and oxygen are thought to be separated in water vapor. The expert view, called the binding model, is based on molecular attraction: water molecules in the liquid state are partially bound together by electrical attraction of the neighboring molecules, whereas molecules in the gaseous state bounce around rather freely. The bubbles in a boiling pan of water are thus water molecules that have broken free of each other to create a small volume of water vapor, and clouds and mist are microscopic droplets of liquid water that have condensed, but are suspended in the air.

Combining component models. Table 3 summarizes all the component models described above. Subjects can combine these component models in different ways. Collins and Gentner (in press) show answers from two subjects who had different combinations of these component models. One subject had a model constructed from the random speed model of water, the rocketship and molecular escape models of escape, the variable-size-room model of the air, the crowded room model of return, and the coterminus model of the liquid-vapor transition. The other subject had a less, consistent and less stable model of evaporation. His view included something like the heat-threshold model of escape, the container model of the air, the recapture model of return, and the intrinsic state model of the liquid-vapor transition. In contrast, as I have indicated, the expert view is made up of the molecular attraction model of water, the rocketship and molecular escape models of escape, the exchange-of-energy model of the air, the aggregation and recapture models of return, and the binding model of the liquid-vapor transition.

Summary

There have been a variety of attempts to identify naive subjects' mental models, but most of these have settled on two or three global views (e.g. Kempton, 1984; Osbourne, 1981; Steinberg, 1983). At a finer grain level of analysis, however, it is possible to identify a number of different components and a variety of mental models for each component. These component models can be combined in many different ways. But frequently two or three combinations predominate, giving rise to the global mental models identified by different researchers.

The componential approach has been pursued by a number of researchers in artificial intelligence (de Kleer 1979; de Kleer & Brown 1981, 1983; Forbus 1981, 1982; Hayes 1984). But none of these researchers had tried to analyze the component models human subjects actually have. What I have tried to show is how a componential analysis is necessary to characterize the way people understand physical systems.



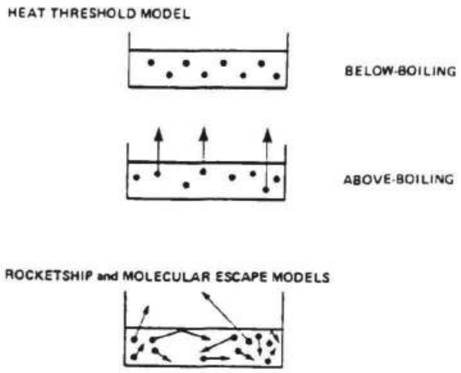


Figure 1. Component models of the behavior of molecules in water

Figure 2. Component models of how water molecules escape from water to air

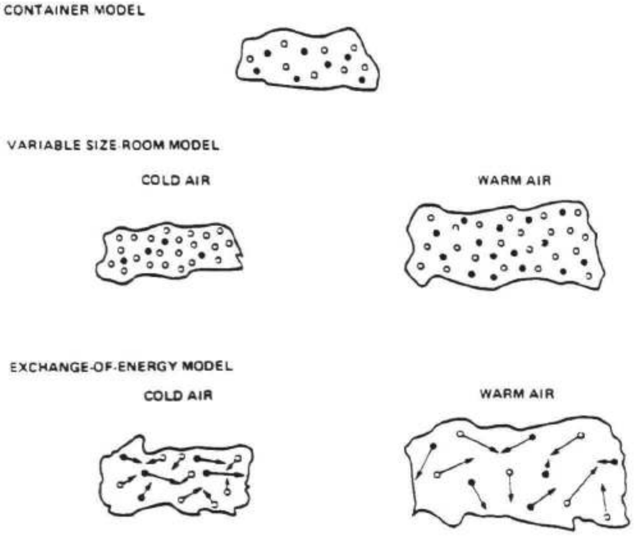


Figure 3. Component models of the behavior of water molecules in air (water molecules are filled circles; air molecules are open circles)

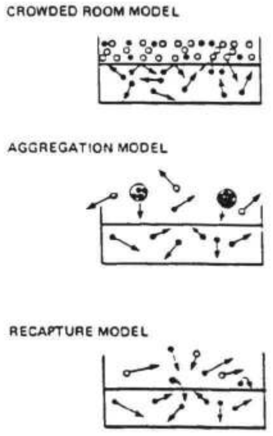


Figure [illegible]. Component models of how water molecules return from air to water

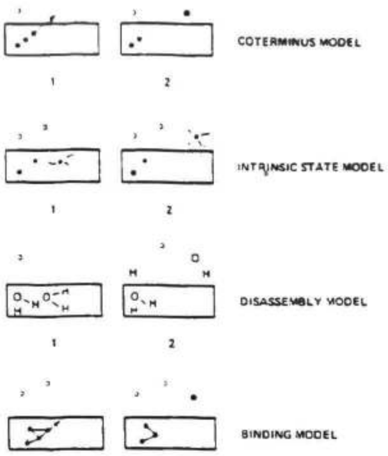


Figure 5. Component models of the liquid-vapor transition

References

Collins, A. & Gentner, D. How people construct mental models. In N. Quinn and D. Holland (Eds.), *Cultural models in language and thought*, in press.

de Kleer, J. The origin and resolution of ambiguities in causal arguments. *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, Tokyo, Japan, 1979, 197–203.

de Kleer, J., & Brown, J.S. Mental models of physical mechanisms and their acquisition. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum, 1981.

de Kleer, J., & Brown, J.S. Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A.L. Stevens (Eds.), *Mental models*. Hillsdale, NJ: Erlbaum, 1983, *7*, 121–153.

Forbus, K. Qualitative reasoning about physical processes. *Proceedings of IJCAI-7*. 1981.

Forbus, K.D. *Qualitative process theory* (A.I.M. 664). Artificial Intelligence Laboratory. Massachusetts Institute of Technology, February 1982.

Fredette, N., & Lochhead, J. Student conceptions of simple circuits. *The Physics Teacher*, 1980, *18*, 194ff.

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

Hayes, P. Ontology for liquids. In J. Hobbs and R. Moore (Eds.), *Formal theories of the common sense world*. Norwood, NJ: Ablex, 1984.

Kempton, W. Two theories used for homes heat control. In N. Quinn and D. Holland (Eds.), *Cultural models in language and thought*, in press.

Osborne, R. Children's ideas about electric current. *New Zealand Science Teacher*, 1981, *29*, 12–19.

Osborne, R.J., & Wittrock, M.C. Learning Science: A generative process. *Science Education*, in press.

Steinberg, M.S. Reinventing electricity. In *Proceedings of the International Seminar, Misconceptions in Science and Mathematics*, Ithaca, New York, June 1983.