ELSEVIER

International Journal of Industrial Ergonomics 22 (1998) 293–297

International Journal of
Industrial
Ergonomics

Identifying mental models of complex human–machine systems

Neville Moray

Lamih-Percotec, Universite de Valenciennes, B.P. 311, Lemont Hoy, 59304, Valenciennes Cedex, France

---

Abstract

The notion of a 'mental model' is widespread but ill defined. Its meaning can be clarified by noting both the cognitive processes of the worker and the context of the task. A mental model is a mapping of the properties of the task to its representation in the mind of the worker. Traditionally, such models have been identified by the use of protocol analysis, but a method has been developed which finds the mapping through its effects on the coupling between human and machine. This method uses Conant's theory of system decomposition based on high-order Shannon information theory. An example of its use is given.

Relevance to industry

Operators make use of mental models of complex systems in process control and manufacturing. A method to identify such models is relevant to productivity, and to fault detection and management, since it gives insight into the way in which operators become coupled, by training and habit, to the dynamics of the processes they control.

Keywords: Process control; Mental models; Model identification; Human–machine systems

---

1. Introduction

As Rutherford and Wilson (1991) point out, the notion of mental model is very confused. At one extreme, we find Johnson-Laird (1983) who uses it to discuss how people reason about abstract logical problems, a task which bears little resemblance to real decision making. At the other extreme, we find those who aim to understand how operators, maintenance people, and in general 'workers' interact with highly complex physical systems such as chemical plants, nuclear power plants, etc. (Bainbridge, 1974, 1996; De Keyser, 1986; Rasmussen, 1986; Rasmussen et al., 1995 for example). In between lie mental models as the basis for interaction with relatively simple systems such as word processors (Green, 1990), and for vehicular control (Veldhuijzen and Stassen, 1977), or to explain experiments on causal reasoning about simple mechanical or devices (Gentner and Stevens, 1983).

PII S0169-8141(97)00080-2

Moray (1997, 1998) has shown that the confusion in the literature can be removed by seeing how workers are coupled both to their task (plant, machine, problem) and the environment in which the task is carried out. For example, in the logical problem-solving research there is no interaction between the 'operator' and the task. (I will use the word 'operator' rather than subject throughout, to emphasize the ergonomists' interest in work situations.) The task is presented to operators who can walk away carrying (a model of) the task in their heads to work on it at leisure, with no coupling back to the task or environment. Nothing the operator does will change the nature of the task, nor will the latter be affected by perturbations in the environment. There are no task dynamics. The properties of the task will not change however long the operator takes to solve it. At the other extreme, there are complex causal dynamics continuously present. If the operator does not respond to the demands of the task, the properties of the task will change. Physical plant will continue to operate in the absence of any intervention by operators, and the values of state variables, etc., will change from moment to moment, sometimes dramatically in the case of faults. When the operator does intervene, he or she changes the state of the plant and to some extent determine the future trajectory of the task dynamics. Both operator and task are subject to perturbations from the environment, such as changes in temperature, noise, changes in the quality of raw material or feedstock, etc., and again the operator is often to some extent coupled with the environment either directly or through the task. In addition, whereas Johnson-Laird seems to think of mental models as existing only in working memory, and being called into existence by the arrival of the task, workers such as Bainbridge typically think in terms of long-term knowledge being a model of the task. One might add that since workers seldom work alone in real tasks, one should include interpersonal and organizational properties of tasks as part of the model, something conspicuously absent from the laboratory studies which use the notion. A graphical summary of three of these relations is shown in Fig. 1.

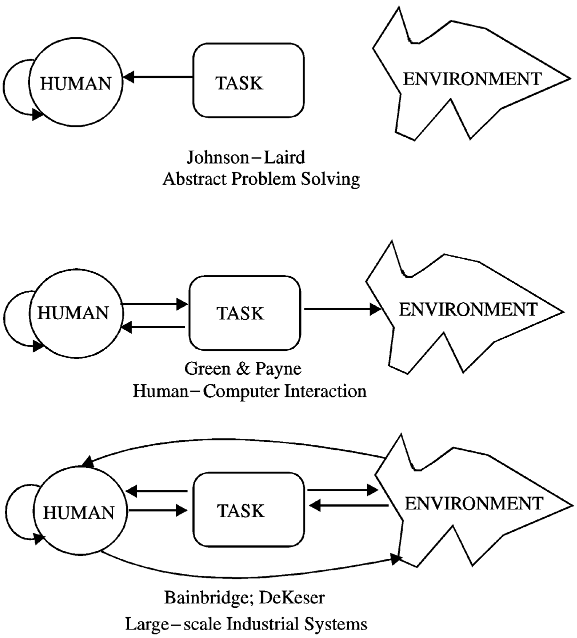

Fig. 1. The coupling of human to task and environment in different cases where 'mental models' have been invoked. The 'self-loop' at the left indicates that the human can think about the task using the model even in the absence of input and output to and from the task. Arrows indicate causal coupling.

2. The nature of mental models

Given that we can understand and remove the confusion in the literature by noting that there are several different ways in which the notion of mental model has been invoked which correspond to different couplings between human, task and environment, what are mental models and how may they by identified? Following Ashby (1956) we note that in a formal sense models are mappings of properties of a set into or onto another set. Typically, the mapping is many-to-one, or formally a 'homomorphism', not an isomorphism. Thus, in almost all mappings some information is lost. A scale model of an aircraft preserves wings, tail, relative size, etc., but loses hydraulics, details of the engine, etc. Similarly, a mental model of a plant preserves those features which have been noted during training, or in reading procedures, or in operating the plant, but loses many details. (An operator may not, for example, know the physical location of various pumps, valves, etc.) Once a model has been formed in the head by a mapping from the physical world, subsequent models may be formed internally by mappings of mappings. Thus, an operator may initially learn all the details of the controls in a control panel, but later come to think of them not as ‘Valve 1, Valve 2, Pump 6,’ etc., but as ‘Cooling system, Steam generator’, etc.. This description in turn can be remodelled into ‘Power Generation, Power distribution’, etc. Thus, operators construct a hierarchical set of models as a series of many-to-one mappings. At the same time an operator may make parallel models at a single level of detail. These models are mappings of different sets of properties. Thus, one may think of a plant in terms of its physical topography, of the causal relations at the levels of fundamental physical processes, or in terms of what actions are needed to control it, of its purposes. These vertical and horizontal mappings are formally a representations of the well-known Abstraction and Means-Ends hierarchies of Rasmussen (1986).

The effect of multi-level modelling is greatly to reduce the mental workload of the operator. Quite simply, higher level models have fewer components. The danger is that because the mappings are homomorphic, not isomorphic, it is impossible to guarantee if one returns to a more detailed (less abstract) model, that one will necessarily arrive at a desired piece of knowledge. One-to-many mappings are inherently ambiguous. A fuller account of how the mapping and the different models are related can be found elsewhere (Moray, 1990, 1997, 1998).

3. Identifying mental models

We now turn to the question of how to identify the properties of a mental model. Traditionally this has been done by inferences based on protocol analysis (see, e.g., Bainbridge (1974) for a particularly clear example). We propose a different and quantitative approach. Our approach is derived from Conant’s (1976) method for system identification and decomposition (Conant, 1976). All real complex systems can be thought of as made up of subsystems which are more or less tightly coupled to one another. By noting how the values of variables change from moment to moment we can calculate the amount of Shannon information transmitted mutually between them throughout the system. Some components will transmit a high proportion of their information to certain others, showing that they are tightly coupled, and form a coherent subsystem. Other individual variables do not so transmit information, and so are only loosely coupled. Furthermore, one can calculate not just the transmission from one variable to others, but from groups of variables to other (groups of) variables, and thus see how subsystems are coupled. Note that we do not need to know in advance what constitutes a ‘subsystem’: it emerges naturally from the calculations.

We now postulate that when an operator constructs a mental model, what is being constructed is usually a model of causal knowledge about how a system works. That is, precisely, what causal couplings exist. This we believe is reflected in two ways: (1) the distribution of attention over parts of the systems and (2) the choice of what controls to use for action, since it is the attention–action cycle which represents the coupling of the operator to the system. By recording the pattern of attention and action, and the way in which the state variables change from moment to moment, we can, by using the measurement of mutual information transmission find how an individual operator believes the system to be decomposable into subsystems which can be used by him or her to control it. That is, we can discover the mental model being used, even if the operator is not aware of its contents.

Such mental models can be very varied. In effect different operators, coupled to a single plant in manual control, must be considered as different human–machine systems. The discovery of mental models is important not so much because models represent the knowledge in the person, but because they result in a single physical system becoming a series of different human–machine systems. It is then presumably the aim of the designer and the trainer to reduce the difference among these joint systems as much as possible in order to ensure that the system overall will function safely and productively.

We have recently begun to explore the virtues and limitations of this approach (Moray et al., 1996). Here there is only time to present one example. Consider the following: There is a pasteurization plant in which the mass and temperature of steam to the pasteurization heat exchanger are normally controlled automatically by a closed-loop controller using a temperature sensor on the output side of the heat exchanger, measuring the temperature of the feedstock emerging. That is, both the mass flow and the energy flow are coupled to the feedstock temperature. But now, suppose that the operator puts the steam mass flow into manual control mode. There will probably continue to be coupling between the mass flow and the feedstock temperature, but it could happen that the operator decides that there is a fault in the mass flow controller, and therefore, puts it at a constant value, leaving control to be closed only through the energy supply. The mental model of the operator has reconfigured the plant. Conant's method allows us to identify this reconfiguration, which is a reflection of the mental model.

Fig. 2a and Fig. 2b show the changes in the plant coupling when a single operator is confronted by the occurrence of a fault in part of plant, taken from a recent series of experiments by Jamieson in our laboratory. Clearly, the mental model has greatly changed after the fault occurred. (There is in addition another such diagram which represents the inherent organization of the plant when running under fully automatic control, which is not shown here.)

4. Conclusions

1. The notion of mental model is less chaotic than the literature suggests, and can be understood by looking not at the inherent nature of the mental model, but at how the notion is used in the contexts of different tasks and environments.
2. The fundamental notion behind all modelling is a mapping of properties in a many-to-one, homomorphic fashion.
3. Operators construct a series of mental models by first performing a mapping from the external world into their heads, and subsequently by performing a series of mappings to form models from models.

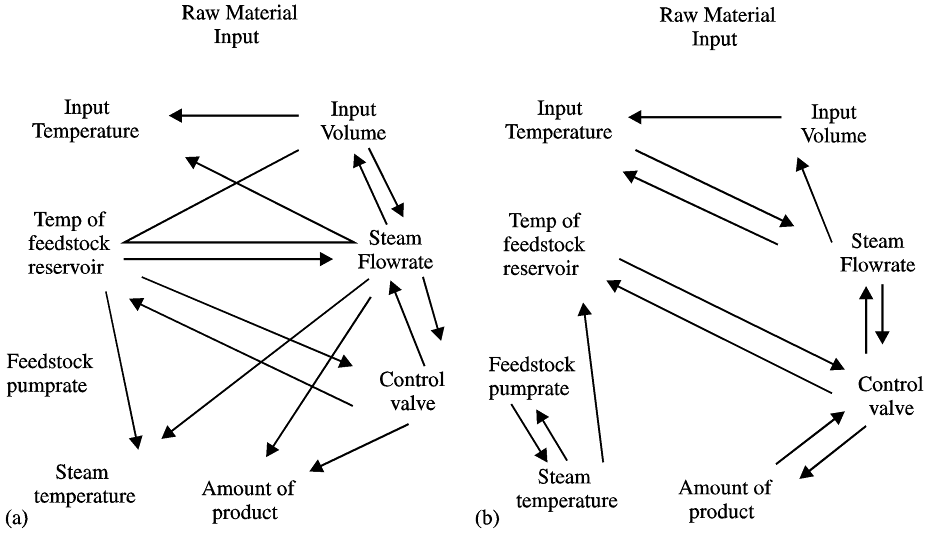

Fig. 2. (a) Couplings between subsystems in a pasteurization plant when an operator is working in supervisory control mode and the plant is working normally. (b). Components that are coupled as a result of manual intervention by the operator following a fault in the feedstock pump. The directions and strength of the couplings are found by Conant's method and reflect the content of the operator's mental model of causality.

4. The formation of mental models results in changes in the way in which operators are coupled to their plants, and this coupling can be identified using Conant's method of information theory analysis.

References

Ashby, W.R., 1956. An introduction to cybernetics. Chapman & Hall, London.

Bainbridge, L., 1974. Analysis of protocols from a process control task. In: Edwards, E., Lees, F. (Eds.), The Human Operator in Process Control. Taylor & Francis, London, pp. 146–158.

Bainbridge, L., 1996. Processes underlying human performance. In: Garland, D.J., Wise, J.A., Hopkin, V.D. (Eds.), Aviation Human Factors. Lawrence Erlbaum, Hillsdalei, NJ, in press.

Conant, R.C., 1976. Laws of information that govern systems. IEEE Transactions on Systems. Science and Cybernetics SMC-6, 240–255.

De Keyser, V., 1986. Cognitive development of process industry operators. In: Rasmussen, J., Duncan, K., LePlat, J. (Eds.), New Technology and Human Error. Wiley, Chichester.

Gentner, D., Stevens, A.L., 1983. Mental models. Lawrence Erlbaum, Hillsdale, NJ.

Green, T.R.G., 1990. Limited theories as a framework for human-computer interaction. In: Ackerman, D., Tauber, M.J. (Eds.), Mental Models and Human–Computer Interaction I. Elsevier, Amsterdam.

Johnson-Laird, P.N., 1983. Mental models. Harvard University Press, Cambridge, MA.

Moray, N., 1990. A Lattice theory approach to the structure of mental models. Philosophical Transactions of the Royal Society of London B 327, 577–583.

Moray, N., 1997. Models of models of ... mental models. In: Sheridan, T.B., van Lunteren, A. (Eds.), Perspectives of Control. Lawrence Erlbaum, Hillsdale, NJ, in press.

Moray, N., 1998. Mental models in theory and practice. In: Gopher, D. (Ed.), Attention and Performance XVII. Tel Aviv, in press.

Moray, N., Jamieson, G., Conant, R., 1996. Symposium at Annual meeting of the Human Factors and Ergonomics Society, September, Philadelphia. USA.

Rasmussen, J., 1986. Information Processing and Human-Machine Interaction: an Approach to Cognitive Engineering. North-Holland, Amsterdam.

Rasmussen, J., Pedersen, A.-M., Goodstein, L., 1995. Cognitive Engineering: concepts and applications. Wiley, New York.

Rutherford, A., Wilson, J.R., 1991. Models of mental models: an ergonomist-psychologist dialogue. In: Tauber, M.J., Ackermann, D. (Eds.), Mental Models and Human-Computer Interaction, vol. 2, Elsevier, Amsterdam.

Veldhuijzen, W., Stassen, H.G., 1977. The internal model concept: an application to modeling human control of large ships. Human Factors 19, 367–380.