The Role of Working Memory in Measuring Mental Models

José J. Cañas

Departamento de Psicología Experimental
Facultad de Psicología
Universidad de Granada
18071 Granada, Spain
delagado@goliat.ugr.es
http://www.ugr.es/~delagado/

Adoración Antolí

Departamento de Psicología Experimental
Escuela Universitaria de formación del profesorado de EGB
Universidad de Granada
52004 Melilla, Spain
antoli@platon.ugr.es

ABSTRACT
There is no agreement on what a mental model is and how to infer the mental model a person has. We are conducting research aimed at solving these problems by developing a model of Mental Model formation. Our basic hypothesis is that a Mental Model is a dynamic representation created in WM by combining information stored in LTM (the Conceptual Model) and characteristics extracted from the environment. Two experiments tested hypotheses derived from the model. Implications for individual and group research are discussed.

Keywords
Mental Models, Working Memory, Knowledge Elicitation.

INTRODUCTION
When a person learns to interact with a system it means she/he acquires a knowledge of its operation and of the structural relationships between its components. Researchers have called this knowledge the 'Mental Model' of the system (Moran, 1981). The existence of Mental Models and its importance during the interaction with the system have been demonstrated in numerous experiments (e.g., Kieras and Bovair, 1984; Cañas, Bajo and Gonzalvo, 1994). Research on group co-operation has also acknowledged the importance of mental models. When members of a group share similar and accurate mental models of group interaction, the group interacts more efficiently and performs more effectively (Cannon-Bowers et al, 1993; Tindale et al, 1996). The concept of Mental Model is particularly important for research on Team behaviour. Teams are groups in which the members work together on the same task to solve a common problem and there is no division of work responsibilities (Cannon-Bowers et al, 1993). To perform the task members of teams must develop a common knowledge that has been called Team Mental Model (Klimoski and Mohammed, 1994) .

Because of this, in the current investigations on design, learning of new interfaces, group co-operation, etc., it is common practice to try to infer what is the mental model that a person or group has.

The investigation of mental models is blocked at present by two problems. Firstly, a theoretical problem exists, which is reflected in the great confusion concerning the definition of a Mental Model. Though this problem has been stressed for a long time (Rouse and Morris, 1986) it still has not been solved satisfactorily. Secondly, a methodological problem exists that, in part, is a consequence of the definition problem. Although many methods have been proposed for inferring which model users has, all have been critiqued as being unreliable (Sasse, 1991).

The research project we are conducting is aimed at solving these definitional and methodological problems by developing a model of Mental Model formation. The basic hypothesis behind our model is that a Mental Model is a dynamic representation created in Working Memory (WM) by combining information stored in Long-Term Memory (LTM) and characteristics extracted from the environment. Methods that are proposed to infer the Mental Model that a person hold must consider that. The methods that are currently used require that the persons perform a task that is different from the real task they perform when interacting with the system or co-operating with other team members. We assume that the person being tested with these methods simulates the real task in her/his WM to perform the elicitation task. Therefore, what it is inferred is the result of this simulation.

This model could explain a common and unexpected result found in Mental Model literature: experts in a field who are supposed to have a good and similar knowledge seem to have different Mental Models

(Cooke and Schvaneveldt, 1987; Navarro, Cañas and Bajo, 1996). For example, Cooke and Schvaneveldt (1987) found that expert computer programmers had less similar mental representation of computer concepts than novices. This result could be explained from our model: even when two people share the same Conceptual Model, they can appear as having different Mental Models because when tested individually they execute different tasks in their WM.

The model could also have implications for the improvement of group discussion. The first step in co-operation is always the elaboration of a common knowledge space (Sauvagnac and Falzon, 1996). However, during group co-operation two people who shared the same Conceptual Model could be cued by the discussion process so that different portions of their knowledge are brought to their WM, making the above elaboration more difficult.

Theoretical Rationale

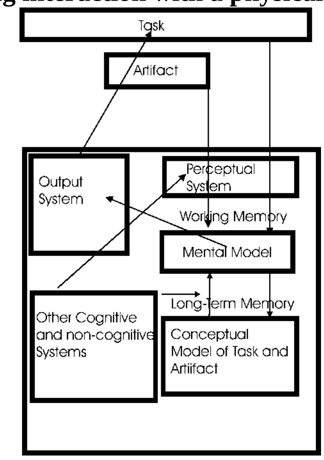
The definition problem

The definition problem is due to the fact that the term Mental Model has been used by researchers who work in different areas and study different tasks. Johnson-Laird has formulated his mental model definition in his attempt to explain the reasoning processes in tasks of syllogisms and language comprehension. To execute these tasks, a person forms in WM a mental representation of the world combining the information stored in LTM with the information of the task characteristics extracted by perceptual processes. This representation is called in this context Mental Model and it is, by its nature, dynamic. Though the information retrieved from LTM, the knowledge that a person has about the world, is important, Johnson-Laird (1983) gave greater importance to the information extracted by perceptual processes of the characteristics of the task (Rasmussen, 1990).

Research on the interaction with physical systems also considers that a person forms a representation in WM combining the knowledge stored in LTM, and the extracted information of the task characteristics (Gentner and Stevens, 1983). However, in this case the information stored in LTM that is relevant for these researchers is related to the knowledge of the structure and the operation of the physical system. Therefore the emphasis is on this representation, which is called Mental Model of the Physical System, and the efforts are placed on investigating how it is acquired and extracted from LTM.

It would be possible to unify both definitions if we make clear the distinction between the information stored in LTM, the Conceptual Model (Young, 1983), and the dynamic representation that is formed in WM combining the information stored in LTM and the extracted information from the environment (task), the Mental Model (see Figure 1). In this way, the unified definition of the Mental Model would emphasize the common characteristic of both definitions: the function of a Mental Model is to simulate the reality in WM.

Figure 1. The role and place of Mental Models during interaction with a physical system



The methodological problem

The dynamic nature of the mental models has an important consequence when considering the methods that we use to measure them. As Staggers and Norcio (1993) have indicated, if a Mental Model is a knowledge structure that is simulated in WM, we must speak of the Mental Model as a process and as the result of that process. When we measure the Mental Model of a person we are measuring the result of the simulation process. This result we take as reflection of the knowledge structure that it is stored in LTM. However, the simulation is accomplished selecting the part of the permanent knowledge that it is relevant for the task. That is to say, all the knowledge is not selected. The part that is selected will depend on the task, the context, the intentions, etc. It is also possible that the mental model as measured might be affected not only by selection but also by transformations performed on the knowledge in order to comply with the elicitation task

An additional problem with the knowledge elicitation methods that we use is that subjects are requested to accomplish a different task from the one that they would accomplish in the real situation. Therefore, when we measure the mental model with a knowledge elicitation task, the person simulates the real task in her/his WM and responses are given based on this simulation. Therefore we do not measure the knowledge stored in LTM but the knowledge which is put in WM depending on the elicitation method we use.

Take for example a knowledge elicitation task such as relationship judgements. This task has been widely used in interface design (Cooke, 1994**).** Subjects are asked to judge how related two components of the interface are. We know that objects could be compared along different dimensions. For instance, the USA and Cuba are very closely related when we compare them taking their geographical location. However, they are unrelated when we think of their political regimes (Tversky, 1977). We could assume that the dimension on which two components are judged to be related would depend on the task the subjects simulate in their WM.

Experimental Rationale

If we teach several people to interact with a system in such a way that all of them are capable of accomplishing any task that we give them, we can assume that they will all have stored the same knowledge (Conceptual Model) of the components of the system and of the relationships among them in their LTM. Then when we request the subject to perform a task in which we infer their Mental Model, e.g., to give us a relationship judgement between two components, they will simulate a task in their WM where those two components are implicated. Their judgement will reflect the relationship of the two components as a function of that task. However, in many occasions two components are implicated in several tasks and in each one the relationship between them will be different. Therefore, judgements will be made based on the particular task that subjects simulate on their WM.

In our experiments subjects learnt to interact with the system until they performed without any error. They also passed a declarative test in which we asked them a set of questions regarding all aspects of the system. Then they performed a knowledge elicitation task (relationship judgements) in which changes were introduced to affect the operation in WM. If performance in the elicitation task depends on these changes, we will be able to say that it was the result of what happens in WM and not of the Conceptual Model which the subjects have, the Conceptual Model being perfect and the same for all the subjects.

Experimental Procedure

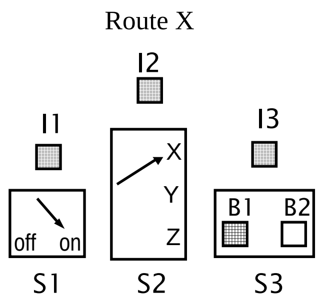
Description of the system

In our experiments, subjects learnt to operate a control panel device displayed on the computer screen. The device was a modified version of the one used by Kieras and Bovair (1984) consisting of switches, pushbuttons, and indicator lights (Figure 2). We said to the subjects that the device was a control panel for an electrical circuit. Their task consisted of making the current to flow from panel S1 to panel S3. They were instructed in the possible three action sequences that allowed them to complete the task:

1. Route X: press button ON in panel S1 (light I1 turned on); switch toggle switch in panel S2 to X (light I2 turned on; press button B1 in panel S3 (light I3 turned on) .
2. Route Y: press button ON in panel S1 (light I1 turned on); switch toggle switch in panel S2 to Y (light I2 remained off); press button B2 in panel S3 (light I3 turned on).
3. Route Z: press button On in panel S1 (light I1 turned on); switch toggle switch in panel S2 to Z (lights I2 and I3 turned on).

This system was sufficiently simple so that the subjects could learn it easily in a short period of time.

Figure 2. Device set for Task 1



The Learning Task

In the first phase of the experiments the subjects learnt to operate the system until they were capable of executing the three action sequences two times without committing any mistakes.

The Declarative Task

Then, subjects undertook a test in which they answered ten questions on the operation of the interface. The purpose of this test was to have a measure of the declarative knowledge that subjects had and to assure us that actually they had learnt to interact with the interface and that, therefore we could suppose that they had acquired the conceptual model of the system.

We eliminated from the experiments all subjects who failed in more than one question of the questionnaire.

The Elicitation Task

In the two experiments that we describe here, subjects completed a relationship judgements task concerning 11 interface items. Subjects were to assign ratings to pairs of items according to how related they thought the items were. The scale ranged from 1 to 6. A rating of one indicated that the items were unrelated, and a rating of six indicated a high degree of relatedness. The subjects were to indicate their responses by pressing the numbers corresponding to their ratings on the keyboard. The instructions emphasised that they should work fast, basing their ratings on their first impression of relatedness.

The judgement matrices were transformed into network representations using the Pathfinder algorithms (Schvaneveldt, 1990). Pathfinder is a graph-theoretic technique that derives network structures from proximity data. The Pathfinder algorithm takes proximity matrices and produces a network in which concepts are represented as nodes and relations between concepts are represented as links between nodes.

Experiment 1

In the relationship judgement task, item pairs are presented in a sequence. The subject sees a pair and judges it, then sees other pair and judges it, etc. The sequence in which the pairs are presented can have an effect on the judgements. According to our model, when a pair is presented the subject simulates in her/his WM a tasks in which that items pair intervene, and depending on the task that she/he simulates will issue her/his judgement. After issuing the judgement a trace of the simulated task remains in WM. When the following pair is presented, the subject resumes simulating a task in her/his WM. However, the task simulated to judge this second pair would depend on the task simulated for the first pair. For example, let us suppose that we are evaluating the mental model of two MSWORD experts. Then, we present two sequences, one to each expert:

1. Expert One: 1. "Print-File" ; 2. "Search-Edit"; 3. "Search-File"
2. Expert Two: 1. "Save-File"; 2. "Open-File; 3. "Search-File"

We could predict that expert Two would rate "Search-File" as being more related than expert One would. File is a menu and an object. The first sequence would lead the expert to think on the relation “is in the menu”. However, the second sequence points to the relation “things that you can do with a file”.

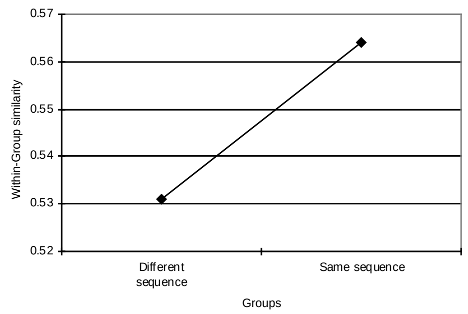
This task is similar to the classical similarity task. Researchers who have worked in the topic of similarity know that the context in which the similarity between two concepts is judged influences the judgement that a subject gives (Goldstone, Medin and Halberstadt, 1997). For example, Medin, Goldstone and Gentner (1993) had groups of subjects rating the similarity of *sunrise and sunset* and *sunrise and sunbeam*. In one condition of their experiment one group of subjects rated the pair sunrise-sunset and another group rated sunrise-sunbeam. In this condition sunrise-sunset was rated as less similar than sunrise-sunbeam. However, in other condition of the experiment, subjects rated both pairs simultaneously. In this condition, sunrise-sunset was rated as more similar than sunset-sunbeam. Medin et al argued that since sunset and sunrise are antonyms they are considered to have little similarity when they are judged in separation. However, this same characteristic causes them to considered closely related when they are judged in the context of sunrise-sunbeam. Therefore, the context influences how we judge the similarity of two concepts.

In this experiment we manipulated the presentation sequence of the system’s pairs of components which the subjects were giving relationship judgements on. For one group the sequence was random and different for each subject. For another group each subject were presented with the same random sequence. Our hypothesis was that the group that had the same sequence would show greater within-group similarity in their judgements than the group with different sequences.

Subjects

Fifty-six subjects participated in the experiment. After eliminating those subjects that did not pass the declarative test, results from forty-three subjects were analysed. Eighteen subjects performed the elicitation task in the same sequence condition and twenty-five in the different sequence condition.

Figure 3. Within-group similarity as a function of sequence of presentation



Results

Pathfinder analysis provided us with a measure of the similarity between two networks called C. This value reflects the degree to which the same node in the two graphs is surrounded by a similar set of nodes. A C value of 0 corresponds to two complementary graphs and a value of 1 corresponds to equal graphs.

We calculated the network similarity between all pair of subjects within one group. Then those C values (253 from the same-sequence group and 300 from the different-sequence groups) were submitted to a one-way ANOVA.

The results of this experiment showed that when subjects judged the concept pairs in the same sequence their ratings had more similar network representations than when they judged them in different sequences, $F(1,451) = 8.25$, $MSe = 0.013$, $p < 0.01$ (see Figure 3). Therefore, we can say that the task simulated to judge a pair left a trace in their WM that influenced the rating of the following pair.

Experiment 2

If the result of the first experiment could be explained by the trace left in WM of the task simulated to judge to a pair of items, we could eliminate that trace by introducing an interference task after subject’s response to one pair.

In this second experiment all subjects were presented with the same sequence of pairs. However, one group of subjects saw one number on the computer screen after they rated one pair of concepts. They had to count backward from that number for 2 seconds. Another group performed the task without counting backward.

We hypothesised that counting backward would erase the WM trace. Therefore two subjects that rate the items with this interference task would have less similar network representation than two subjects that perform the elicitation task without counting backward.

Subjects

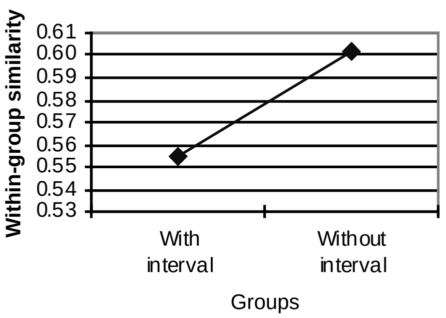
Forty-eight subjects participated in the experiment. After eliminating those subjects that did not pass the declarative test, results from forty subjects were analysed. Twenty subjects performed the elicitation with the retention interval and twenty without the interval.

Results

As in the first experiment, we calculated the network similarity between all pairs of subjects within each group. The C values (190 from each group) were submitted to an One-Way ANOVA.

As shown in Figure 4, the group that performed the elicitation task without counting backward showed more within-group similarity than the group that performed the task counting backward, $F(1,378) = 16.28$, $MSe = 0.013$, $p < 0.001$. Therefore, counting backward after rating a pair of items erased the trace that the simulated task had left in WM. Then, when a new pair of items had to be rated, the probability that two subjects simulated the same task decreased.

Figure 4. Within-group similarity with and without retention interval



CONCLUSION

The variables that we introduced to affect the contents WM had an effect on the [erformance of the subjects in the elicitation task. Though all subjects learnt to interact with the system perfectly and they were capable of answering questions that were put to them on this, their ratings of relationship were affected by what occurred during the elicitation task. Since the inferences that we make on the mental model are based on these judgements, our inferences would be affected by these variables.

Therefore, we claim that what we measure with our elicitation tasks is the content of WM, the Mental Model, and not the Conceptual Model stored in LTM. Research on Mental Model should be conducted with a model of the elicitation task that is used. This model would take into account the tasks that subjects simulate in their WM. Also, several elicitation tasks should be used in conjunction to get a better picture of the Conceptual Model.

Implications for Team Mental Model Research

Klimoski and Mohammed (1994) have called attention to the great confusion that exists about the contents of Team Mental Models. Researchers have used the concept to mean indifferently: artefacts that team members use; the task to be performed; a particular problem facing the team; the knowledge, skills, abilities, etc. relative to team functioning; and representations of environment events and projected future states. Our definition would include all these contents. A team Mental Model would be constructed by combining a conceptual model of artefacts, world and technical knowledge, skills and abilities (procedural knowledge stored in LTM), with information extracted from the environment. What is combined depend on the task to be performed by the team to solve a particular problem and it is used to project future states of the environment and artefacts.

Klimoski and Mohammed (1994) have also addressed the question of what does it mean to “share” a Mental Model. Team members could have identical representations, a distributed configuration of representations (no overlap), or a configuration of overlapping representations. Although most researchers would agree that a shared Mental Model does not imply identical representations (Cannon-Bowers et al., 1993), the question of how much overlapping would allow us to say that team members shared a common Mental Model remains open to research.

Our results suggested that we cannot answer this question with methods of measuring Mental Model that test subjects individually. We should think of the Mental Model considering the task that team members perform. Mental Models are created while team members interact with the system to reach a common goal (Orasanu and Salas, 1993). Therefore, we should develop methods that capture the representation that team members share when working on the co-operation task.

Problems of validity of measures have been also addressed by researchers on Team Mental Models (Hinsz, 1995). However, proposing a new method (e.g. belief association matrix) without considering the dynamic nature of Mental Model and that tests subjects individually would not solve the problem.

We need a method that captures the content of , let us say, the 'Collective Working Memory'. Our goal for the future should be precisely that one.

REFERENCES

Cañas, J. J., Bajo, M. T. and Gonzalvo, P. (1994). Mental Models and computer programming. *International Journal of Human-Computer Studies*, 40, 795-811.

Cannon-Bowers, J. A., Salas, E. and Converse, S. A. (1993). Shared mental models in expert team decision making. In Castellan (Ed.) *Individual and group decision making*. Hillsdale, NJ: LEA.

Cooke, N.J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, *41*, 801-849.

Cooke, N.J. and Schvaneveldt, R.W. (1987). Effects of computer programming on network representations of abstract programming concepts. *International Journal of Man-Machine Studies*, *29*, 407-427.

Gentner, D. and Stevens, A.L. (1983) *Mental Models*. Hillsdale: NJ: LEA.

Goldstone, R. L., Medin, D. L. and Halberstadt, J. (1997). Similarity in context. *Memory and Cognition*, 25, 237-255.

Hinsz, V. B. (1995) Mental models of groups as social systems: Considerations of specification and assessment. *Small Group Research*, 26, 200-233.

Johnson-Laird, P.N. (1983). *Mental Models. Cambridge: Cambridge University Press.*

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

Klimoski, R. and Mohammed, S. (1994). Team Mental Model: Construct or Metaphor?. *Journal of Management*, *20*, 2, 403-437.

Medin, D.L., Goldstone, R.L., and Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254-278.

Moran, T.P. (1981) An applied psychology of the user. *Computing Surveys*, 13, 1-11.

Navarro, R., Cañas, J.J. and Bajo, M.T. (1996). Pictorial aids in computer use. In T.R.G. Green, J.J. Cañas and C. Warren (eds). *Proceedings of the 8th European Conference on Cognitive Ergonomics*. Granada. EACE.

Orasanu, J. and Salas, E. (1993). Team decision making in complex environments. In G. Klein, J. Orasanu, R. Caldewood and C.E. Zambok (Eds.) *Decision making in action: Models and Methods. Norwood, NJ: Ablex.*

Rasmussen, J. (1990). Mental models and the control of action in complex environments. D. Ackermann and M.J. Tauber (Eds.) *Mental Models and Human-Computer Interaction*, vol. 1. Amsterdam: North-Holland.

Rouse, W. B. and Morris, N.M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Review, 3*, 349-363.

Sasse, M.A. (1991). How to t(r)ap users' mental models. In M.J. Tauber and D.Ackermann (Ed.) *Mental Models and Human-Computer Interaction*,vol. 2. Amsterdam: North-Holland.

Sauvagnac, C. and Falzon, P.(1996). Adapting models to case: A study of collective problem solving. In T.R.G. Green, J.J. Cañas and C. Warren (eds). *Proceedings of the 8th European Conference on Cognitive Ergonomics*. Granada. EACE.

Schvaneveldt, R.W. (1990). *Pathfinder Associative Networks: Studies in Knowledge Organization*. Norwood, NJ: Ablex Publishing.

Staggers, N. and Norcio, A.F. (1993). Mental models: concepts for human-computer interaction research. *International Journal of Man-machine studies*, *38*, 586-605.

Tindale, R.S., Smith, C.M., Thomas, L.S., Filkin,J. and Sheffey, S. (1996). Shared representations and assymetric social influences process in small groups. In E.H. Witte and J.H. Davis (Eds.) *Understanding Group Behavior: Consensual Action by Small Groups*. Hillsdale, NJ.: Erlbaum.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327-352.

Young, R. M. (1983). Surrogates and mappings: two kinds of conceptual models for interactive devices. In D. Gentner and A.L. Stevens (Eds.) *Mental Models*. Hillsdale: NJ: LEA.