An explanation of human errors based on environmental changes and problem solving strategies

José F. Quesada[1], José J. Cañas[2] and Adoración Antolí[3]

Department of Experimental Psychology
University of Granada
Campus de Cartuja
18071 Granada, Spain

[1]jose@geneura.ugr.es, [2]delagado@goliat.ugr.es, [3]antoli@platon.ugr.es
[2] http://www.ugr.es/~delagado

ABSTRACT

People that show good performance in problem solving tasks make also errors. Psychological theories of human error predict that those errors are to some extend the consequence of the difficulties that these people have to adapt to new environmental conditions. This paper describes two experiments and a research methodology designed to test this hypothesis.

Keywords
Human error, Complex Problem Solving, Microworlds, Transitions Between Actions, Firechief.

INTRODUCTION

Most psychological theories (i.e. Norman, 1981; Norman y Shallice, 1980; Rasmussen, 1983; Reason, 1990; Hollnagel, 1998) proposed to explain and predict human error share certain characteristics: (1) people loose conscious control when they increase their ability at performing a task; (2) there is a hierarchical structure (schemas, semantic network or control levels) in which higher levels include, organize and control lower levels; (3) practice and elaboration let to a representation which hides the process details that can led to a lack of flexibility. (4) there is a trade - off between quick, fluid actions and controlled, flexible actions.

These psychological theories seems to agree on the idea that in order to avoid an human error you need to realize that the situation has changed to be able to `log out´ the automatic processing mode and come into the controlled processing mode. To detect the situation change and the necessity of a non - routine response, it is necessary to come into a higher level of attentional control, where you access the new situation and plan the action to be taken. You need to perceive environmental cues in a different way, reinterpreting them. How the person represent the task and the set of strategies employed to deal with it determine how easily she/he shifts attention to the new environmental conditions.

For example, Rasmussen (1983) proposed a theory that can be used to frame this idea. He distinguished three level of processing: (1) skill based level, for activities we do in an automatic way, (2) rule based level, for situations in which our experience give us a response in a known situation, and (3) knowledge based level, for new situation in which there are no rules and we need to plan a different response. Within this framework, practice in a task lets to more automatic activities at the rule and skill based levels. However, when the task conditions change a person must do activities accomplished at the knowledge based level to adapt to those new conditions. In general, when a person is skillful would have problems to detect the environmental changes and to change her/his strategy to adapt to those changes.

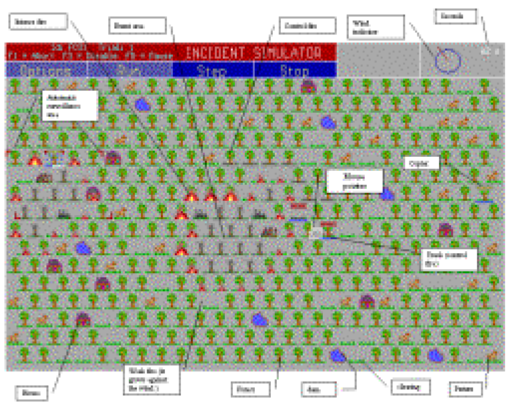
The present research was designed to test and pursue this hypothesis by developing an experimental program in which people learn to perform a complex problem solving task. The conditions in the task remains constant for some time and people have the opportunity to develop strategies to deal with them. At one point, those conditions change and the effect of that change on their performance is observed. We hypothesized that performance would be impaired when the environmental change affects the particular strategy people develop.

To prove this hypothesis it is necessary to elicit planning mistakes or errors related to changes from the inferior levels to knowledge based level. This is only possible in problem solving tasks where planning and decision making abilities are more important for optimum performance than perceptual and motor abilities.

Microworlds are complex problem solving tasks that are the appropriate research environment to test this hypotheses. Microworlds are based on simulation of task changing dynamically and they are prepared to reproduce the important characteristic of the real situations (the state of the problem changes autonomously as a consequence of the actions of the subject and the decisions must be taken in real time) but letting opened the possibility to manipulation and experimental control.

Our work is based on *Firechief*, a microworld generation program created by Omodei and cols. (Omodei and Wearing, 1995),

Figure 1: Screenshot of Firechief (Omodei and Wearing, 1993).



In *Firechief* participants find a screen that simulates a forest where the fire is spreading. Their task is to extinguish the fire as soon as possible. In order to do so, they can use helicopters and trucks (each one with particular characteristics) that can be controlled by mouse movements and key presses. The different cells (see figure 1) have different values of flammability and points: houses are more valuable than forests, for example. The participant's mission is to keep saved as more forest's extension as possible, but preserving more the most valuable cells and preventing the trucks to be burnt. Helicopters move faster and drop more water than trucks, but the latter can make control fires. Trucks are unable to extinguish certain fires, and they will be destroyed if they are sent there. The fire is more intense and spreads faster in the direction the wind blows. Participants can see a window with their overall performance score at the end of the trial. This figure is calculated adding every safe cells and subtracting the value of the trucks burnt. The task is complex and participants feels interested from the beginning to the end. At the same time, is possible to control experimentally every single feature of the system, and to prepare experimental situations for checking a wide variety of hypotheses.

There are four commands that are used to control the movement and functions of the appliances: (1) Drop water on the current landscape; (2) start a control fire (trucks only) on the current landscape segment; (3) move an appliance to a specified landscape segment; (4) search in a specified portion of the total landscape area. Commands are given by first selecting the desired vehicle (by moving the mouse cursor into the landscape segment containing it) and then pressing a key in the keyboard. At the end of each trial, the program saves the command sequence that the participant issues in that trial.

In the two experiments described in this paper participants performed the task in a situation in which the environmental conditions remained constant for some time. The wind was blowing always toward the East in the first sixteen trials. In the last four trials the wind change direction. If our hypothesis was correct, the participant with better performance would be those that experience a decrease in their performance, while those with worse performance would adapted to the change. However, this effect would depend on the problem solving strategy adopted by each participant. There could be strategies that allow participants to keep a good performance level in spite of the change, while other strategies will prevent them to adapt to the new situation. Since the fire spreads faster in the direction the wind blows, the strategies most affected by the change would be those that are based on predictions of where the fire will spread.

EXPERIMENT 1

In the area of traditional Problem Solving, where it exists a limited problem space, a usually well-defined goal, and an only way of reaching the goal in the smaller number of steps is relatively easy to identify the strategies that a person adopts. However, in most Complex Problem Solving tasks this identification is more difficult since it does not exist the optimum strategy, and furthermore the protocols of the participant are so wide in data that they have probably been produced by more than one simple strategy at the same time.

For this reason we have devised a method of analysis to identify the strategy that the participants adopted in our experiment. The method consists basically in comparing a participant behavior with that of a simulated person that would adopt a hypothetical strategy.

The protocol output by Firechief at the end of each trial contained the sequence of commands that a participant issued during that trial (i.e. move a truck, drop water). Those commands could be used to construct a **matrix of transitions between actions**. Rows and columns in the matrix represent the command and the cells contain the number of times that one command follows another. This matrix contains important information about problem solving strategies since transitions between actions reflect how a person issue the commands (Howie and Vicente, 1998). Therefore, our method for inferring participants- strategies is based on the analysis of matrices of transitions between actions and is done following a number of steps:

1. Construct empirical matrices of transitions between actions

For each trial of each participant, we obtained a protocol file where all the actions were registered, in a temporal sequence. Afterwards, we extracted every transition between two actions, and put them in a symmetrical matrix with size equal to the number of possible actions.

2. Design a set of theoretical, simple strategies based on a task analysis.

The *Firechief* program has a simulation module that allows the implementation of problem solving strategies. This module permits to introduce code in Pascal, and provides a function library to facilitate the design of those strategies. When recompiled, the simulation module takes the programmed strategy and accomplishes the task as a participant that has adopted it would. Therefore, the program generates equivalent protocol files to those that would be generated when a human participant accomplished the task. That is to say, we could have a protocol from a simulated participant that has performed the task with a single hypothetical strategy. And from this protocol we could also obtain a matrix of transitions between actions.

At this state of our research we have devised three possible strategies participants might be using:

Move and drop water (SLRN): People using this strategy move appliances to the closest unattended fires and drop water there. Trucks are not sent to fires too fierce where they could be destroyed

Control Fires (SCON): This strategy only can be used by trucks. It finds the closest fire and then sends an appliance to deliberately light small fires in a location two segments away from it in a randomly chosen direction. Before that, the algorithm checks that the location is unburned, unoccupied and no burning.

Setting up an automatic vigilance area (MISC): An automatic vigilance area is a rectangular zone that the participant can draw in the map after selecting an appliance. If the appliance has water and a fire is found in an adjacent segment within the area, the appliance moves into that segment and automatically drops water on the fire. If the appliance is empty of water and a dam is found, automatically fills with water. If none of these conditions apply, the appliance will move one segment in a random direction and the above search process repeated.

These strategies were simple, easily distinguished and orthogonal in the sense that the matrices of transitions generated by them did not correlate.

3. Correlate empirical and theoretical matrices.

When we introduced the theoretical matrices as predictors of participants' matrices of transitions between actions it was possible to identify which one of them was used by her/him. Then, to evaluate the possibility that one participant had used one particular strategy we calculated the similarity between her/his empirical matrix and that obtained from simulating the strategy. A significant correlation between those two matrices mean that to some extent that strategy was responsible for her/his performance. Therefore, we performed a multiple regression analysis with the empirical matrix as the dependent variable and the simulated strategies as the predictors. The Betas in the analysis represented the partial correlation between the strategies and the performance of the participant. For example, if one participant adopted the SLRN and SCON strategies in one trial, we should find significant betas (alpha = 0.05) for the matrices representing these strategies.

4. Classify participants according to which strategy they used.

We built a rectangular matrix representing each participants' similarities with the theoretical strategies. Rows represented the participants and the three columns represented the strategies. One cell in the matrix had a value of 1 if that participant used the corresponding strategy. Finally, we performed a cluster analysis on that matrix to group participants with similar strategies.

Therefore, what we have after applying this method is a classification of participants into groups based on the similarity of their strategies. This grouping could be used to as a quasi-experimental independent variable in a factorial design to evaluate its interaction with other manipulated independent variables.

METHOD

Participants

Thirty-seven students at the University of Granada participated in the experiment as part of class requirement.

Procedure

Participants were asked to play 20 trials, where 16 of them kept wind direction constant and the last 4 had variable wind conditions (wind changed from east to west slowly). They did not know that beforehand.

RESULTS

We did the analysis using only the 4 trials before the wind change and the 4 trials in which the change was experienced.

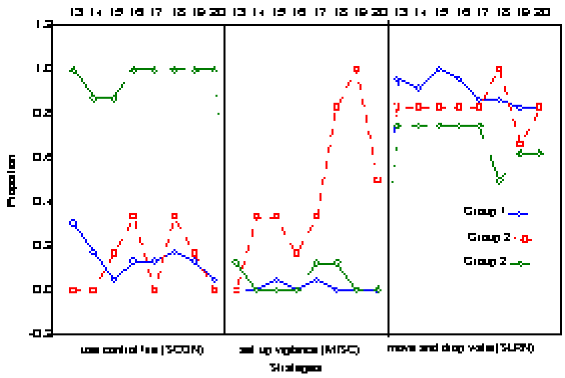
In order to group participants in accordance with their strategies we did a cluster analysis with the data of significant correlation between their strategies and the theoretical ones in each trial, coded as zero and one.

The cluster analysis that gave better discriminant results threw 3 groups (see Figure 2). The groups had unequal number of participants: 23 participants belonged to group 1, 6 belonged to group 2 and 8 belonged to group 3. The grouping could be interpreted as follows:

Group 1 Participants whose strategies rarely correlates with SCON or MISC. They mainly adopt a strategy similar to SLRN.
Group 2 Participants who use mainly SLRN, but react to the wind change increasing MISC- like strategies.
Group 3 These participants use neatly SCON in every single trial, though they also use SLRN and MISC a little bit.

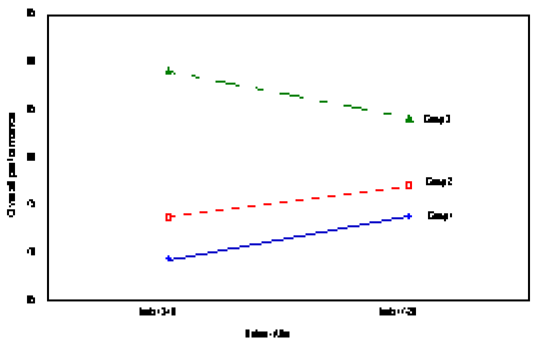
To support this interpretation with statistic arguments, we did three analysis of variance (ANOVA), one for each strategy, with three independent variables, groups, trials and before-after change. The former was a between groups variable and the other two were within-subject variables. In this analysis, the SLRN strategy did not discriminate between groups neither showed significant effects in trials or in wind change. The SCON strategy discriminated between groups, $F(2,39) = 73.32$, $Mse = .2357$, $p<0.01$, and showed an interaction between trials and before-after change, $F(3,102)=13.84$, $Mse = .0719$, $p < 0.05$. This interaction express little variations in groups 1 and 2 that are not relevant to our objectives. Finally, the MISC strategy also discriminated between groups, $F(2,34)= 63.218$, $Mse = .0059$, $p<0.01$, and showed an effect of before-after change, $F(1,34)= 40.075$, $Mse=0.357$, $p<0.01$, and trials, $F(3,102)=5.192$, $Mse = .0457$, $p<0.01$. The interaction of groups by before-after was also significant, $F(2,34)=28.717$, $Mse=1.025$, $p<0.01$ and an interaction of groups by trial, $F(6,102) = 6.784$, $p < 0.01$. These results can be interpreted saying that the only group which increase their use of MISC is group 2, and they do this after the change. The LSD tests showed that the interaction was not significant for groups 1 and 3.

Figure 2: Proportion of times correlation between theoretical and empirical strategies becomes significant (trials 13 - 20)



Once we have grouped participants, we used these groups to perform an ANOVA using the overall performance scores [1] as the dependent variable. The independent variables were groups (between groups, 3 levels), trials (within-subject, 4 levels) and before-after change (within-subject, 2 levels). Results showed a significant effect of group, $F(2,34) = 6.84$, $Mse = 786.71$, $p < 0.01$ (observed power: .897) and a significant effect of the interaction between group and before-after change, $F(2,34) = 3.52$, $Mse = 155.04$, $p < 0.05$ (observed power:.618, see figure 3).

Figure 3: Group by before - after wind change interaction.



Thus, the different group formed in accordance with the differential use of strategies, were also different in overall performance. Planned comparisons showed that group 3 had superior overall performance scores than group 1 and 2 .$F(1,34)=9.84$, $mse=786.713$ $p< 0.05$. Group 1 and 2 were not significantly different, $F<1$. The wind change affects the groups differently. The change worsens group 3 performance, who were the best participants, although this difference did not reach significance $F(1,34) = 2.58$, $Mse = 155.036$, $p=0.116$. In group 1 participants gets better, $F(1,34)=6.961$, $Mse = 155.036$, $p<0.01$. In group 2 participants did not experience any change $F<1$.

The variable before-after difference in significance was probably owed to the bigger size of group 1(n=23), which declared significant an increment of the same magnitude as the group 2 experienced (n= 6).

EXPERIMENT 2

Experiment 2 was designed to replicate experiment 1, but with an important modification. We decided to eliminate the physical possibility of doing automatic vigilance area, and, thus, of doing a strategy similar to MISC. We had two reasons for doing that: (1) the most interesting results in experiment 1 were on group 3, that is, people who used control fires (SCON). This strategy could be responsible for the decrease in performance, so we were interested in forcing participants to use it. A simple, indirect way to do so is by eliminating one of their alternatives (MISC). (2) The MISC strategy gave participants the opportunity of delegating in the automatic search for a fire good amount of their work.

---

[1] Defined as the sum of all cells that remain unscathed subtracting the value of all burnt trucks. It was expressed as a percentage of the total area.

METHOD

Participants

Twenty-four students at the University of Granada participated in the experiment as part of class requirement.

Procedure

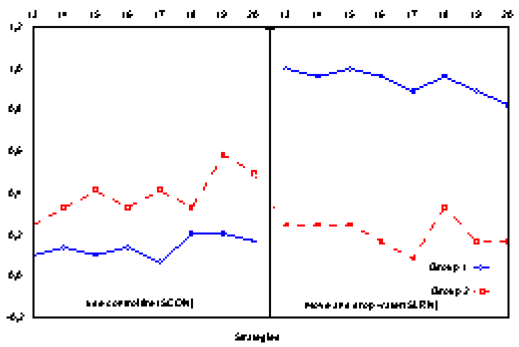
Like in experiment 1, participants were asked to play 20 trials, where 16 of them kept wind direction constant and the last 4 had variable wind conditions (wind changed from east to west slowly). They did not know that beforehand.

The only difference with experiment 1 was that participants were not allowed to use automatic vigilance areas. That option was unable during the trials, and it was omitted in the instructions that participants read.

RESULTS

Again, we did the analysis using only the 4 trials before the wind change and the 4 trials in which the change was experienced. To group participants in accordance with their strategies we did a cluster analysis with the data of significant correlations between their strategies and the theoretical ones in each trial, coded as zero and one.

Figure 4: Proportion of times a correlation between theoretical and empirical strategies becomes significant (trials 13 - 20)



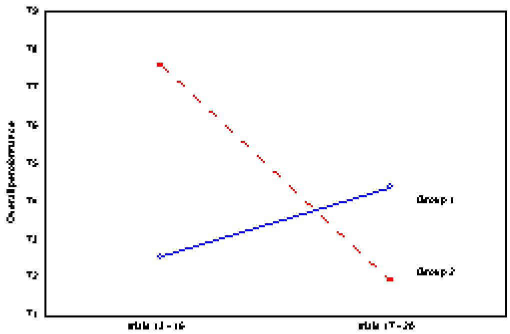
The cluster analysis that gave better discriminant results threw 2 groups of 12 participants each. The grouping could be interpreted as follows (see figure 4):

Group 1: Participants whose strategies significantly correlated with SLRN, that is, they normally moved their appliances and dropped water.

Group 2: Participants whose strategies were more based on SCON (used control fire) than in SLRN.

To interpret cluster analysis results and get a statistical reinforcement to these statements, we did two ANOVAs,, one for each theoretical strategy. We used three independent variables, groups, trials, and before-after change. The results showed that both SCON and SLRN discriminated between groups, $F(1,23) = 5.117$, $Mse = .715$, $p < 0.05$, and $F(1,23) = 191.07$, $Mse = .1407$, $p < 0.01$ respectively.

Figure 5: Group by before - after wind change interaction.



Using these 2 groups created in cluster analysis, we did an ANOVA using overall performance scores as dependent variable. The independent variables were Groups (between groups, 2 levels), Trials (within- subject, 4 levels) and Before-After change (within- subject, 2 levels).

Results (see Figure 5) showed that group 2 which started with a performance much better than group 1, was affected by the change. They performed worse than group 1 in the second level of the variable Before-After, as showed the significant effect of the interaction between Groups and Before-After Change,, $F(1,22)= 5.5156$, $Mse = 121.4$, $p < 0.05$ (observed power, .612). The interaction between Before-After Change and Trials was also significant $F(3, 66) = 2.96$, $Mse = 57.8$, $p <0.05$, but this was not relevant to our objectives.

This results indicated that, as in experiment 1, the group who used control fire commands was the best before the change and the worst after it. Thus, participants who employed a strategy similar to SCON had a tendency to be affected by the change in the wind direction. Participants in Group 1, whose performance was under group 2 before the change, kept a constant improvement in spite of the change, and finished the experiment with a performance better than group 1.

DISCUSSION

The results from both experiments confirmed our hypothesis. After some time performing a task a person acquired knowledge about the environmental conditions and developed problem solving strategies appropriate to those conditions. When the environment changed some people were affected by that change and showed a decrease in performance. Those people used the SCON strategy and were performing better than the other before the change.

In order to explain these results we need to consider what each of these strategies required from participants' cognitive processing.

The SRLN strategy simply consists on moving appliances to the closest fire and dropping water there. It does not require making predictions on where a new fire will start. Therefore, the direction of the wind did not matter much for reaching a good performance. Something similar happens with the MISC strategy. A person using it defines an area around one already started fire and sends an appliance there to search for it. Therefore, it does not require any prediction at all. However, a person using the SCON strategy need to predict in which direction the wind will be blowing and issue a control fire according to that prediction. The fire will spread depending on the wind direction. Therefore, the person selects the location to issue a control fires depending on that prediction.

Therefore, it seems to be an easy explanation for the relation between the environmental change and the problem solving strategies. However, it remains to be explained why participants that used the SCON strategy performed better before the change.

The strategies simulated in these two experiments were very simple and need to be elaborated to explore this hypothesis within a general model of decision making in Complex Problem Solving tasks. Therefore, in order to find a complete explanation for our results, our next step in this research program will be to make a more complete model of how people make predictions and design possible strategies based on that model.

REFERENCES

Hollnagel (1998). *Cognitive reliability and error analysis method*. Elsevier . Tokyo, Japan.

Howie and Vicente (1998). Measures of operator performance in complex, dynamic, microwolds: advancing the state of the art. *Ergonomics*, *81*, 485-500.

Norman, D. A. and Shallice, T. (1980). Attention to action: Willed and automatic Control of behavior,. In R .J. Davidson, G. E. Schwartz and D. Shapiro, Consciousness and self regulation, pg. 1 - 15, New York. Plenum times.

Norman, D. (1981). Categorization of action slips, Psychological review, volume 88, pg. 1 - 15.

Omodei , M.M. and Wearing, A.J. (1995). The FireChief microworld generating program: An ilustration of computer-simulated microworlds as an experimental paradigm for studying complex decision-making behavior. *Behavior Research Methods, Instruments and Computers*, *27*, 303-316.

Rasmussen, J. (1983) Skills, Rules and Knowledge: Signals, Signs, and symbols and other distinctions IEEE transactions on human, systems and cybernetics volume 3,pg. 653 - 688.

Reason, J. (1990) Human error. Cambridge University Press, New York.