ELSEVIER

Reliability Engineering and System Safety 53 (1996) 319–325

0951-8320/96/$15.00

PII: S0951-8320(96)00042-7

Trust between man and machine in a teleoperation system

I. Dassonville, D. Jolly & A. M. Desodt

Centre d'Automatique de Lille, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France

The work we present deals with the trust of man in a teleoperation system. Trust is important because it is linked to stress which modifies human reliability. We are trying to quantify trust. In this paper, we'll present the theory of trust in relationships, and its extension for a man–machine system. Then, we explain the links between trust and human reliability. Then, we introduce our experimental process and the first results concerning self-confidence. © 1996 Elsevier Science Limited.

1 INTRODUCTION

A teleoperation system is a system composed using a simplified approach by:

1. A master universe
2. A slave universe
3. A space between these two universes.

In the master universe, we find the operator and a master arm. In the slave universe, we find the slave arm and some sensors. In the space between these two universes, transmissions of information, computers, and decision support systems are found.

Such a system can be driven by three different modes: the automatic mode where the computer controls the system, the manual mode where the operator drives the system alone, and the semi-automatic mode, where some 'degrees of freedom' are driven by the computer, and others by the operator.

In fact, different parameters are important in the control of such a system: vigilance, workload, trust... All these parameters modify the quality of the relation between man and machine as told by Sheridan.[1]

Trust between man and machine is a particularly imprecise and uncertain variable, because we can't directly use a tool to measure it, and this variable is formed by two components: the static part and the dynamic part. The dynamic one is the most difficult to study, and will be studied in a dynamic way, it's the reason why we decided to use theories convenient for the specificity of our problem.

2 MODELLING UNCERTAINTY

According to the dictionary, uncertainty is the state of what is uncertain, unpredictable or badly known. As a matter of fact, the uncertainty is defined as opposed to what is certain, foreseeable, or known. Today, many theories aim to fulfil the representation of uncertainty: theory of possibilities, theory of fuzzy logic, theory of Dempster–Shafer, theory of fuzzy measures.

If we imagine a question whose answer is unknown, we know that the right answer to this question is in a specific set of thinkable solutions: the set Ω, called 'frame of discernment'. With different sources of information, the likelihood of each solution A (or set of solutions) can be estimated: we can assign a degree of belief m (called mass or bpm: basic probability mass). The set of m is called bpa: basic probability assignment. With those masses, Bel (Belief) and Pl (Plausibility) can be defined.[2]

$$\mathrm{P1(A)} = \sum_{F \cap A \neq \varnothing} \mathrm{m(F)}$$

$$\mathrm{Bel(A)} = \sum_{F \subset A} \mathrm{mF}.$$

In fact, the interest of belief and plausibility functions resides in the fact that this theory manages the uncertainty and then allows us to have different kinds of sources of information.

The management of uncertainty is possible thanks to a set of rules. A rule uses different beliefs to create a sole belief. For example, a particular rule is that of Dempster: if the frame of discernment is Ω, and if there are two sources of information, modelled by their BPA $m1$ (for A) and $m2$ (for B), the rule creating a new belief (represented by its bpa m) is:

$$\forall C \subset \Omega, C \neq \varnothing$$

$$m(C) = \frac{\sum\limits_{Ai \cap Bj + C} m1(Ai)m2(Bj)}{\sum\limits_{Ai \cap Bj \neq \varnothing} m1(Ai)m2(Bj)}.$$

Then, starting from two opinions, the rule permits us to have a 'summary' of those opinions. With such an approach, the increase of the knowledge is allowed. In our dynamic process which will measure trust, such a theory is useful in taking into account new events occurring in the system.[3] Other tools will appear in this study, such as fuzzy integrals often used in man machine systems.

3 TRUST BETWEEN MAN AND MACHINE AND ITS LINK WITH RELIABILITY

3.1 Trust

When we consider the definition of trust (in relationships) in the literature,[4,5] we realise that a lot of parameters compose and influence this variable. In fact, this particular aspect is explained by the multidimensional character of trust. Two sociologists explain it: Barber (1983) and Rempel (1985). The first explains it by a compromise between three kinds of hope linked to the three dimensions: (1) persistence of natural and moral laws; (2) technically competent performance; and (3) fiduciary responsibility. The other one explains it by three dimensions: (1) predictability; (2) dependability; and (3) faith.

In fact, in a context of a man–machine relation, Bonnie Muir in Ref. 6 demonstrates that Barber's dimensions of persistence of natural laws, technical competence, and fiduciary responsibility, could be identified, respectively, with the reliability of machines, the predictability of a machine's behaviour, and with the degree to which the behaviour of machines is in accordance with their designer's stated intentions. She suggested also that the dimensions of trust suggested by Remple could be interpreted in the framework of machine performance. She also found a strong relation between an operator's judgment of the < Competence > of a machine, defined as the extent to which it does its job correctly, and the amount of trust placed in the machine. In addition, her work on trust, in a supervisory context, has demonstrated that the variable trust is an authentic causal variable in a man–machine system.

So, we can work with bases of the theory of trust in our particular context. As a matter of fact, it is a particular context because the operator must give his trust in the system, and, at the same time, he is in the system! He is implicated in the control of the system. So, he must give his opinion on himself in the system.

As the theory of trust is suitable for a relation of trust of a man in a man, and for a man in a machine, this theory will also suitable for our problem (Fig. 1).

3.2 Links between trust and human reliability

Trust of the operator in a man machine system is a complex phenomenon because it is a parameter modifying human reliability and, so, the safety of the system and, at the same time, it is modified by the reliability of the components of the system. Faults affect trust and behaviour. Study of human reliability supposes the analysis of the characteristics of the work situation and of the operator. These two parameters are interdependent and the links between them must be analyzed carefully, particularly concerning perturbations conditioning the work of the operators.

A particular point must be emphasised in this context: stress of the human operator.

Stress is a global interactive process integrating stimuli given by the task requirement, responses that are the working of the operators, processes of perception and evaluation of the situation.

Concerning the links between trust and stress some indications can be found in the paper of D. Lhuilier.[7] This paper relates experiments made on simulation of control of a complex system during incidental situations. The experiment induces an uncertainty between the task requirements and the impossibility of satisfying them. A task can be felt as putting the operator to the test. The fear of this putting to the test seems to be increased when the operator has an expectation level corresponding to a weak level of expected success, that is to say, when the ideas he has give him no or little self confidence. The idea of oneself seems to be linked with two parameters: experiments of success or failure that the operator has had in the past during a similar task, and the global idea of himself. The operators, when speaking about their work, say, if they have a feeling of success, that their trust and insurance increase. On the contrary, after failure, they have doubts about themselves, and on their ability to control the situation. The same reaction characterises not only the operator alone but also the whole team. So, self-confidence will directly influence stress and at the same time the reliability of the system influences self confidence.

The exploration of mechanism and of consequences of emotional pressure is necessary to study human reliability, for it can be at the source of mistakes caused by a disruption or an inhibition of mind processes. This stress leads to mistakes in the management of accidental situation. Moreover, a defence reaction to the stress can change the

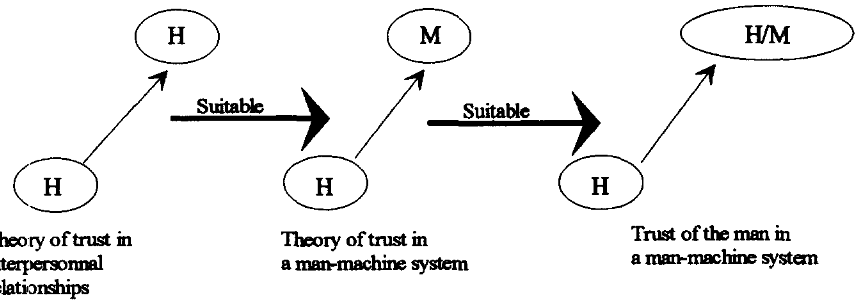

Fig. 1. From trust in interpersonal relationships to trust in a man-machine system.

evaluation of the situation so that the operator cannot adapt a good strategy to cope with the faults. Lee[10] found that the first experience of unreliability in a system which has been very reliable has a devastating effect on some operators. They initially assume that the error must be due to something they have done, especially if the control mode is predominately manual. Only when errors appear repeatedly do they accept the fact that the system is at fault. In fact, the significance of the situation being fundamental in the birth of stress, cognitive aspects can play a defensive function in modifying the judgement on the situation. So, mistakes are generated because of a defence reaction which deforms reality. In any case, erroneous commands misdirect the robot and may lead to potentially life threatening situations.[8]

4 OUR EXPERIMENTAL PROCESS

4.1 Description of the simulated system

Before the presentation of our method to follow trust in a system realising the simulation of a teleoperation system, we'll present this experimental process in order to understand the method used. We work on simulation: it is necessary for the portability of the experiment. As in every experimentation on human factors, preliminary tests on population are necessary so as to validate the protocol and so we had to bring our system in various places: this would be impossible with a real teleoperation system. The idea has been to create a simulation as near as possible to the real phenomena.

As we saw, a teleoperation system is composed of three important entities (master and slave universes and the space between the two). To simulate the master universe, we take an operator who manipulates a joystick (simulation of the master arm). To simulate space between the two universes: a P.C., to simulate the slave universe: a cursor (to simulate the slave arm) which executes joystick's orders, and to simulate the task: an operator who conducts as fast as possible, using a joystick, the cursor from a point A to a point B, within a trajectory, without mistake. A mistake is made when the cursor touches the side of the trajectory.

In the right top corner (Fig. 2), a window is reserved for the profile of the trajectory. In fact, the trajectory is not always planar. There are some slopes (up and down). So the cursor undergoes some acceleration and some deceleration. The system can be driven in 3 modes: (1) the automatic mode: where the cursor executes the trajectory alone (the computer controls the trajectory), (2) the manual mode: the joystick gives orders to the cursor (the operator is alone to control the trajectory), (3) the semi-automatic mode: where the computer takes into account the acceleration and deceleration of the cursor. In this last mode, for the operator, the trajectory seems to be plane.

4.2 The faults which appear during the simulation

The variable trust can only be studied thanks to the creation of situations that make it change. We create such situation of distrust, caused by a lack of reliability of the system, thanks to some failures placed in the simulation. We simulate a real teleoperation system, having as a slave arm a PUMA robot. We studied the failures that could occur in such a system so as to create the simulation. You can find in Ref. 9 a good review on robot reliability, which

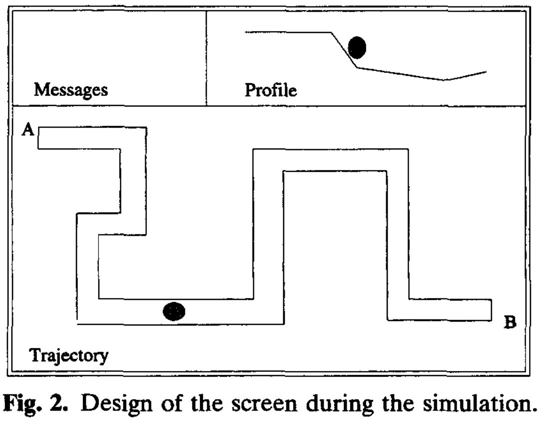

Fig. 2. Design of the screen during the simulation.

describes all the failures that can be encountered with such robots.

So, the distrust is created from some real faults that could occur in a real teleoperation system. We can quote three kind of situations:

- distrust generated by a problem in the master universe
- distrust caused by a problem in the slave universe
- distrust created by a problem on the control system placed between the two universes.

4.2.1 Concerning the master universe
The master universe is composed of a joystick, simulating the master arm, and the human operator. The fault created seems to affect the joystick working: it is a simulation of the deterioration of one degree of freedom of the joystick. The operator has the impression that one of the degrees of freedom does not respond or responds very badly to the orders he gives.

4.2.2 Concerning the slave universe
The distrust is caused by the apparition during the experimentation of an underlying fault, undetectable directly. It is a failure that causes the robot to quickly accelerate into wild motions (in real life these failures may cause considerable damage). In our simulation a sudden acceleration or deceleration of the cursor bring failures in the experiment.

4.2.3 Failure in the universe separating the slave and master arm
We create a deterioration of the image that the operator has in front of him: a part of the image gets out of shape.

This image concerns the information feedback whose control should be insured by the computer linking the slave and master universes. This could be in a real situation a hardware of software failure in the information control system which would generate such a problem.

The anomalies are pre-programmed and always appear at the same time (for each operator). These anomalies will incite the operator to change the driving mode (and will permit us to follow the variation of trust). For the operator, the anomalies appear to be random.

5 THE STUDY OF THE EVOLUTION OF TRUST

5.1 The tools

In order to follow trust, which is a psychological variable, we decided to take psychological tools. Different tools are possible, but we chose to use two particular ones: questionnaires, and scales. In fact, before executing the task, the operator must answer a series of questions, and must give his opinion on the difficulty of answering them.

For example:
I'm afraid when I'm alone with some people I don't know:
a: often b: now and then c: rarely
Give your feeling:

10	- - - - 0
It's easy to answer	It's very difficult to answer

The aim of the first questionnaire is to examine three components of personality which seems to be very important for our experimental research: (1) self-confidence, (2) trust in others, (3) trust in machines, in general. It appears that these three components must be studied separately because in recent papers[10,11] of Lee & Moray, self confidence and trust are parameters that influence mutually. Then, before starting the experimental process, we must qualify the *a priori* trust of the man. To determine the *a priori* trust, we use factors of the 'Test 16 Pf' from Cattel.[12] This is a personality test which is known and tested (knowledge of probabilities of this test). So, for self-confidence and trust in others, we use this test. For trust in machines, in general, we didn't find any existing test.

To follow trust during the task, we use scales (to follow the dynamic part of the variable trust). The aim of those scales is to qualify (1) trust in the joystick, (2) trust in the movement of the cursor, (3) trust in the PC, (4) trust in the programming, and (5) trust in the whole system.

For example, to evaluate trust in the joystick, the operator must give his opinion on: (1) the reliability of the joystick, (2) the performance of the joystick, (3) the predictability, and his feeling on the difficulty to answer.

These scales appear in each essay at the same time from one operator to another.

At the end of the experimental process, we give scales for the last time and a final questionnaire based upon the same principle as the first one to evaluate the operator's trust in the teleoperation system.

A questionnaire brings a series of answers notified a or b or c, and scales on the difficulty of answering for the operator, shown in Fig. 3.

5.2 Study of the population concerned by the experiments

The questionnaire submitted at the beginning of the experiment should allow the evaluation of:

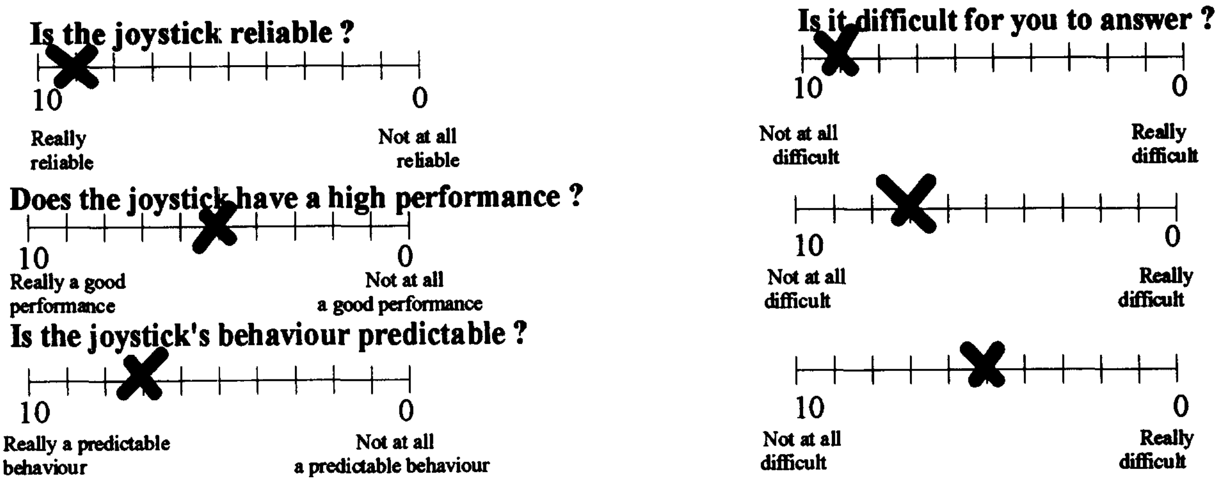

Fig. 3. An example of operator's response.

- self confidence (1)
- interpersonal trust (2)
- trust in the machines (3).

It is composed of questions issuing from the <16 Pf test > well known in psychological world, concerning items (1) and (2). However, there existed no questionnaire concerning item (3), so we had to create one.

In order to validate this test, we had to test it. On the advice of psychologists, we submitted it to two populations:

- one population doing literature studies
- one population doing scientific studies.

The questionnaire is validated if the response curves obey a Gaussian distribution, and if the responses on the questionnaire on trust of machines allowed us to distinguish between the two populations.

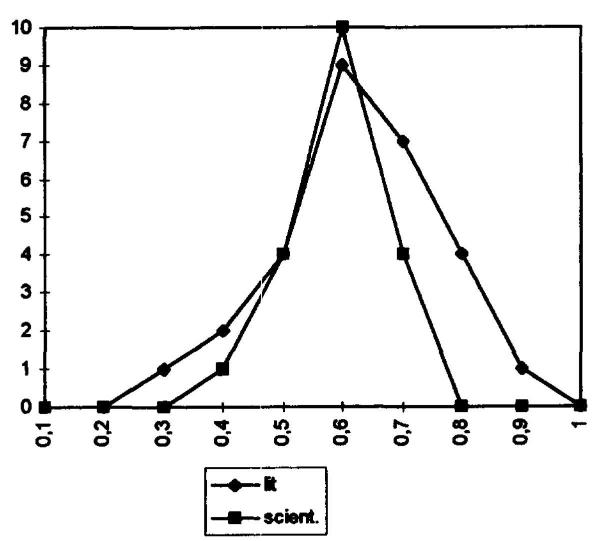

Fig. 4. Comparative response of literary and scientific population a the questionnaire.

We see in Fig. 4 that the demonstration is conclusive.

The results obtained show that scientific population is more self confident than the literary population but that it is less confident in the machines.

This observation leads us to a reflection on the population to be chosen for real experimentation. In fact, a teleoperation system includes a teleoperation operator who is a man experienced in the control of the teleoperation system, he knows the system he has in his hands. Furthermore, he is not only used to this system but he also has knowledge and culture obtained from its formation and past experience on the system.

So, the experimentation has to be done with subjects used to the co-operating work with personal computer, joystick, automatic machines, and that have also a scientific culture.

We used a population of students resuming studies in electrical engineering after having obtained a diploma of 'Graduated technician'. They are, therefore, scientific students that have at the same time the fundamental knowledge of the system and an experiment in industrial control of automatic systems.

6 RESULTS

After the operator answers the questionnaire, we have the *a priori* self-confidence, trust in others, and trust in machines in general. During the task, we follow the evolution of the trust in the joystick, of the trust in the movement of cursor, in the PC, in the software, and in the whole system.

6.1 Results concerning the initial questionnaire

The experimental protocol has been tested on the population of students described in Section 5.2. The analysis of the psychologic questionnaire at the beginning of the experimentation shows that trust in the machines and self-confidence follow a Gaussian distribution. In comparison with the curves obtained during the preliminary experiments, we see that this population is more self-confident than the scientific students tested. Between these two populations, there are, on average, 4 years difference. In fact, self-confidence is constant during the experimentation but not during the transition from adolescent age to adult age. However, the curves for trust in machine for these two kind of scientific students are quite identical.

The Gaussian distribution obtained for self-confidence makes us distinguish three kinds of people:

- the self-confident ones
- the doubtful ones
- the ones that are not self-confident.

It seems *a priori* impossible that those three classes of people give the same evaluation of trust during the experiment. We try to show this.

The questionnaire allows a refinement of this classification because if we consider more precisely the components of test 16 PF we get 5 classes.

6.2 Responses to the scale of evaluation during the experiment

We can see that there is a good correlation between the initial test concerning self-confidence and the evaluation scales: we find for the individuals composing each of the previous classes the same results concerning the scales. We can now describe these classes:

6.2.1 Class 1:
People who are not self-confident and are unable to evaluate oneself: the responses of the scales are in average < perhaps > (Fig. 5)

6.2.2 Class 2:
People hesitating between: <I do not trust myself> and <I doubt whether I trust myself> (Fig. 6). In response to some questions the operator thinks he knows how he responds but on many more questions he does not know if he answers correctly.

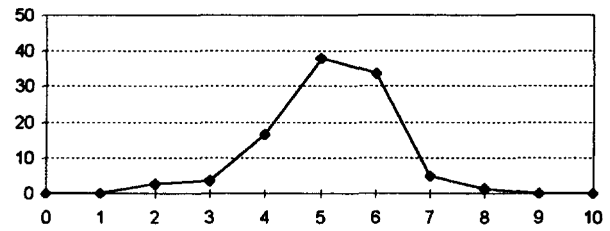

Fig. 5. Curve representative of class 1

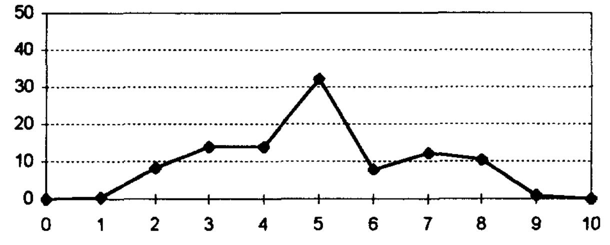

Fig. 6. Curve representative of class 2

6.2.3 Class 3:
People who don't know if they can trust themselves (Fig. 7).

The operator thinks he has answered correctly to half the questions but he does not know if he has given good answers to the other questions. He doubts his ability to respond, he does not know if he is or is not self-confident.

6.2.4 Class 4:
People hesitating between <I doubt if I am self-confident> and <I am self-confident>. The operator can say if he has answered correctly to most of the questions but on some questions he has a doubt (Fig. 8).

6.2.5 Class 5:
People that are self-confident (Fig. 9).

The curve of the response to the scales corresponds to the appreciation of someone who thinks they have been able to answer the questions.

6.3 Future developments

In fact, between the time $T0$ (time of the administration of the starting questionnaire) and the time $T1$ (time of answering the scales), trust changes, and the evolution could be represented by a vector. It is the research of these vectors which is the principal aim of the experimental process. In fact, we suppose that the trajectory of the evolution of trust in the multiaxis graph (Fig. 10) will be the function of the *a priori* trust determined previously.

Concerning the great number of parameters tested, some aggregations between these parameters are possible in order to obtain, for example, the evolution

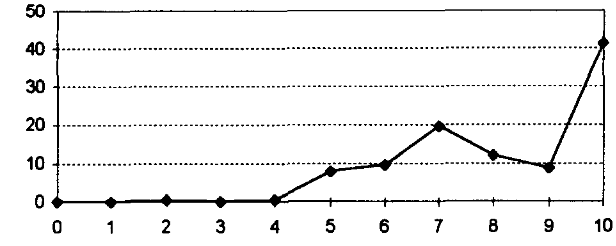

Fig. 7. Responses of individuals of class 3

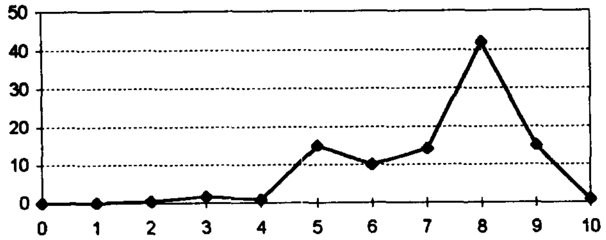

Fig. 8. Responses of individuals of class 4

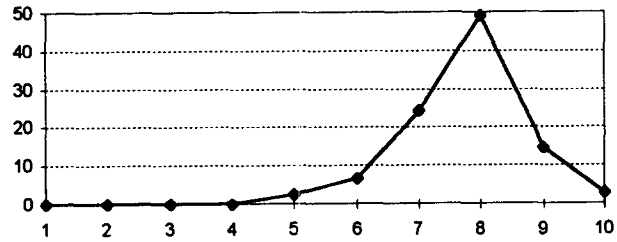

Fig. 9. Responses of individual of class 5

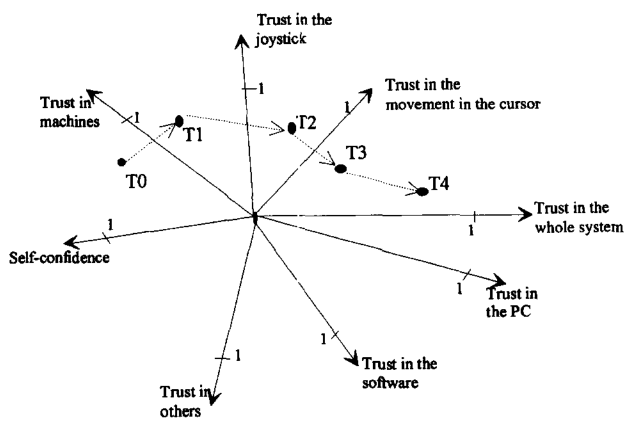

Fig. 10. Multiaxis graph for the representation of trust

of trust in the whole system as a function of the appearance of failures. Lee[9] identified the causal factors in trust using a multiple regression model, and found that trust varied as a function of the occurrence of a fault and the system performance level measured as percentage of productivity.

7 CONCLUSION

The first results show the importance of self-confidence. The dynamics of reactions confirm the place of this main element in the global compartment in front of the system. Previous studies had already put in evidence the importance of this item, and the point of our work is to show its basic aspect on individual behaviour because it allows us to class the subjects and, thus, their reaction in front of a machine and their behaviour in front of a misworking system. Unreliability changes productivity, trust and self-confidence, and those variables in turn change the mode of control. It is not sufficient to use merely productivity and the presence of a fault to predict operator performance. Subjective relationships to the machines must be included, so, for each operator, it would be included. For each operator, it would thus be possible to take into account their likely level of stress so as to take dispositions to maintain a correct human reliability.

REFERENCES

1. Sheridan, T. B., Supervisory control. In *Handbook of Human Factors*, (ed. G. Salvendy) Wiley, NY, 1987, pp.1243–1268.
2. Shafer, G., *A Mathematical Theory of Evidence*, Princeton University Press, NJ, 1976.
3. Iganaki, T. & Fujioka, T., Decision support for multiple fault, diagnosis under vague information. In *12th Eur. Ann. Conf. on Human Decision Making and Manual Control*, University of Kassel, Germany, 22–24 June 1993.
4. Heider, F., The psychology of interpersonal relations. 1958-*L63*/2389 *BLLD*, University of Kansas, 1958.
5. Rotter, J. B., Interpersonal trust, trust worthiness, and gullibility. *American Psychologist* **35** (1980) 1–7.
6. Muir, B., Operator's trust in and use of automatic controllers in a supervisory process control task. Doctoral thesis, University of Toronto, 1989.
7. Lhuilier, D. & Grosdeva, T., Stress et conduite de système complexe. *Le travail humain*, **55** (1992) 155–169.
8. Graham, J. H., *Safety, reliability and human factors in robotic systems*, Van Nostrand Reinhold, NY, 1991.
9. Visinsky, M. L., Cavallaro, J. R. & Walker, I. D., Robotic fault detection and fault tolerance: a survey. *Reliab. Engng System Safety*, **46** (1994) 139–158.
10. Lee, J. & Moray, N., Trust, control strategies and allocation of function in human–machine systems. *Ergonomics*, **35** (1992) 1243–1270.
11. Lee, J. & Moray, N., Trust, self confidence, and operator's adaptation to automation. *Int. J. Human-Computer Studies*, **40** (1994) 153–184.
12. Cattell, R. B. & Eber, H. W., Test du 16 PF de R.B. Cattell. Report edited by Les editions du Centre de Psychologie Appliquée, Paris, 1972.