CHAPTER 14

MICROWORLDS BASED ON LINEAR EQUATION SYSTEMS: A NEW APPROACH TO COMPLEX PROBLEM SOLVING AND EXPERIMENTAL RESULTS[1]

Joachim Funke

University of Bonn, Germany

ABSTRACT

The method of computer-simulated scenarios has recently been introduced to study how people solve complex problems. This paper describes a special approach to constructing such microworlds by means of linear structural equation systems. Subjects' task in the experimental situation is to first identify in a knowledge acquisition phase the causal structure of an hitherto unknown system. In a later knowledge application phase they have to control this system with respect to a given goal state. Knowledge that was acquired on the task is assessed both by means of causal diagrams – a method developed within this project and proven to be very useful – as well as by the degree of successful control performance. Three experiments on special attributes of such systems (active interventions versus observations only, effects of different degrees of Eigendynamik, the influence of different degrees of side effects) illustrate the approach. The mentioned factors have considerable influence on identification and control of the system SINUS. The conclusion deals with the advantages of an experimental approach in this area.

Following the pioneering work of Dietrich Dörner (1980, 1987, 1990, 1991) starting in the mid-seventies, several computer-simulated scenarios (some call it "microworlds"; e.g. Brehmer, 1992; Brehmer & Dörner, in press) have been developed and applied in correlational as well as in experimental studies on

---

complex problem solving (for a review see Funke, 1988, 1991). For instance, in a computer-administered microworld called LOHHAUSEN, a subject has to take the role of an omnipotent mayor of a little town (Dörner, 1980, 1987). In other work, a subject plays the role of a manager of a little shop (TAILORSHOP) or of an engineer in a Third World country (MORO; Putz-Osterloh & Lemme, 1987). In general, the new approach deals with the exploration and control of complex and dynamic systems by human individuals.

According to Dörner, subjects in such situations have to cope with the following task requirements: (1) they must deal with the *complexity* of the situation and with the *connectivity* of the variables involved since typically not only a few variables have to be handled with (LOHHAUSEN consists of about 2000 variables!); (2) they must deal with the *intransparency* or opaqueness of the situation since typically not all information that is needed is available; (3) they must deal with *dynamic developments* of variables which – over time – change their states autonomously and make it necessary to anticipate trends; (4) in contrast to simple tasks having only a single goal they must deal with *multiple goals* some of which may contradict others (e.g., as a manager: pay high wages due to the trade-union's request and at the same time maximize the company's profits).

This paper describes an approach in the area of complex problem solving developed in Bonn during the three years of the DYNAMIS project. It can be seen as an attempt to establish an experimentally and systematically oriented line of research on complex problem solving which should overcome some of the problems of early research (see the critical aspects mentioned by Eyferth, Schömann & Widowski, 1986, or by Funke, 1984). During its early phase, research was less coordinated and less rigorous with respect to traditional criteria of scientific precision. The main intention was the establishment of a new research paradigm, the method of computer-simulated scenarios. The Bonn approach represents a second generation of research activity no longer under pressure to argue for the existence of certain phenomena, but being able to introduce the first lines of ordinary research in a more settled phase of scientific development.

The paper starts in part 1 with an outline of the DYNAMIS research philosophy, including a description of the dynamic task environment and of the dependent variables which measure quality of system identification and control. Part 2 deals with three experiments being in line with the presented

research philosophy. Part 3 summarizes the results and also give some ideas for further activities.

The DYNAMIS Approach to Complex Problem Solving

Early work in the research domain of complex problem solving suffered from certain weaknesses, some of them being:

♦ In most cases the definition of a subjects' solution quality turned out to be highly arbitrary: how should one, e.g., determine the success of a town mayor? If one adds up the number of employees in the town, the energy used and the amount of money in the bank: what kind of measure would that be? What about its reliability and validity?

♦ Influence of previous knowledge on dealing with a microworld was at the same time assumed to be of high importance, but has never been controlled for. Even at the end of a simulated period it was absolutely unclear if and what subjects would have learned during the session.

♦ Each microworld was a world by its own. Only on a very global level comparisons to other microworlds were possible. Because of their idiosyncratic structure, phenomena turned out to oscillate greatly: sometimes certain effects were observed, sometimes not. Also, due to the missing replications, it was unclear how stable the results, often found with small samples, would be.

♦ From the beginning, subjects had to control the microworld without any opportunity to test certain hypotheses about assumed dependencies in the system. Due to this procedure, acquisition of knowledge could not be separated from its application. Also, if some subjects run the risk of hypothesis testing it could happen that according to their intervention and according to its nonlinear structure the microworld was brought in such a bad shape that they never escaped such an attractor despite good problem solving attempts.

The line of research done in our Bonn laboratory therefore established the following principles: (1) It should always be possible to define the quality of a solution by comparing it with an *optimal solution strategy*. (2) The situation should realize the *features of complex problems* (complexity, connectivity, intransparency, Eigendynamik [i.e., autonomous changes without intervention], and multiple goals) as far as possible. Also, different microworld situations should be comparable with respect to these criteria.

(3) A *detailed diagnostic procedure* should reveal the subject's development of hypotheses about the system. This implies that subjects have to be prompted repeatedly about the causal structure they assumed to the system. (4) There should be a clear distinction between a phase of *knowledge acquisition* (mainly realized by encouraging the subjects to explore the system) and a phase of *knowledge application* in which certain states of the problem space should be reached by the subjects as quickly as possible. In this last phase, performance measures should precisely indicate the quality of a subjects´ intervention.

The DYNAMIS Shell for Scenarios

Trivially, before you can control a complex system, you must learn how it works. To study experimentally the acquisition, as well as the application, of knowledge we confront our subjects with dynamic computer-simulated scenarios. As a universal tool for constructing these scenarios a computer program called DYNAMIS serves as a shell, with which the experimenter can implement in a simple way different types of simulated systems which all have in common one formal background. This general frame is a *linear equation system* (see e.g., Steyer, 1984) which consists of an arbitrary number of exogenous (=x) and endogenous (=y) variables according to the following equation:

$$y^{t+1} = A \cdot y^t + B \cdot x^t \tag{1}$$

where y^{t+1} and y^t are vectors representing the state of the endogenous variables at times t+1 and t; x^t is a vector representing the values chosen by the subject for the exogenous variables; A, B are matrices containing the weights for the variables.

A set of measures for formally describing such systems has been suggested (e.g., Hübner, 1989). An equation system is constructed according to theoretical considerations about the presumed influence of certain system attributes on task complexity (e.g., the effect of Eigendynamik or the influence of side effects or effects due to different interdependencies). It is *not* intended to simulate a domain of reality adequately, because that kind of simulation puts too many constraints on the attributes of the system to be useful for basic research on problem solving. Consequently, most of the simulated systems used in our research group have been "artificial". With respect to a distinction made by Hays and Singer (1989) one can say that what we want our systems to possess is not physical fidelity, but rather functional fidelity. As an example see the SINUS system shown in Fig. 1.

Subjects are told that this fictitious system consists of living creatures from a distant planet called SINUS. The "endogenous" variables are introduced as creatures labeled "Gaseln" (y_1), "Schmorken" (y_2) and "Sisen" (y_3), the "exog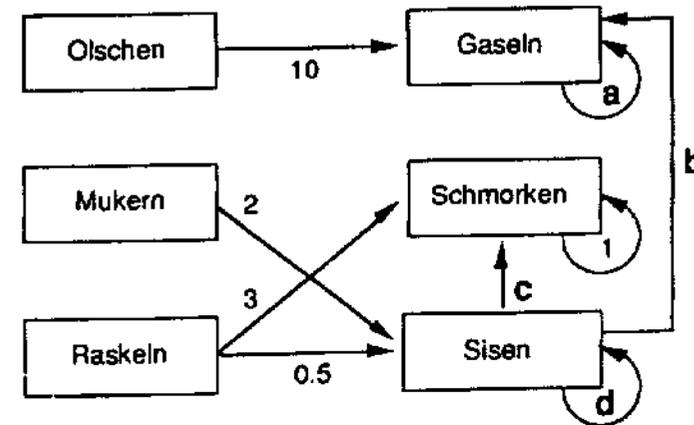enous" creatures are called "Olschen" (x_1), "Mukern" (x_2) and "Raskeln" (x_3). The system has the following structure (parameters a, b, c, and d represent variable weights, with a=1, b=0, c=0.2, and d=0.9 being the standard set):

Figure 1: Causal structure of the system SINUS. The weight parameters in the standard configuration are set to a=1, b=0, c=0.2, and d=0.9, but are changed due to the experimental purposes.

$$y_1^{t+1} = 10.0 \cdot x_1^t + a \cdot y_1^t + b \cdot y_3^t, \tag{2}$$
$$y_2^{t+1} = 3.0 \cdot x_3^t + 1.0 \cdot y_2^t + c \cdot y_3^t, \tag{3}$$
$$y_3^{t+1} = 2.0 \cdot x_2^t + 0.5 \cdot x_3^t + d \cdot y_3^t \tag{4}$$

The task for the subjects is first to explore the system (i.e., to find out the causal links between the system variables) and then to control the endogenous variables (=the numbers of y-creatures) by means of the exogenous variables with respect to a set of given goal states. Parameters a to d are manipulated depending on the experimental conditions (see below).

General Experimental Procedure

In our experiments, subjects pass through at least two phases. In the first phase, the *knowledge acquisition phase*, subjects can explore the system and its behavior as they like (learning by exploration; see also Moray, Lootsteen, & Pajak, 1986; Shrager & Klahr, 1986). They can take actions (i.e., make an intervention on one or more of the exogenous variables) and observe the resulting effects in the endogenous variables. Figure 2 shows how the SINUS microworld is presented to subjects.

SINUS	BLOCK				
Week	1	2	3	4	5
State:					
Gaseln	1600	1700	1800	1900	2000
Schmorken	900	957	1013	1055	1096
Sisen	300	293	286	281	306
Intervention:					
Olschen	10	10	10	10	?
Mukern	12	11	13	28	?
Raskeln	-1	-1	-5	-5	?

Press "space bar" to select an intervention, choose a value
and then press "return"

Figure 2:   Screen display of the numerical version of DYNAMIS when presenting system SINUS after four week (= trials) on the first block. The upper part shows the state of the three endogenous variables, the lower part shows past interventions.

Exploration is possible within four blocks following one after the other. Each block consists of a certain number of trials (referred to as "weeks"in the cover story) which all depend on each other. From one block to another the system is reset to the same starting values. From time to time we measure the knowledge that has been acquired so far by asking subjects for a graphical representation of their structural knowledge ("causal diagrams"). In the second phase, called *knowledge application phase*, the subject has to reach a defined system state and try to maintain the variable values as close as possible to the values defined as goal states. In this phase, we measure the

quality of the operator's control by assessing the distance between the current and the goal values for all endogenous variables.

Some comments on measuring structural knowledge and system performance seem necessary at this point because this is central to our studies. A review on techniques for knowledge assessment can be found in Kluwe (1988). Also, Rouse and Morris (1986) discuss some of the diagnostic problems in more detail.

Measures for quality of identification and control

Starting with *control performance quality*, the goal is to determine how well a given goal state is approximated by the operator's interventions. The classical approach requires the measurement of the deviation from the target system state in terms of the root mean squares criterion (RMS). This indicator reflects the mean deviation, independent of sign. The weights of individual deviations become increasingly higher the farther away they are from the target state. A good discussion of the frequently used RMS criterion can be found in Poulton (1973) and Bösser (1983).

Our solution for this problem is a *logarithmic transformation* of the goal deviation. This transformation leads to an evaluation of distances which is from our point of view more efficient: larger distances are no longer weighted more heavily. Rather, they are considered less important by this transformation because of an assumed decrease in measurement reliability with increasing goal distance. The transformation, thus, reduces the error variance that increases as a function of the operator's distance to the goal state.

In the experimental section the variable "QSC" refers to this kind of dependent variable ("Quality of System Control"). A low QSC score represents a good score because it results from low discrepancies between goal values and the values subjects reached on the endogenous variables through their control behavior.

Measuring the *structural knowledge* an operator has acquired about a system also requires some kind of distance or similarity measurement. In this case the distance exists between the structural relations hypothesized by subjects and those implemented in the system. For this purpose, the subject marks on a sheet (or in some versions directly on the screen) the assumed causal relationships at certain points in time (either at the end of each trial or at the end of a block of trials). This results in a subjective causal structure similar to the real one shown in Fig. 1.

Quantification of structural knowledge requires the following steps: for each causal specification of a subject one first counts whether it belongs to one of three classes of knowledge (relational, sign, or numerical; for a similar classification, see Plötzner & Spada, in press) and whether it is correct or false. Then, for each of the three levels one can determine the "quality of system identification" (QSI) in terms of the difference between "hits" (HI) and "false alarms" (FA), weighted by a "guessing" probability (p) according to the following scheme, which closely resembles the discrimination index P, from the two-high threshold model for recognition memory (see Snodgrass & Corwin, 1988; the proposed "correction for guessing" dates back to Woodworth, 1938):

$$QSI = (1-p) \times [HI/max(HI)]-p \times [FA/max(FA)], \quad -p \leq QSI \leq (1-p) \quad (5)$$

The guessing probability for numerical parameters in a dynamic system could, for instance, be set to zero. In this case all hits are counted relative to the maximal number of hits, max(HI). If one sets the guessing probability to 0.5 in the case of sign knowledge (assuming that plus and minus relations are considered as equally probable by the subject), then errors lead to a reduction in the QSI index for that level.

The index for structural knowledge, which serves as a dependent variable in the following experiments, is called "QSI" ("Quality of Identification"). A high QSI score reveals a good score because of high correspondence between implemented and assumed causal relations; it results from an additive combination of the QSI-values for all three knowledge levels. An evaluation study done by Müller (in press) demonstrates considerable reliability and, thus, sufficient psychometric quality of this index.

Experimental Studies on System Properties[2]

In the following section three experiments on the role of different system properties serve to illustrate the approach just outlined. The focus of the experiments is on the role of active intervention into a system vs. pure observation (Exp. 1), on the influence of different degrees of Eigendynamik (Exp. 2), and on the influence of side effects (Exp. 3). For each of the experiments the presentation includes a description of the independent, as well as dependent, variables, subjects, material, and procedure, hypotheses and

---

[2] This section follows partly the description of results given in Funke (1992).

result, and a short discussion. Then, in the next section, a general discussion picks up the interesting results and connects them with results from other studies.

Experiment 1: Active Intervention vs. Pure Observation

Independent and dependent variables. In this first experiment (for more details see Funke & Müller, 1988) learning by active interventions was compared to learning by pure observation of the system's development (Factor 1: intervention vs. observation, I vs. O). This factor points to the question if active regulation is really a necessary precondition for knowledge acquisition about dynamic systems. If there is reason for the assumption of different modes of learning (e.g., Berry & Broadbent, 1984, 1987), then different results for knowledge and performance have to be expected under the two treatments. In addition to the activity factor, the effect of a diagnostic tool (subjects had to predict the system's next state) was compared to a no-prediction condition (Factor 2: prediction vs. no prediction, P vs. NP). The reason for this selection was to test the hypothesis if the diagnostic questions show interference with the task or if this additional prediction request leads to a deeper understanding of the system's structure. The amount of verbalizable system knowledge subjects had acquired (QSI, as measured by the "causal diagram" at the end of exploration) and the control quality (QSC, as measured via the distance of the actual to the specified goal states) served as dependent variables.

Subjects, material, and procedure. Subjects were 32 college students from Bonn University who participated in fulfillment of course requirements. Both factors with two levels each were crossed completely yielding four different experimental groups. In each of the four conditions eight subjects were run individually. This allows for detection of "large effects" (f=0.40, according to Cohen, 1977) with α=0.10 and β=0.30 for main effects. In the I- and O-condition the method of experimental twins was used: Each subject in the O-condition observed exactly that system data which another subject (the twin) under the I-condition had produced (yoked-control design). So there was no difference with respect to the self generated or observed information about the system between the I- and O-conditions.

The microworld used was SINUS with parameters a=1, b=0, c=0.2, and d=0.9 in Eq. (2), (3), and (4). The system had to be manipulated during five blocks of seven trials each. During the first four blocks subjects could freely explore

the system. During the fifth block all subjects (both the I- and the O-group) were to reach and maintain a previously specified goal state.

Results. It was expected that the I-group should be superior to the "observers" with regard to amount of knowledge as well as to control quality. Also, the "predictors" should accumulate more knowledge than the "non-predictors". Path-analytical evaluation of the data supported these expectations only partially: The I-group was indeed better in controlling the system (significant standardized path coefficient $\beta=0.42$, $p \leq 0.10$ from I to QSC), but seemed to know less than the "observers" ($\beta=-0.30$, $p \leq 0.10$ from I to QSI). "Predictors" acquired more verbalizable knowledge than "non-predictors"(mean QSI: 1.02 vs. 0.57, $F_{(1,28)}=5.50$, $p \leq 0.10$). Knowledge about the system was generally a good predictor of control performance ($\beta=0.41$, $p \leq 0.10$ from QSI to QSC). Interestingly, there was a negative relationship between the time spent on the task and the quality of performance.

Discussion. The results demonstrate the effectiveness of both task manipulations. Active interventions allow for better system control. However, this effect is not accompanied by an increase in "externalizable" knowledge. Similar dissociations have been reported by Broadbent, FitzGerald, and Broadbent (1986), Berry and Broadbent (1984, 1987), and Putz-Osterloh (1987), for a critique see Sanderson (1989). Concerning the second factor, requiring subjects to predict the next state increases the amount of knowledge as revealed by QSI. Detailed analyses of the so-called "experimental twins" – pairs of subjects who had to cope with the same system situations either actively or passively – indicated a high interindividual variability: there were no significant correlations between the twins' QSI and QSC scores, thus showing the importance of person-specific ways of information processing.

Experiment 2: Effects of Eigendynamik

Independent and dependent variables. In this second experiment the effect of different degrees of "Eigendynamik" was analyzed. Eigendynamik means that an endogenous variable at time t has an effect on its own state at time t+1 independent of exogenous influences which might add to the effect. These autonomous system changes represent a central feature of dynamic systems compared to static ones, where changes can occur only due to active interventions of an operator. In dynamic tasks, Eigendynamik implies the existence of forces which are independent from the operator and which have to foreseen with respect to future goal values of the system. Eigendynamik requires from the operator to cope with temporal developments, either

increasing or decreasing the values of state variables, thus producing the necessity to think about the system's next states not only in terms of the planned interventions, but also in terms of the system's activity itself. Eigendynamik can easily be detected in situations where the operator does not make interventions; but such situations seldom occur because people think erroneously that they only learn about a system by actively influencing it instead, of looking at its behavior without disturbances. In case of exogenous control activities, the separation of endogenous Eigendynamik from exogenous interventions becomes much harder.

To realize different degrees of "Eigendynamik" within system SINUS, parameters a and d from Fig. 1 and Eq. (2) to (4) were changed in three steps: $a=1$, $d=1$: a control condition without any Eigendynamik (Condition 0); $a=1$, $d=0.9$: one variable with Eigendynamik (Condition 1); $a=1.1$, $d=0.9$: two variables with Eigendynamik (Condition 2). Parameters $b=0$ and $c=0.2$ were held constant. Dependent variables were QSC for control performance and QSI for verbalizable knowledge.

Subjects, material, and procedure. A total of 24 paid males doing their civil service served as subjects. Under each of the three conditions eight subjects were run individually. Assuming $\alpha=0.10$ and "large effects" ($f=0.40$), the power $1-\beta$ proves to he at 0.50 in this case for the main effect (Cohen, 1977). SINUS was used to simulate the system with the characteristics described above. The system had to be manipulated during five blocks of seven trials each. During the first four blocks subjects could freely explore the system. During the fifth block all subjects were to reach and maintain a previously specified goal state.

Results. It was expected that with an increase in Eigendynamik the amount of acquired knowledge as well as the degree of control over the system should deteriorate. Analysis of variance revealed only a significant effect for QSC ($F_{(2,21)}=3.23$, $p \leq 0.10$; mean QSC for Eigendynamik of 0, 1, and 2 are 3.86, 3.70, and 5.18), but not for QSI ($F_{(2,21)}=1.12$, n.s.). Thus, increasing Eigendynamik leads to a less good control of the system variables, but the causal dependencies are equally well detected under all three conditions (but see the restrictions of this interpretation because of medium power).

Discussion. Eigendynamik has been previously reported to have an important effect on the operators' behavior (see de Keyser, 1990). The results of the present study show that the degree of knowledge acquisition does not seem to be influenced by Eigendynamik. In contrast, the control of the system varied as a function of Eigendynamik. Particularly under the condition of two

variables with Eigendynamik control of the system turned out to be much harder. This points to the fact that knowledge acquisition and knowledge application require different abilities, which under certain circumstances lead to a dissociation of both measures.

Experiment 3: Identification of Side Effects

Independent and dependent variables. In this third experiment, the effect of three different degrees of side effects was analyzed. Side effects play a major role in the complexity of system identification and control because in most cases the side effects cannot be observed directly, but only via other indicators. At the same time, side effects have –by definition– smaller effects on the system variables than the strong main effects, thus making it more difficult to detect their existence and their subtle influences on the other system variables.

Side effects were operationalized as minor effects from one endogenous variable on to another. In this case, parameters a and d from Fig. 1 and Eq. (2) to (4) remained unchanged (1 resp. 0.9), but parameters b and c were changed in three steps: b=0, c=0: a control condition without side effects (Condition 0); b=0, c=0.2: one side effect (Condition 1); b=0.5, c=0.2: two side effects (Condition 2). Dependent variables were again QSC as a measure for control performance and QSI as indicator for acquired and verbalizable knowledge.

Subjects, material, and procedure. Under each of the three conditions eight male subjects (Bonn University students) were run individually. According to Cohen (1977), assuming $\alpha=0.10$ and "large effects" (f=0.40) power $1-\beta$ proves to be 0.47 for the main effect. The system used was again SINUS with the changes described above and with the following change of the procedure: During the first four blocks exploration was not limited by number of trial but by time (15 min per block). During the fifth block all subjects were required to reach and maintain the previously specified goal states over seven trials without time pressure.

Results. The expected influence of side effects on knowledge acquisition was confirmed by a significant negative path coefficient ($\beta=-0.35$, $p \leq 0.10$) from the side effect predictor to QSI (mean QSI for 0, 1, and 2 side effects are 1.14, 1.26, and 0.77, $F_{(2,21)}=1.74$, n.s., respectively; see Fig. 3). Also, the effect from knowledge onto control quality reached significance ($\beta=0.73$, $p \leq 0.10$ from QSI to QSC; mean QSC for 0, 1, and 2 side effects are 2.39, 2.86, and 4.72, $F_{(2,21)}=4.01$, $p \leq 0.10$, respectively; see Fig. 3). The number of trials in blocks

1 to 4 had (contrary to our expectation) no predictive value for QSC or QSI, but this conclusion is taken only as preliminary because of medium power.
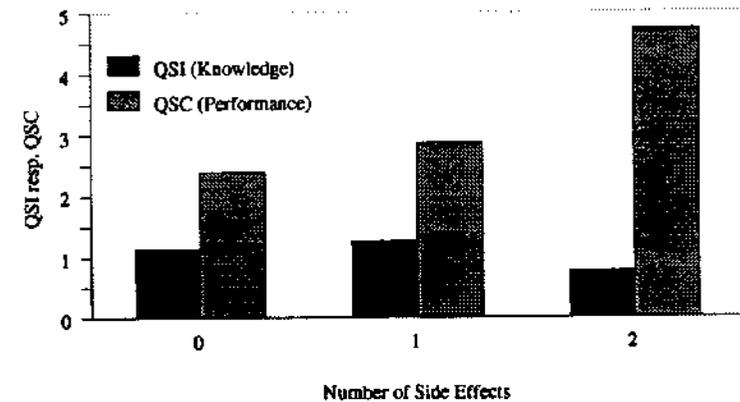

Figure 3   Number of Side Effects and its influence on the verbalizable knowledge (QSI: larger values indicating more knowledge) and control performance (QSC: larger values indicating less control).

Discussion. As in the two previous experiments, the manipulation of another system attribute shows an effect on knowledge acquisition as revealed by the QSI measure and, again, the amount of knowledge predicts the quality of system control QSC. This result is in line with Conant and Ashby (1970) according to which good control has to be the consequence of a good model.

Discussion

This section deals first with a summarizing discussion of the experimental results. The second part is concerned with a discussion of the DYNAMIS approach as a general approach to study complex problem solving.

Discussion of Experimental Results

The three experiments described above have something in common: they all demonstrate the differential effects of subtle changes in system attributes and task requirements on the dependent variables:

(1)   Active intervention leads to a better control performance, but verbalizable knowledge decreases (Exp. 1).

(2)   To let subjects predict the next system state increases their amount of verbalizable knowledge (Exp. 1).

(3)   Knowledge predicts performance (Exp. 1, 3).

(4)   Growing Eigendynamik deteriorates control performance but not the available knowledge (Exp. 2).

(5)   Growing side effects reduce control performance and available knowledge (Exp. 3).

Thus, comparing side effects with Eigendynamik as two important variables contributing to the complexity of systems it shows up that Eigendynamik can be more easily detected, but less easily controlled. In the case of side effects the situation is more complicated: subjects are not able to detect the cause of changes in the system variables correctly and, thus, have little chance of good control. As Müller (1993) points out, these problems in the identification of system structure lead to a phenomenon called "compensatory assumptions". According to this hypothesis, subjects try to build up a model which explains the available system data by means of incorrect models; the incorrectness comes from simple, but wrong causal paths which compensate for the right, but complicated ones. In his data analysis, Müller (1993) could demonstrate that subjects either have the right model about the implemented side effects (which seldom occurs) or have a certain simple wrong model which explains the system changes with compensatory assumptions.

Even if subjects would not detect different degrees of side effects or Eigendynamik – an argument used recently by Strohschneider (1991) in a critique of the presented experimental approach – the effects of this manipulation on the dependent measures cannot be denied. Also, the result of Brehmer (1989; Brehmer & Allard, 1991) point to the critical role of system characteristics. In their studies, introducing different degrees of feedback delay leads to a detrimental problem solving behavior (see also Funke, 1985; Sterman, 1989).

Discussion of DYNAMIS approach

Besides the results of the above reported experiments, there are some general features of the Bonn approach to complex problem solving worth discussing. The main progress made in the three years of the DYNAMIS project can be summarized as follows:

(1)   It was possible to develop and run an experimentally oriented research methodology in the area of complex problem solving. This guarantees causal interpretation of the reported effects.

(2)   A formal framework exists for the description and construction of arbitrary dynamic systems with continuous variables. This guarantees comparability between studies with different systems (as done in Exp. 1, 2, and 3).

(3)   Measures for quality of system control and quality of system identification have been derived which show acceptable psychometric quality. This guarantees –in combination with acceptable power of statistical tests– that the non-appearance of certain effects cannot be attributed to unreliable measures.

(4)   The effects of subtle changes in system attributes and task demands have been demonstrated experimentally with respect to their consequences for identification and control.

(5)   A first step towards a taxonomy of influence factors in dealing with dynamic systems has been made as an attempt to integrate the results (for more details see Funke, 1990).

Until now, there has not been impressive theoretical progress in this research domain. But the instruments for doing research and a corresponding framework have been developed which seem to be the basis for theoretical work. As often occurs in the development of science, the preparation of analytical tools was a necessary step for further insights (for example, even the traditional theory of finite state automata could be a useful and new tool for problem solving research, see Buchner & Funke, 1993; Funke & Buchner, 1992). This step has been done.

Concerning the general research strategy, it seems more useful to manipulate critical variables in system structures and in presentation modes than to create numerous of new systems which are completely unrelated and offer no solid basis for comparisons. Also, replications of reported effects are quite necessary – this requirement also applies to the experiments reported here.

Collecting data without theoretical assumptions produces puzzling situations in which spurious correlations may suggest significant effects where no effects are present. Only the strategy of analyzing the effects of selected variations based on some minimal theoretical premises – the experimental method – can offer new insights into the principles and mechanisms that govern complex human problem solving. For this purpose, the research strategy outlined above offers a method for the systematic construction and variation of stimulus material with well known characteristics which can used in future experiments.

References

Berry, D. C. & Broadbent, D. E (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology. 36.* 209-231.

Berry, D. C. & Broadbent, D .E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological Research, 49,* 7-15.

Bösser, T. (1983). Eine nichtlineare Regelstrategie bei der manuellen Regelung [A nonlinear control strategy at manual control]. *Zeitschrift für Experimentelle und Angewandte Psychologie, 30.* 529-565.

Brehmer, B. (1989). Dynamic decision making. In A. P. Sage (Ed.), *Concise encyclopedia of information processing in systems and organizations* (pp. 144-149). New York: Pergamon Press.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81.* 211-241.

Brehmer, B. & Allard, R. (1991). Dynamic decision making: The effects of task complexity and feedback delay. In J. Rasmussen, B. Brehmer & J. Leplat (Eds.), *Distributed decision making: Towards a basis for modeling* (pp. 319-334). Chichester: Wiley.

Brehmer, B. & Dörner, D. (in press). Research with computer simulated microworlds: Escaping the narrow straits of the laboratory as well as the deep blue sea of the field study. *Computers in Human Behavior.*

Broadbent, D. E., FitzGerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology. 77,* 33-50.

Buchner, A. & Funke, J. (1993). Finite state automata: Dynamic task environments in problem solving research. *Quarterly Journal of Experimental Psychology, 46A,* 83-118.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Conant, R. C. & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of System Science, 1,* 89-97.

de Keyser, V. (1990). Temporal decision making in complex environments. *Philosophical Transactions of the Royal Society London. B327.* 569-576.

Dörner, D. (1980). On the difficulty people have in dealing with complexity. *Simulation & Games. 11,* 87-106.

Dörner, D. (1987). On the difficulties people have in dealing with complexity. In J. Rasmussen, K. Duncan. & J. Leplat (Eds.), *New technology and human error* (pp. 97-109). Chichester: Wiley.

Dörner, D. (1990). The logic of failure. In D.E. Broadbent, J. Reason, & A. Baddeley (Eds.), *Human factors in hazardous situations. Proceedings of a Royal Society discussion meeting* (pp. 463-473). Oxford: Clarendon Press.

Dörner, D. (1991). The investigation of action regulation in uncertain and complex situations. In J. Rasmussen, B. Brehmer, & J. Leplat (Eds.), *Distributed decision making: Cognitive models for cooperative work* (pp. 349-354). New York: Wiley.

Eyferth, K., Schömann, M. & Widowski, D. (1986). Der Umgang von Psychologen mit Komplexität [Psychologist's style of coping with complexity]. *Sprache & Kognition, 5,* 11-26.

Funke, J. (1984). Diagnose der westdeutschen Problemlöseforschung in Form einiger Thesen [Diagnosis of problem solving research in the FRG in form of some theses]. *Sprache & Kognition, 3,* 159-171.

Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle [Control of dynamic systems by means of construction and application of subjective causal models]. *Zeitschrift für Psychologie, 193,* 435-457.

Funke, J. (1988). Using simulation to study complex problem solving. A review of studies in the FRG. *Simulation & Games, 19,* 277-303.

Funke, J. (1990). Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen [System attributes as influence factors in dealing with dynamic systems]. *Sprache & Kognition, 9,* 143-153.

Funke, J. (1991). Solving Complex Problems: Exploration and Control of Complex Systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and Mechanisms* (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum.

Funke, J. (1992). Dealing with dynamic systems: Research strategy, diagnostic approach and experimental results. *German Journal of Psychology, 16,* 24-43.

Funke, J. & Buchner, A. (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas [Finite automata as instruments for analyzing knowledge based problem solving processes: Introduction of a new research paradigm]. *Sprache & Kognition. 11,* 27-37.

Funke, J. & Müller, H. (1988). Eingreifen und Prognostizieren als Determinanten von System-identifikation und Systemsteuerung [Intervention and prediction as determinants of system identification and system control]. *Sprache & Kognition. 7,* 176-186.

Hays, R. T. & Singer, M. J. (1989). *Simulation fidelity in training system design. Bridging the gap between reality and training.* New York: Springer.

Hübner, R. (1989). Methoden zur Analyse und Konstruktion von Aufgaben zur kognitiven Steuerung dynamischer Systeme [Methods for analysis and construction of tasks for cognitive control of dynamic systems]. *Zeitschrift für Experimentelle und Angewandte Psychologie, 36,* 271-238.

Kluwe, R.H. (1988). Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen [Psychological methods for data acquisition about human knowledge]. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp. 359-385). München, FRG: Psychologie Verlags Union.

330 *J. Funke*

Moray, N., Lootsteen, P., & Pajak, J. (1986). Acquisition of process control skills. *IEEE Transactions on Systems, Man, & Cybernetics, 16,* 497-504.

Müller, H. (1993). *Komplexes Problemlösen: Wissen und Reliabilität* [Complex problem solving: Knowledge and reliability]. Bonn: Holos.

Müller, H. (in press). Complex problem solving: The evaluation of reliability, stability, and some causal models. In R. Steyer, H. Gräser, & K. F.Widaman (Eds.), *Consistency and specificity: Latent state-trait models in differential psychology.* New York: Springer.

Plötzner, R. & Spada, H. (in press). Analysis-based learning on multiple levels of mental domain representations. In E. De Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), *Computer-based learning environments and problem solving.* Heidelberg: Springer.

Poulton, E. C. (1973). *Tracking skill and manual control.* New York: Academic Press.

Putz-Osterloh, W. (1987). Gibt es Experten für komplexe Probleme [Are there experts for complex problems]? *Zeitschrift für Psychologie, 195,* 63-84.

Putz-Osterloh, W. & Lemme, M. (1987). Knowledge and its intelligent application to problem solving. *German Journal of Psychology, 11,* 286-303.

Rouse, W. B. & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100,* 349-363.

Sanderson, P. M. (1989). Verbalizable knowledge and skilled task performance: Association, dissociation and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 729-747.

Shrager, J. & Klahr, D. (1986). Instructionless learning about a complex device. *International Journal of Man-Machine Learning, 25,* 153-189.

Snodgrass, J. G. & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117,* 34-50.

Sterman, J. D. (1989). Misperception of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43,* 301-335.

Steyer, R. (1984). Causal linear stochastic dependencies: An introduction. In J. R. Nesselroade & A. von Eye (Eds.), *Individual development and social change: Explanatory analysis* (pp. 95-124). New York: Academic Press.

Strohschneider, S. (1991). Kein System von Systemen! Kommentar zu dem Aufsatz "Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen" von Joachim Funke [No system of systems! Comment on the paper "System characteristics as determinants of dealing with dynamic systems" by Joachim Funke]. *Sprache & Kognition, 10,* 109-113.

Woodworth, R. S. (1938). *Experimental psychology.* New York: Holt.