Does Experience or an Education in System Dynamics Help People to Solve Simple, Dynamic Problems? - A Laboratory Experiment

Stefan Grösser, Dipl.-Kfm. Techn.
M.Phil. Candidate in System Dynamics at the University of Bergen[1]
Zeppelinstrasse 8/8, 73066 Uhingen, Germany
Phone: +49.173.7887771
Fax: +49.7161.352295
stefan.groesser@web.de
www.stefan-groesser.de.vu

Abstract

The Methodology of System Dynamics claims to promote understanding of complex systems. Accepting this claim, the question arises 'Does experience or an education in System Dynamics help people to solve simple, dynamic problems?'. It guides the conduction of our experiment. The first hypothesis about no influence of additional information for problem solving has to be accepted. The performances of two different information treatment groups are not significantly different. Our second hypothesis, that people with and without experience in System Dynamics will have the same performance, has to be rejected. A significant difference between the performances of experienced people and people with no or little experience exists. A possible reason for this circumstance is that an education in System Dynamics doesn't immediately, but over a longer time horizon, enables people to comprehend dynamic systems. At last, the experimental design will be discussed and several weaknesses will be pointed out.

Keywords: *Experiment, Applicability of System Dynamics, Hypothesis, Dynamic Problem, Education, Comprehension*

1. Introduction

Decision makers must increasingly deal with complex decision situations. Available information is, in general, highly uncertain and can be misleading. Providing vividness is the essence of good decision making and leadership. In dynamic environments, leaders need tools which help them to achieve this vividness (Ritchie-Dunham *et al.* 2001). One tool, which increases the clarity and thus, manageability of complex systems, is System Dynamics. System Dynamics, originally developed at the Massachusetts Institute of Technology, is nowadays thought in several education institutions all over the world (Morecroft 1999a, Morecroft 1999b, Davidsen 1999). Even though the science of System Dynamics is in the academic field well established and accepted, it is little used in the business world, possibly because the problem analyst can't apply the System Dynamics Methodology adequately to the problem at hand. In this article, we focus on the relationship between increased capabilities in managing simple dynamic problems and experience respectively education in System Dynamics. An 'In-Between Subject' experimental design is used to test a hypothesis about the relationship be-

[1] The author gives credit to Prof. Dr. E. Moxnes, Professor at the University of Bergen/Norway, for his support.

tween the degree of System Dynamics experience/education and the performance in managing a simple dynamic problem. In addition, an information treatment is used to test a second hypothesis about the available information in a decision situation and the performance in problem solving. A discussion of the main insights from this experiment, weaknesses of the experimental design and possible further research conclude the article.

1.1 Research Question

By knowing that the science of System Dynamics claims to be capable of managing dynamic complex systems (e.g. Forrester 1958 and Sterman 2000) and the circumstance that System Dynamics is not widely used in management business (Warren 2004), it is reasonable to assumed that there is, at least, an additional parameter involved, which impede the widespread application of System Dynamics to solve dynamic business problems. Warren indicates already a number of possible reasons (Warren 2004). However, the question about the effects of experience/education in System Dynamics for problem solving has, so far, not been raised. Do people comprehend the principles of the System Dynamics Methodology and, more important, are they able to apply the knowledge to specific tasks? In this paper, we want to point out a possible gap between theoretically claimed and actual personal capabilities to understand and make use of System Dynamics. Therefore, we will analyze the research question: 'Does experience/education in System Dynamics help people to solve simple dynamic problems?'. This research experiment could be a triggering impulse in System Dynamic related management education, if there are significant differences existing because of System Dynamics.

1.2 Motivation for Further Research

A System Dynamics oriented literature review put forth that System Dynamics has been used more copious as a problem solution method then as a research tool. Serving the second purpose, some System Dynamics based experiments were conducted to test the 'Misperception of Feedback'- (Sterman 1989a, Sterman *et al.* 1996), or the 'Misperception of Dynamics'-hypothesis (Moxnes 2000). Both hypotheses state that decision makers, who do not understand feedback concepts, are unlikely to perceive the feedback loops, time delays, and nonlinearities that create system's dynamics. Another application, for which System Dynamics is used to perform scientific experiments, is the management of supply chains. Sterman et al. use System Dynamic simulations to manage supply chain behavior (Sterman *et al.* 1993 and Sterman 1989b). They conclude that managers misperceive the dynamics of supply chains facilities, which confirms the 'Misperception of Feedback' hypothesis. In addition, a series of experiments regarding the perception of simple, dynamic systems have been undertaken (Sweeney *et al.* 2000, Kainz *et al.* 2002 and Sterman *et al.* 2003). They all use an experimental design different from ours; their subjects are people who are not educated in the System Dynamics Methodology. We need subjects, which are both experienced/educated and not experienced/educated in the usage of System Dynamics.

Increases System Dynamics the capability to understand dynamic systems? This basic question has not been, until now, subject of scientific research. For this, there is at least one explanation possible: it is implicitly assumed that System Dynamics contributes to the understanding of dynamic systems and because of a supposed apparentness the question is not posed. As an overall result, the paper contributes to a validation process of the System Dynamics Methodology.

2. Design of the Experiment

In this chapter, firstly, the System Dynamics simulation, its components and the model boundary are described. Secondly, the experimental design, hypotheses and treatments are displayed. Thirdly, special aspects regarding the experimental design are shown.

2.1 System Dynamics Model

The System Dynamics model sketches a business company, in which the subject plays the role of a personnel manager. The task is to generate as much profit as possible, each month and in total. The company policy should be sustainable over a time period longer than the announced 30 month time horizon. An additional task is to find the sustainable business policy as fast as possible. The participant has to execute two decisions each month: 1) Number of personnel hirings and 2) Number of personnel layoffs. The simulation can only be played once. It is announced that the player with the highest score[2] will gain a symbolic prize.

In the following, the subject's role and the company's environment will be explained. The company sells a single product on the market. The product is unique and protected by patents. Thus, it is impossible for competitors to compete within the same market. The essential part for selling the product is market demand. The selling price for the product is fixed at 30 dollars per unit. Under normal circumstances, a sufficient profit margin will result. The only limiting factor, regarding the production capacity, is the amount of employees the company has. There are no physical restrictions for production capacity. Furthermore, there exists no delay between production and demand fulfillment. It is assumed that if customers have been supplied with the product, that they will leave the market and have no further connection to the market; neither with the monthly demand growth nor with the already existing market demand. It is assumed that one customer inquires only one unit. The market demand is therefore expressed in terms of units required.

The System Dynamics model for the experiment consists of the three model parts: Total Demand (Figure 2), Employees (Figure 3) and Bank Account (Figure 5). Each of the model parts will be described in the following.

Total Demand

The model part for the total demand simply consists of the stock 'Total Demand', which is increase by 'demand growth rate' and drained by 'fulfillment rate'. The delay time is one year. The 'Demand Growth Rate by Total Demand' is an inverse parabola function, which depends on the parameter 'Total Demand'. Figure 1 shows the dependency of the demand growth rate on the total demand. It is simply assumed that the market demand growth rate will fall to zero units per year, if the 'Total Demand' reaches a value of 6.300 units. The demand stops to increase, because of the unavailability of the product. Demand growth rate will, on the other side, reach zero units per year, if the 'Total Demand'-stock is depleted to zero units. This is because it is assumed that there exist no marketing expenditures or other effects, which could boost the demand growth rate. Therefore, only network effects can contribute to the increment of the demand growth rate. When there is no demand group existing, i. e. the demand is fully supplied, then network effects will be rather small and, thus, the same will be true for the growth rate of the demand.

[2] For more information about the score, refer to the section 'Performance' in chapter 2.1.

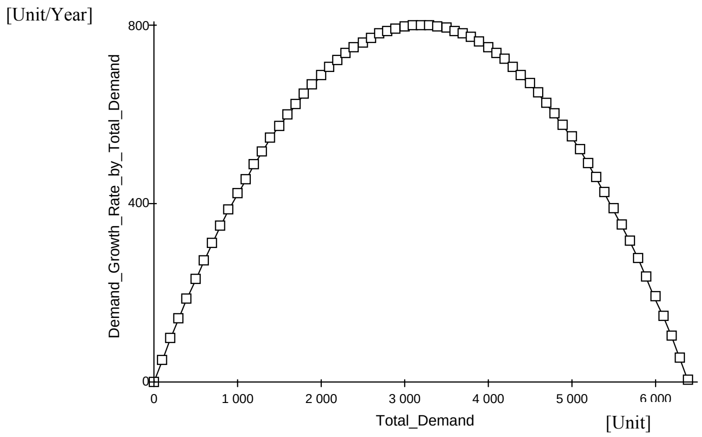

Figure 1: Demand Growth Rate depending on Total Demand

The 'Fulfillment per Employee' is a fixed value of four product units per employee per month. The amount of 'Employees' and the 'Fulfillment per Employee' determine the 'Total Capacity'. It is assumed that the employees, once they are salaried, have immediately their full productivity. Hence, it is apparent that the fulfillment rate is only depending on the number of employees. The variable 'Aux Dem' acts as a fuzzy minimum function to ensure model robustness (cf. appendix, Figure 29).

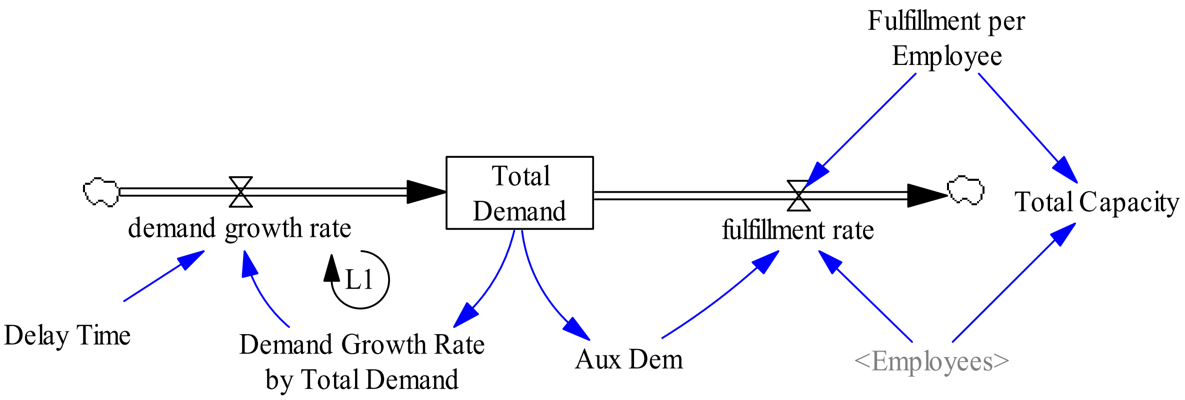

Figure 2: Stock and Flow Diagram: Total Demand

Employees

The employment sector consists of two stocks 'Available Workers' and 'Employees', which are depending on each other (Figure 3). The stocks are initialized with 999 employees for the 'Available Workers' stock and 1 employee for the 'Employees' stock. The decision variables 'Hiring' and 'Layoff' control the belonging flow rates. Through the variables 'Aux Serv' (Figure 30, appendix) and 'Aux Ava' (Figure 31, appendix), fuzzy minimum functions are introduced, which ensure model robustness. Recruitment time and layoff time depend on the workers available and the employees respectively. Figure 4 (left side) shows the graph function for the recruitment time. The recruitment time will increase, if the number of available workers at the labor market is decreasing.

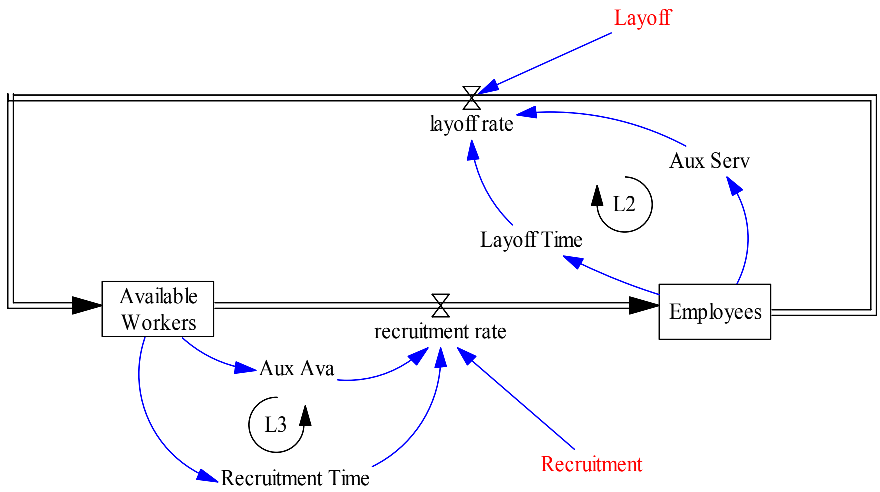

Figure 3: Stock and Flow Diagram: Employees

Layoff time depends on the amount of employees. The greater the number of employees the company has, the greater will be the layoff time. It is assumed that a greater number of employees will lead to creation of institutions, e. g. labor unions, which complicate and therefore lengthen the layoff process. Figure 4 (right side) shows the dependency graph for the layoff time.

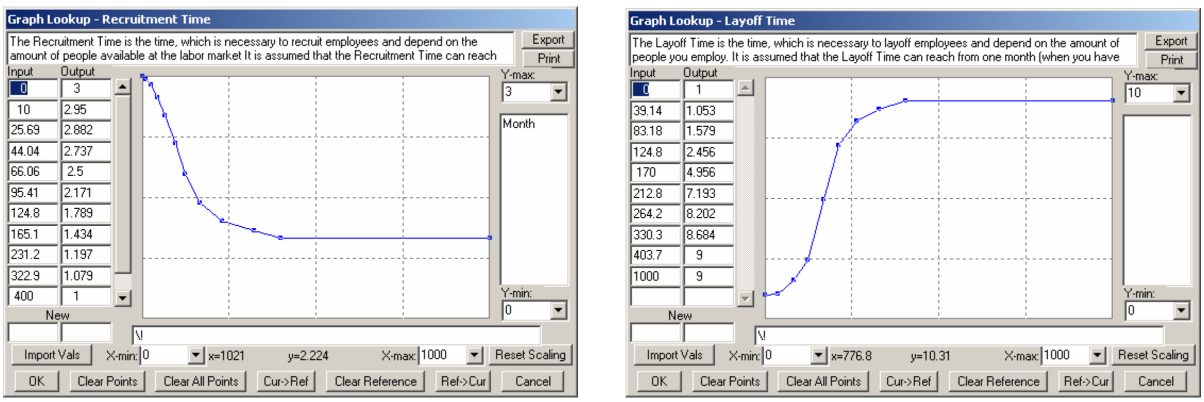

Figure 4: Recruitment Time and Layoff Time

Bank Account

The bank account sector calculates the participant's performance. Revenues in Dollar per month increase the 'Bank Account' level and total costs, also in Dollar per month, decrease it. The monthly profits are calculated as the difference of revenues and total costs. The product price is fixed to 30 Dollars a unit. The fulfillment rate determines the revenues gained. Total costs incorporate costs for layoff, for recruitment as well as for employees salary. 'Costs per Recruitment' is fixed to 50 Dollars per instance; 'Costs per Layoff' is 200 Dollars per instance and the monthly salary per employee is 100 Dollars. Figure 5 shows the stock and flow representation.

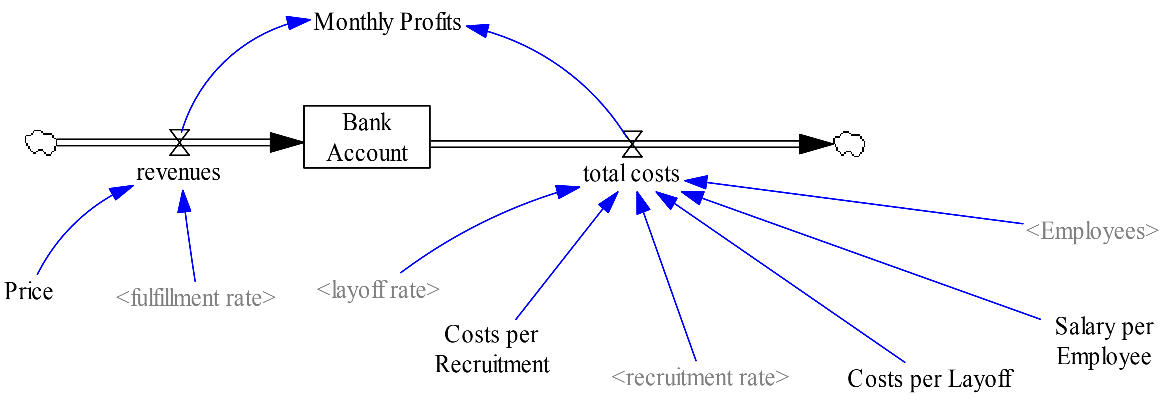

Figure 5: Stock and Flow Diagram: Bank Account

Feedback Loops
The System Dynamics model has three feedback loops L1, L2 and L3 (see Figure 6), which cause the dynamic problem in the experiment. From this three, feedback loop L1 has the strongest effect on the total model behavior. As Figure 1 shows, 'Total Demand' determines the demand growth rate. In the left part of the parabola function, the demand growth rate is increasing with an increased total demand. Therefore the causal relationship is positive. If the total demand reaches a value of 3200 units or above, the relationship will change to a negative one, i.e. the higher the total demand, the lower the demand growth rate will be. Because of these two effects, the feedback loop description in Figure 6 shows an ambiguous effect sign (+ / -) between total demand and demand growth rate by total demand. Furthermore, the higher the demand growth rate by total demand, the higher the demand growth rate will be which increases in turn the total demand. L1 is a reinforcing loop until the shift to a balancing loop occurs, when the maximum value of the parabola is passed.

The feedback loops L2 and L3 come only into account, when the player lays off employees, or the amount of available workers is considerably low. Put in other words, feedback loop L2 is only active, when layoff is done; feedback loop L3 is only active, when the participant has exceeded the optimal number of employees by more than double the optimal size. L2 captures the phenomenon that a higher amount of employees create a higher layoff time, which in turn leads to a lower layoff rate. This represents the resistance against layoffs, which is generated by labor unions. L3 depicts the balancing effect of a limited available labor force. If a high amount of workers is available on the labor market, the hiring time will be low. With a low hiring time, the hiring rate will be high. The higher the hiring rate, the lower the amount of available workers will be. The balancing loop represents the effect of restrictive resources. In both cases, the feedback loops tend to balance the system's behavior, because the participant is quite far away from the optimal trajectory. In both occurrences, the feedback loop L1 is not longer active. Figure 6 shows the three feedback loops.

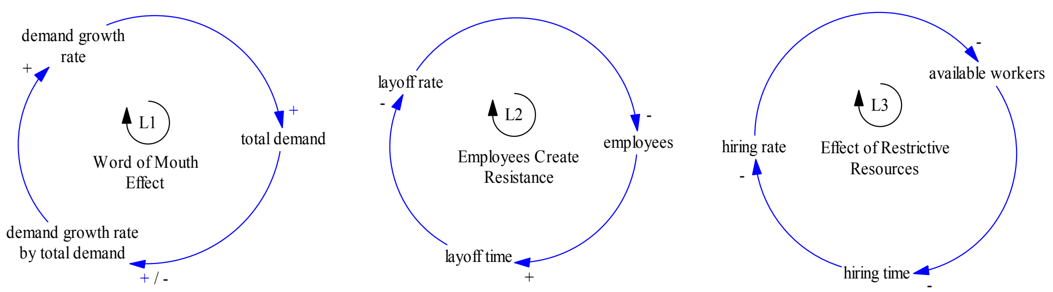

Figure 6: Feedback Loops

Understanding the word of mouth effect, which is mainly represented by the variables 'Demand Growth Rate by Total Demand' and 'Total Demand' (Loop L1), is essential for a successful personnel management policy. To comprehend this relationship was the task for the participants during the simulation. Figure 7 depicts the causal loop representation of the word of mouth effect.

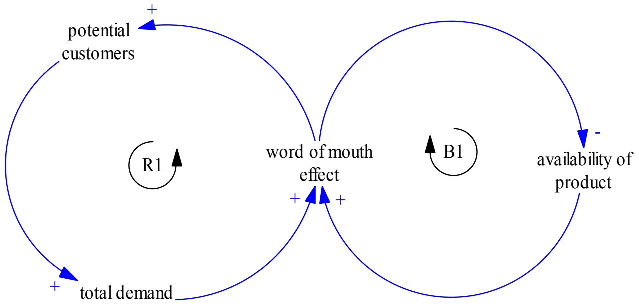

Figure 7: Detailed Causal Representation of the Demand Growth Rate

These causal loops are mainly incorporated in the variable 'demand growth rate by total demand'. By interaction between the subjects, an existing total demand will result in a word of mouth effect, which leads to increased potential customer about the product. The total demand will increase, because the potential customers are now aware of the product. The reinforcing loop generates the left half of the parabola function (Figure 1). The second half of the function is caused by the balancing loop. The higher the word of mouth effect, and given that all satisfied customers will leave the market and will not contribute to potential customers awareness, the lower will be the availability of the product. Consequently, the word of mouth effect will be reduced, because the potential customers can get the product.

Performance

Three different performance measures are used: the Bank Account at the end of the simulation, the sustainability of the chosen policy, and the quickness in reaching the sustainable policy.

All decisions taken have direct effects on the company's bank account, because either decision parameter is straightly connected with financial weights. Thus, the level 'Bank Account' accumulates all decision effects in financial terms. Furthermore, the level can only be changed by either revenues or total costs.

The sustainability of a policy can easily be judged by the deviation of the players' achieved stable level of total demand and the optimal sustainable level of total demand. The same is valid for the quickness in reaching the optimal level. Moreover, the simulation is stopped after 21 months in order to avoid windfall gain effects at the announced end of the simulation.

In consequence, the condition of the level 'Bank Account' supplemented with judgments about sustainability of the policy and the quickness in reaching the policy, offer an elaborated measure to evaluate the subject's performance.

Optimal Policy

The optimal policy for this experiment is provided to compare the results achieved by the subjects. As stated in the introduction, the personnel policy should be sustainable and it should be reached as quickly as possible. The optimal personnel policy is a single hiring action of 200 employees at month two; this is equal to the third decision (Figure 8).

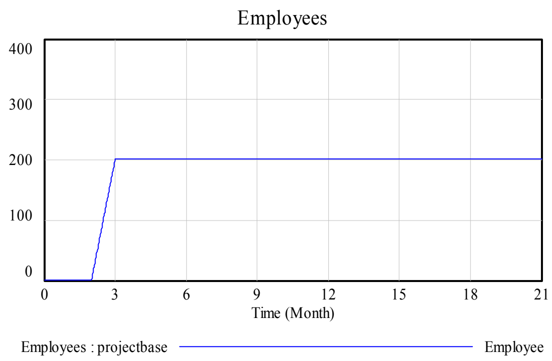

Figure 8: Hiring for the Optimal Personnel Policy

The maximal sustainable amount of fulfillment per month is reached at approximately 800 units per month. The total demand is then stable at around 3.200 requested units (Figure 9: left side). If this policy is employed, the maximal Bank Account of 64.388 Dollars will be reached (Figure 9, right side). At the same time, this policy ensures a sustainable level of total demand. Moreover, the optimal level value is reached in the quickest possible manner.

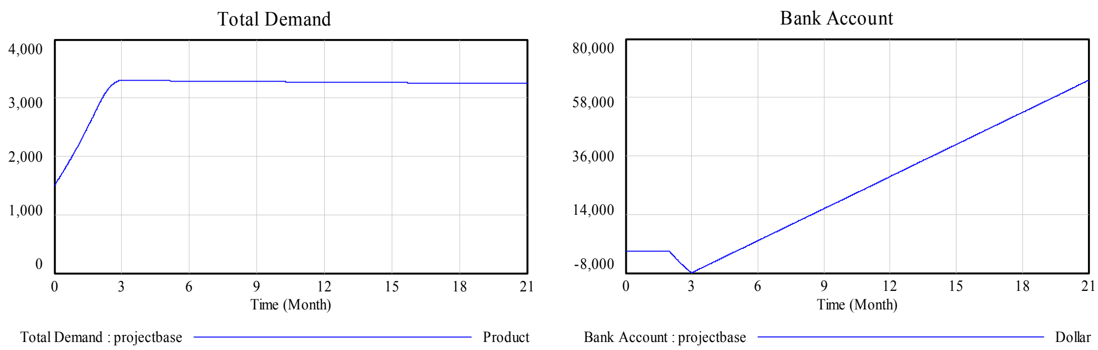

Figure 9: Total Demand and Bank Account for the Optimal Personnel Policy

2.2 Experimental Design, Treatments and Hypothesis

Experimental Design

As pointed out in the introduction, the research question is "Does experience/education in System Dynamics help people to solve simple dynamic problems?" For this experiment a 'In-Between Subject' experimental design is used. The separation dimension is 'Information Treatment'. In this experiment, the different treatments refer to different information available. The introduction of several goals, also a possible treatment parameter, was not executed.

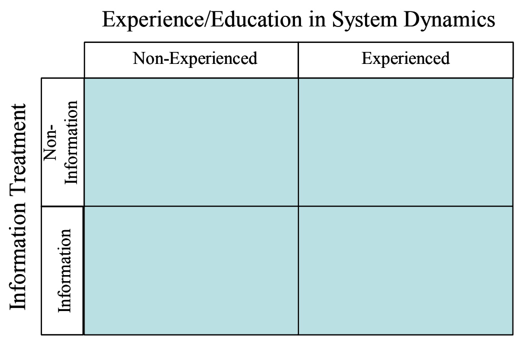

Figure 10: 'In-Between Subject' Experimental Design

Treatments

The dimension is 'Information Treatment' with two differentiations: 'Non-Information Treatment' and 'Information-Treatment'. Both participant groups received briefing information during the introduction phase. The difference in information available is that the information treated group was enabled to behold additional information screens with explicit representation of the stock and flow diagram of the System Dynamics model. In these screens, all for the simulation relevant variables were shown along with detailed information about their relationship with other variables (e.g. Figure 11).[3]

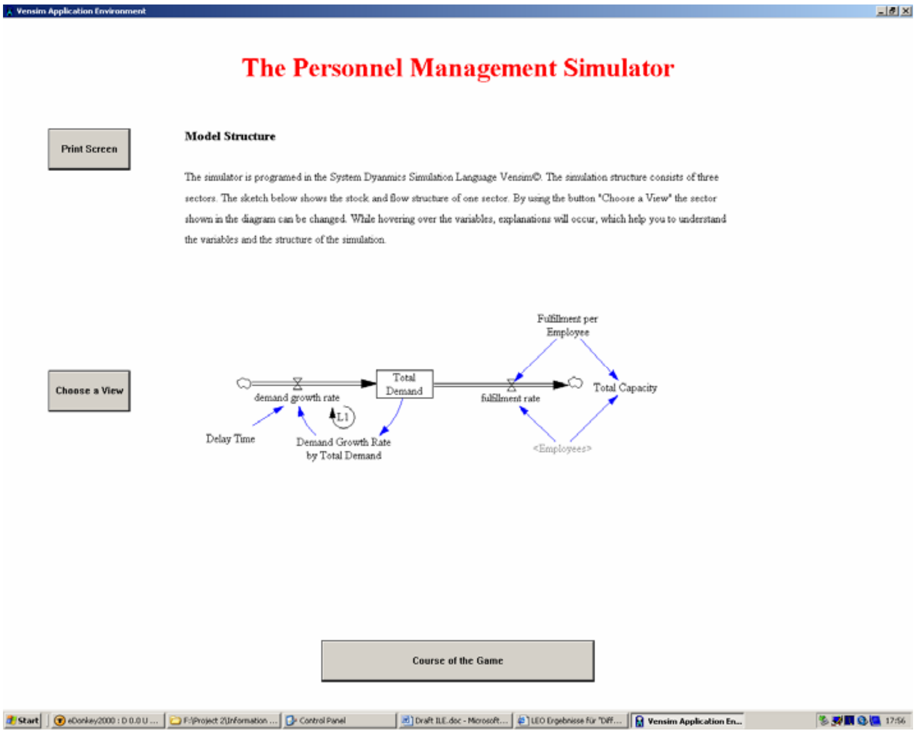

Figure 11: Occurring Screen in the Information Treatment Version

Other Research Dimension

Another research dimension is 'Degree of Experience/Education in SD' with two instances: 'No Experience/Education in SD' and 'Experience/Education in SD'. The 'No Experience'-group consist not only of people without previous contact to System Dynamics, but also of people with experience/education in SD up to 1.5 years. The total amount of participants is not equally divided into the four groups, because it is easier to reach and convince people with 'No Experience in SD' than people with 'Experience in SD' to participate in the exercise.

Hypotheses

In order to make the research question operational, a formulation of statistical testable hypotheses is necessary. Several hypotheses were formulated and tested through this experiment:

H1.0: Performance of the group of persons with information treatment is equal to the performance of the group with non-information treatment

[3] All information screens for the Information Treatment version are included in the appendix.

H1.1: Performance of the group of persons with information treatment is not equal to the performance of the group with non-information treatment

H2.0: The performance of persons, who are educated in System Dynamics is equal to the performance of persons, who are not educated in System Dynamics
H2.1: The performance of persons, who are educated in System Dynamics is not equal to the performance of persons, who are not educated in System Dynamics

2.3 Other Design Issues

The system dynamics model was designed with the Vensim© programming language. A simulation interface was created in order to make the application more accessible and more user friendly for people, who don't have experience in System Dynamics related software environments. Moreover, the Venapp© application was created to enable different information treatments. Only by usage of the Venapp© application structure was it possible to separate the System Dynamics model and the user control interface. The model core was protected; a normal user couldn't access the quantified model.

Because of the configuration of the experiment, it was possible to distribute the simulation files to the participants via email. Several advantages come along with this manner of experiment conduction: 1) participants can take as much time as the want or need for performing the simulation, 2) participants are independent from time and location, 3) no unintended influence by the experiment facilitator will occur, and 4) participants can use every support they want, except help from other experiment participants.

However, some disadvantages or restrictions occurred: 1) Vensim DSS© is necessary to run the simulation, 2) the simulation could be played more than one time, 3) motivation effects by the facilitator and other experiment participants are not present, 4) advanced Vensim© modelers could gain access to the model, use the quantified stock and flow representation and calculate the optimal solution.

In the following paragraph, the disadvantages will be discussed in more detail. The restriction that Vensim DSS© is necessary for running the simulation was successfully relaxed by providing laptops or computers with the software installed. In the chosen experiment design, it is not possible to avoid that the simulation could be played several times. However, it is literally directed in the simulation introduction to play the simulation only once. Discussing the third disadvantage, we assume that no additional motivation stimulus is needed for the participants to play the experiment, because the subjects volunteered. This shows a certain level of interest in the experiment and therefore intrinsic motivation. The third disadvantage is assumed, not being weightily. The same is supposed for the fourth disadvantage. Advanced Vensim© modelers would take on the challenge of understanding the system structure and behavior without cheating.

As a supplementary method, a post questionnaire was used to gather information about the participants like age, actual and previous field of study, risk readiness, time needed to play the simulation, the chosen strategy, etc.[4]

[4] The questionnaire is attached in the appendix.

3. Results

In the following chapter, the results of the executed experiment will be shown. Firstly, standard statistical data analyses are executed. Secondly, quantitative and inferential statistics is used to test the hypotheses.

3.1 Overview

The subjects for the experiment were recruited mainly among students with varying backgrounds at the University of Bergen. In addition, due to the experimental design, several System Dynamics students and professionals as well as persons without education in System Dynamics from Norway and Germany have participated. The experiment was conducted in March 2005. In total, 19 subjects (n = 19) participated. Due to budgetary restrictions, it was not possible to motivate more persons to participate in the experiment, and as a consequence of this, it was not possible to control the sample composition in such a way that influences originated by the participants background were certainly neutralized.

Standard Statistical Analysis

Descriptive statistical analysis assists as first method to summarize and comprehend the statistical data.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Bank Account	19	-449096	64450	-52721,74	145251,116
Experience	19	0	1	,42	,507
Time Played the Game	19	3	120	20,84	26,525
Risk Readiness	19	2	5	3,68	1,003
Valid N (listwise)	19				

Figure 12: Standard Descriptive Statistical Measures

In Figure 12, Bank Account, the most important performance measure, shows a wide, mostly negative, range of values. Noteworthy is, that the maximum value of 64.450 Dollars is greater than the optimal value of 64.388 Dollars. This indicates that at least one participant choose a strategy, which approximately fit the optimal solution. However, it shows also that the policy is not sustainable and over a longer time horizon the player would have had to change it. The histogram about the performance depicts the sample population with respect to three performance ranges (Figure 13). 31% of the participants have a result lower than zero. Only four participants out of 19 employed a quite successful strategy.

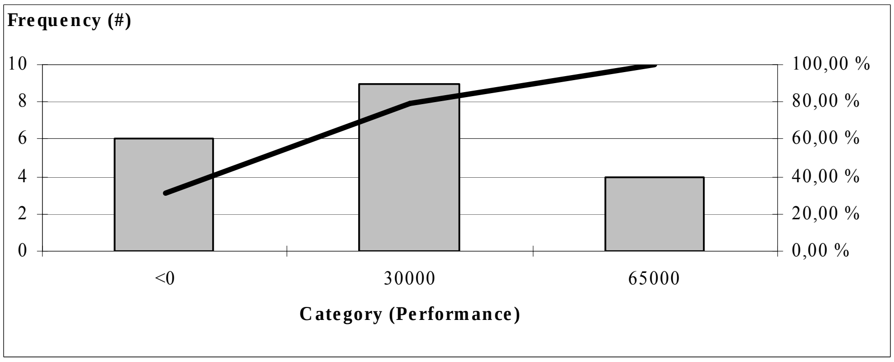

Figure 13: Histograms about the Frequency in Performance

Figure 14 present the mean value of the variable 'Bank Account' with respect to the category of experience. Three categories are used: persons with no experience in System Dynamics (No-System Dynamics; $n_1 = 4$), people with limited experience in System Dynamics (System Dynamics Novice, up to 1.5 years experience, $n_2 = 7$), and advanced users of System Dynamics (System Dynamics Professionals, more than 1.5 years of experience, $n_3 = 8$).

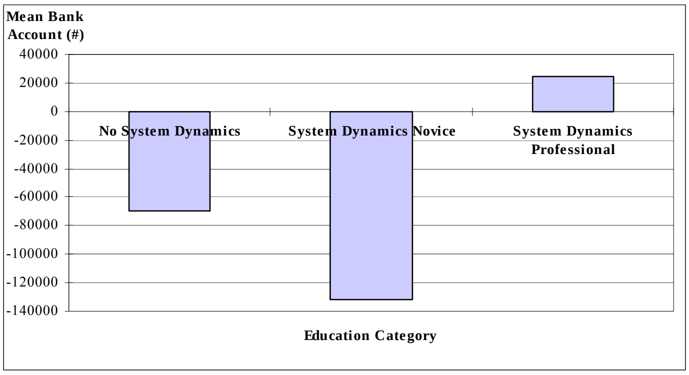

Figure 14: Mean Value of Bank Account over Experience Level in SD

The group with little experience in SD performs on average worse than the non experienced group. The group with strong background in System Dynamics has, on average and as the only group, a positive value for the variable 'Bank Account'.

Figure 15 depicts the mean value of the variable Bank Account for the three categories introduced above. In addition, we distinguish for each experience group between the information treatment. The white pillar in the diagram represents the performance of the group not provided with additional information. The gray pillar stands for the group, which has been provided with information.

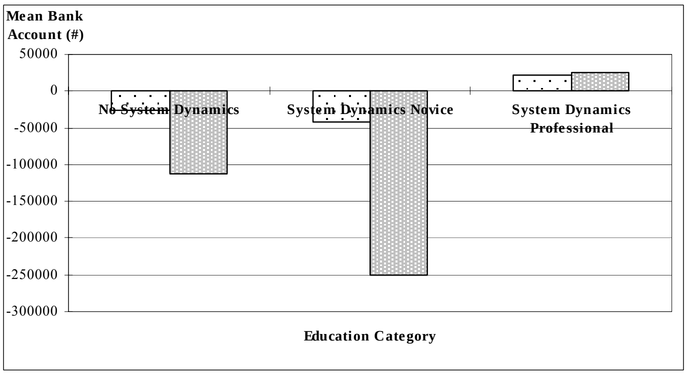

Figure 15: Mean Value of Bank Account over Experience in SD and Information Treatment

In Figure 15, the same relationship as in Figure 14 can be seen: the group with little System Dynamics experience performed worse compared to both other groups. It appears that participants supplied with the information treatment (gray pillars) were not able to use the additional information in a beneficial manner. Remarkably is that the well-experienced group performed both with and without additional information equally well.

Inferential Statistics

Total Sample

For the analyses of the data, a confidence interval of 5% is selected. Figure 16 and Figure 17 show the 5% confidence interval for the chosen 95% confidence level of the total sample (n = 19) regarding the variable 'Bank Account', 'Employees' and 'Total Demand'. The dashed line represents the optimal policy.

It is highly probable that the variable 'Bank Account' has a value below zero over the time (Figure 16, left side). This is because most of the participants hire workers too aggressive and too early in the simulation (Figure 16, left side). A high number of employees lead to a strong reduction of the market demand. Consequently, the total demand is below the sustainable value, and it grows only with a small slope. Most participants have realized this and laid off employees, that lead to additional costs. At the end, most subjects had too few employees. They did not use the market potential. Instead, the negative feedback loop of L1 is active and reduces the total demand accordingly (Figure 17).

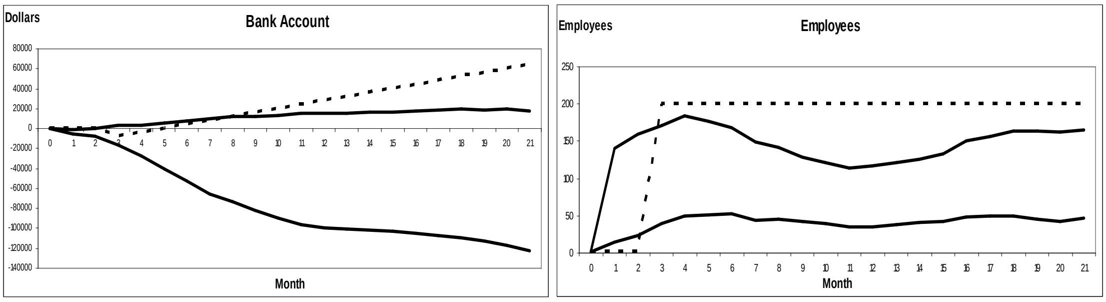

Figure 16: Confidence Interval for 'Performance' and 'Employees'

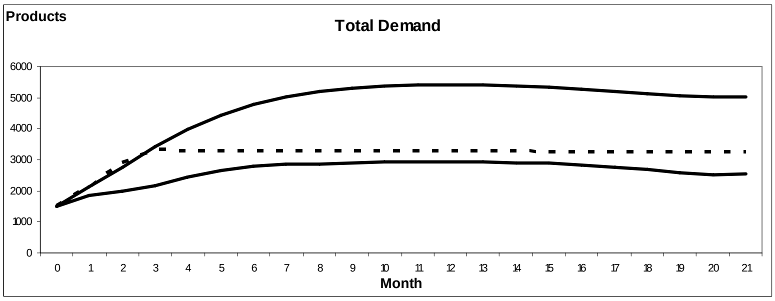

Figure 17: Confidence Interval for 'Total Demand

Information Treatment

For the information treatment, the total sample of n = 19 is divided into two groups with n_1 = 9 (Information) and n_2=10 (Non-Information). Figure 18 and Figure 19 show the 5% confidence interval for the 95% confidence level for the different groups regarding the variable 'Bank Account', 'Employees' and 'Total Demand'. The thick line represents the group, which gained additional information. Consequently, the thin line sketched the confidence interval for the 'Non-Information' group. The dashed line represents the optimal policy.

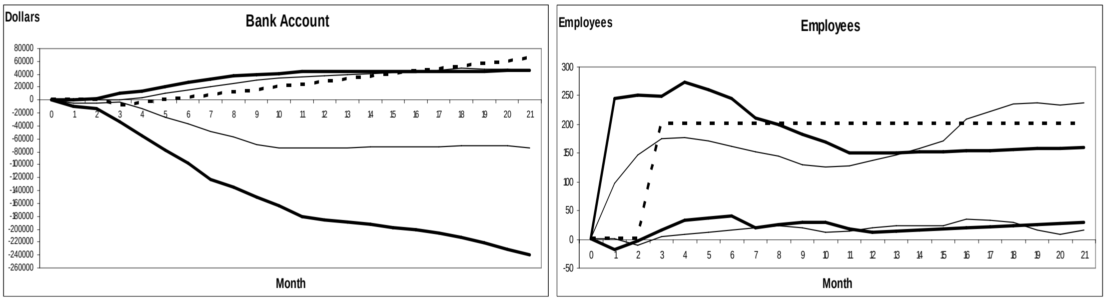

Figure 18: Confidence Interval for ‘Performance’ and ‘Employees’

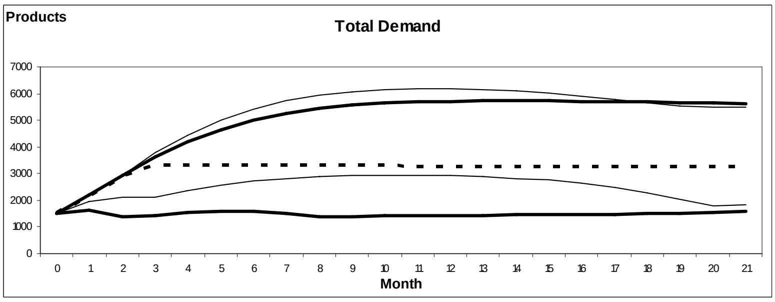

Figure 19: Confidence Interval for ‘Total Demand’

Figure 15 showed it in already: the additional information, provided by the information treatment, tempts the participants to execute more aggressive decisions. The amount of employees is, at the first twelve time steps, higher than the group with non-information treatment (Figure 18, right side). Due to great layoff expenditures, the span of possible values for the Bank Account widens, which leads to a higher probability to have high negative values in the information treatment case compared to the non information treatment (Figure 18, left side).

Experience/Education in System Dynamics

To depict the sample’s property of experience/education in System Dynamics, the total sample (n_{tot}=19) is divided into two groups with n_1=12 (No Experience in SD or Experience up to 1.5 years) and n_2=7 (more than 1.5 years experience in SD). Figure 20 and Figure 21 show the 5% confidence interval for a 95% confidence level for the different groups regarding the variable ‘Bank Account’, ‘Employees’ and ‘Total Demand’. The thick line represents the group with experience in SD. Consequently, the thin line sketched the confidence level of the ‘No Experience’ group. The dashed line represents the optimal solution.

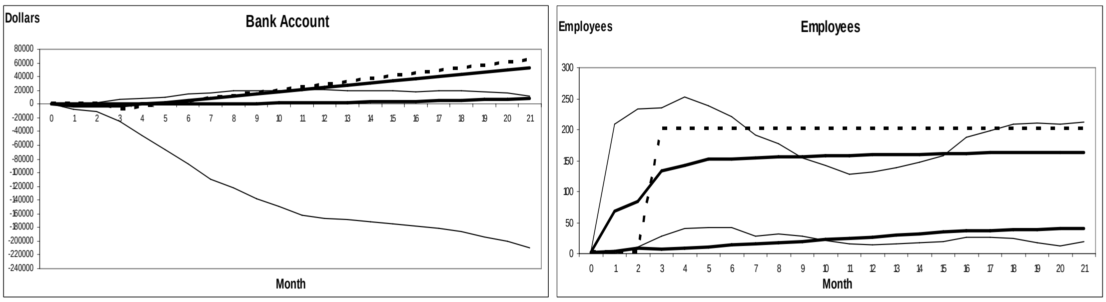

Figure 20: Confidence Intervals for ‘Performance’ and ‘Employees’

The group experienced in System Dynamics can outperform the non-experienced group significantly. The variable ‘Bank Account’ has only a small confidence interval with a width of 60.000 Dollars (Figure 20, left side). Compared to all the other confidence intervals, this is noteworthy. We can see that the experienced participants have not aggressively increased their amount of employees (Figure 20, right side). Simultaneously, even they were not conservative in their decisions; they achieved a difference which is highly significant. However, the experienced group employees too few workers, which lead to a high and self restricting total demand (Figure 21).

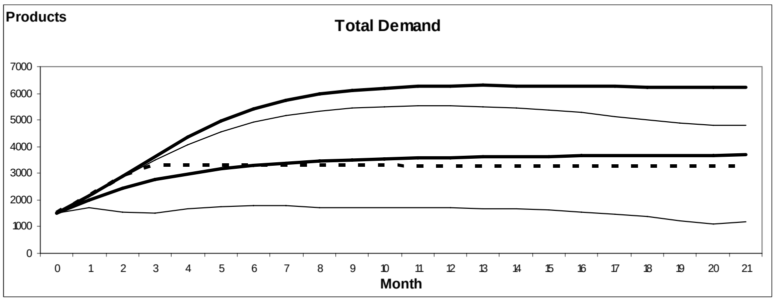

Figure 21: Confidence Interval for 'Total Demand'

3.2 Testing of Hypotheses

In the following, both hypotheses are tested. The results will be discussed in chapter 4.

Null Hypothesis 1.0 and Alternative Hypothesis 1.1
H1.0: Performance (Information Treatment) = Performance (Non-Information Treatment)
H1.1: Performance (Information Treatment) ≠ Performance (Non-Information Treatment)

Bank Account is the equivalent for performance.[5] Treatment1 is the treatment variable regarding information treatment: treatment1 = 1 represents 'Information Treated'; treatment1 = 0 stands for 'Not Information Treated'. Figure 22 summarizes the information about the sample used for testing of hypothesis H1.0. Figure 23 depict the results of the independent sample test.

Group Statistics

	Information_Treatment	N	Mean	Std. Deviation	Std. Error Mean
Bank Account	0	10	-13305,01	79479,581	25133,650
	1	9	-96518,11	190430,119	63476,706

Figure 22: Group Statistics for Sample Test #1

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Bank Account	Equal variances assumed	12,562	,002	1,268	17	,222	83213,101	65640,532	-55276,3	221702,5
	Equal variances not assumed			1,219	10,476	,250	83213,101	68271,463	-67973,2	234399,4

Figure 23: Independent Sample Test of Hypothesis H1.0

For the independent sample test of hypothesis H1.0, the level of significance is set to $\alpha = 0.05$. The standardized t-value ($\alpha = 0.05$; $d_f = 17$) is 2.11. The sample t-value is 1.268. The p-value is 0.222; i. e. the probability to incorrectly reject the hypothesis is 22.2%. Therefore, hypotheses H1.0 has to be accepted.

---

[5] See Chapter 2.1 section 'Performance' for details.

Null Hypothesis 2.0 and Alternative Hypothesis 2.1
H2.0: Performance (People Educated in SD) = Performance (People Not Educated in SD)
H2.1: Performance (People Educated in SD) ≠ Performance (People Not Educated in SD)

Bank Account is the equivalent for performance. 'Experience' is the sample property regarding education/education in System Dynamics: Experience = 0 represents 'No Experience'; Experience = 1 stand for 'Experience in SD'. Figure 24 summarizes the information about the sample used for testing of hypothesis H2.0. Figure 25 depict the results of the independent sample test.

Group Statistics

	Experience	N	Mean	Std. Deviation	Std. Error Mean
Bank Account	1	8	24446,07	25793,339	9119,322
	0	11	-108844	171119,664	51594,520

Figure 24: Group Statistics for Sample Test#2

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Bank Account	Equal variances assumed	23,605	,000	2,169	17	,045	133289,86	61466,334	3607,229	262972,5
	Equal variances not assumed			2,544	10,620	,028	133289,86	52394,241	17465,145	249114,6

Figure 25: Independent Sample Test of Hypothesis H2.0

For the independent sample test of hypothesis H2.0, the level of significance is set to $\alpha = 0.05$. The standardized t-value ($\alpha = 0.05$; $d_f = 17$) is 2.11. The sample t-value is 2.169. The p-value is 0.045; i. e. the probability to incorrectly reject the hypothesis is 4.5%, which is lower than α. Therefore, hypothesis H2.0 must be rejected; Hypothesis H2.1 has to be accepted.

4. Discussion

In the following paragraph, we discuss the results of the experiment and the experiment design. In addition, we carve out weaknesses of the experiment and infer improvement potential.

Discussion of Statistical Test Results

<u>Hypothesis H1.0</u>
The t-value of the sample is lower than the tabulated t-value. Therefore, we have to accept the hypothesis H1.0. The performance of individuals with additional information is not significantly different from the performance of individuals with only standard information. This is an interesting finding. It is generally assumed that people with experience/education in System Dynamics can used the stock and flow representation of a system to understand the behavior better. In fact, people with little experience in System Dynamics performed, on average, worse than people without experience (Figure 15). Possibly, this comes from a too fast and instantaneous judgment of the simple system used in the experiment. Most people in this group needed only up to ten minutes to read, to understand and to play the simulation. It seems that people with little experience in System Dynamics tend to judge systems according to the system's dimension, not to the system's dynamics.

Hypothesis H2.0

The t-value of the sample is higher than the tabulated t-value. Thus, the null hypothesis H2.0 has to be rejected. At the same time, the alternative hypothesis H2.1 has to be accepted. The level of experience/education in System Dynamics has a significant effect regarding the performance and, thus, the solving of simple but dynamic problems. Figure 15, Figure 20 and Figure 21 gave some indications in advance. It seems that people with experience in System Dynamics have a higher degree of problem awareness. The System Dynamics model in the experiment incorporates a medium level of uncertainty, because only one model variable is not explained in detail. Apparently, the experienced System Dynamic individuals have adapted their decision strategy to this level of uncertainty. Their decisions, even though they were not conservative, were less aggressive than the decisions of the 'not experienced' control group.

Why have the people with little experience in System Dynamics performed worse than subjects with no experience? There are several reasons possible for this circumstance: 1) the methodology of System Dynamics was not understood properly, 2) it is difficult to apply System Dynamics to the problems at hand, 3) it is difficult to mentally deduce the system behavior in simple dynamic systems, or 4) it takes time to comprehend the concept of System Dynamics. Most of the well-performing experienced subjects are Ph.D. Students. It is possible that Ph.D. students are more interested and willing in comprehending an unknown problem than first year master students.

Discussion of the Experimental Design

In order to have significant experiment results, meaning to reduce nuisance factors, it is necessary to gain control over subjects' characteristics. According to Friedman et al., the defined participant's objective of the simulation must satisfy the following characteristics (Friedman *et al.* 1994): monotonicity, salience, dominance. Monotonicity claims that the subjects must prefer more reward medium to less, and not become satiated. The salience criterion requires that the rewards received depend on the results of subject's action. At last, the dominance postulate says that changes in subject's utility during the experiment should be because of the reward medium; other influences should be negligible.

The monotonicity criterion is assumed to be satisfied, because the required behavior is proved by basic economic theory (e.g. Smith 1982). On the other hand is it possible that the saliency criterion is not fulfilled. As a payoff for the participants, a symbolic prize is given to the individual who performed best. However, this prize is fixed in value. Thus, no matter how successful the individual is, the participant can't increase the worth of the prize it will get. The dominance criterion is most properly not fulfilled. We assume that our symbolic prize is inappropriate to reduce motivational influences from other sources. In short, the symbolic prize, which was chosen due to the limitation in the author's budget, seems to be not adequate to ensure control over the subjects' behavior. However, since in the experiments only volunteering people have participated, it is to suppose that the intrinsic motivation level is quite high. Therefore, no external motivator would be needed.

In the following, we focus on the discussion of experimental problems and measurement errors. Duheim-Quine problems are common during experimental executions. Misleading instructions, influence from the facilitator, lack of anonymity of participants, lack of time and other resources, insufficient payoffs are possible instances of Duheim-Quine problems. In fact, the matter discussed earlier is a problem of this sort. However, the chosen experimental design allows preventing several problems. The instructions have been checked and discussed with several people, both with and without experience in System Dynamics. Feedback about the experiment shows that the instructions were well understood. The problem that people had to formulate their own instruction didn't occur. Thus, it is assumed that the instructions are not misleading. The same is assumed for the influence of the facilitator, because no person was engaged as facilitator. The simulation was sent to the participants via email. Moreover, the problems arising from lack of time and other resources are assumed to be rather small, because the participants had, firstly, the needed resources, mainly simulation software and equipment, available and, secondly, they could execute the simulation on their own schedule within a time frame of seven business days. An effect due to lack of anonymity could, indeed, arise, because the authors know most of the participants personally.

Furthermore is it highly possible to have sampling errors in the experiment. There are several uncontrolled nuisances present, for example subjects' attention to the task, which are not controllable. In addition, the sample was not totally randomized and it was not balanced among the experience groups.

For the experiment, a System Dynamics model is used which simplifies the reality to a certain extend. Obviously, the structure of reality can't be reproduced by such a model (Sterman 2002). We decided to include in the simulation model not all effects, which can be quantified for this topic. We did not consider, for example, the effect of marketing expenditures on demand growth rate, because this would have complicated the management task and would have lead to a lower level of comprehension of the network effect in the demand system. Even though, several effects, which exist in a real instance, are not taken into account, the experiment has external validity regarding the incorporated effects. The understandability of the instructions, which some participants assertively express, shows the achievement of the model purpose.

Further Research

For this experiment, we used a design, which allows us to reach more students by sending the simulation files to them via email. The disadvantages belonging to this design have been discussed earlier. If assumed that the individuals behave according to the instruction, the following question arises. Does a 'free experimental design' create different results than an experiment with a 'classical' design? We assume this is a subject being worth studied.

5. Conclusion

In this article, we discuss an experiment by which two hypotheses could be tested. The first hypothesis states that, even though differences regarding the degree of information about an uncertain, dynamic problem exist, the outcome is not significantly different. Given that the information is useful, it is obvious that the subjects don't take the information into account. The second hypothesis states that there is no difference in successfully solving simple but dynamic problems regarding to the degree of experience respectively education in System Dynamics. By a statistical independent sample test, it is shown that a significant better performance in dynamic problem solving depends on the level of experience or education in System Dynamics. However, further research has to be done to confirm the findings. Especially the weaknesses of this experiment design, for example not fulfilling the basic payoff criteria salience and dominance as well as the low number of participants in the experiment, must be improved. Even though the experiment has weaknesses, it gives ideas about further research potential. Especially the effect of the 'free experimental design' on subjects' performance is worthwhile being studied.

Appendix

Questionnaire

Please fill out this questionnaire after having done the 'The Personnel Management Simulator' task. I ensure you that the data will not be used for any other purposes than for scientific and debriefing purposes.

I. Personal Data

1. What is your age?

   ____

2. What is your field of study?

   ____

3. Please indicate your previous academic education.
   (e.g. B.Sc. in Engineering)

   ____

4. Please describe your academic education in mathematics.
   (e.g. two - one semester courses)

   ____

5. How would you assess your own risk readiness?
   Please use the provided scale:
   (From 1 to 7, with #1 = lowest value and #7 = highest value)
   and cross the field under the number, which you feel most comfortable with.

1	**2**	**3**	**4**	**5**	**6**	**7**

Risk Aversion Risk Seeking

6. Do you have any experience in System Dynamics?
(If yes, please answer question #7 also.)

____

7. If yes to question #6, please describe the type of education that you have in System Dynamics. If you have an education in System Dynamics, please indicate the length and the level of your education.
(e.g. 1st year master student, 2nd year master student, Ph.D. student, etc.)

____

II. The Business Management Game

1. How long did it take you to play to game?

____

2. Indicate briefly your strategy to reach the objective of the game?

____

____

____

____

3. Did you change your strategy during the game? If so, please describe your changes.

____

____

____

____

4. Do you want to make additional comments about your decisions, during the game?

____

____

____

____

____

III. Suggestion to Improve the Business Experiment

1. If you have any suggestions to improve the game, the interface, etc., then please write down your comments.

2. Are you interested in the results of the experiment?

Information Treatment Screens

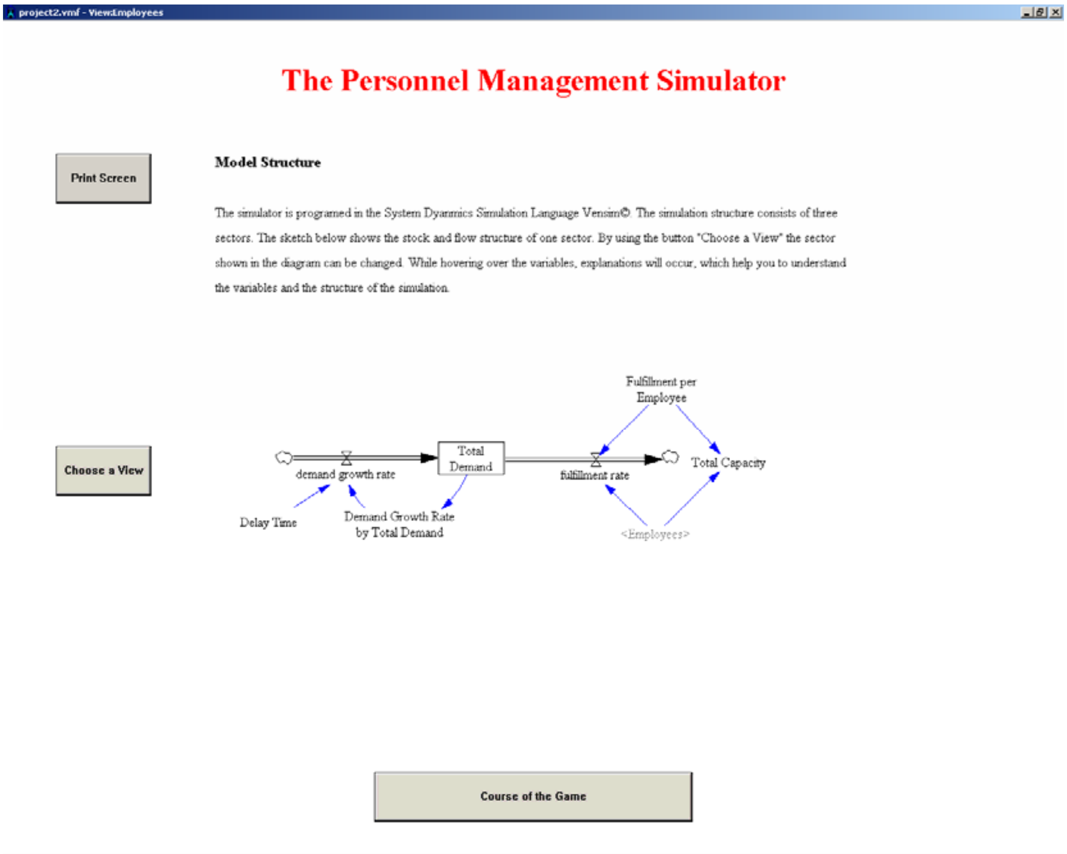

Figure 26: Information Treatment Screen (1)

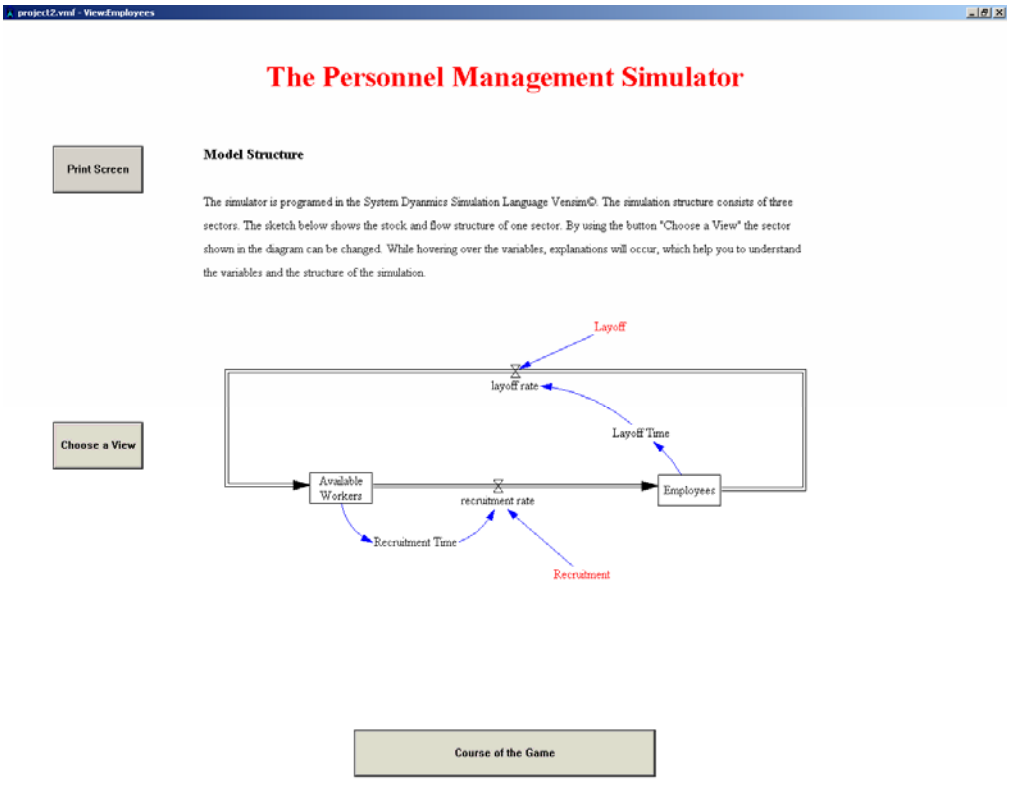

Figure 27: Information Treatment Screen (2)

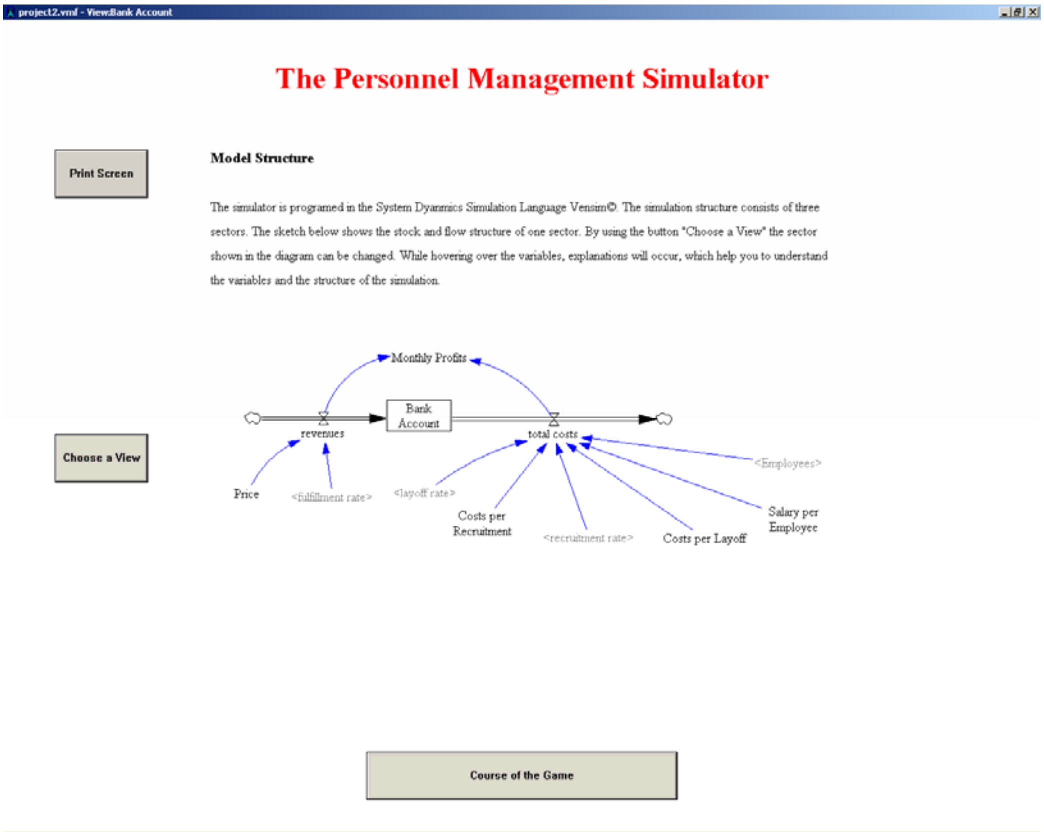

Figure 28: Information Treatment Screen (3)

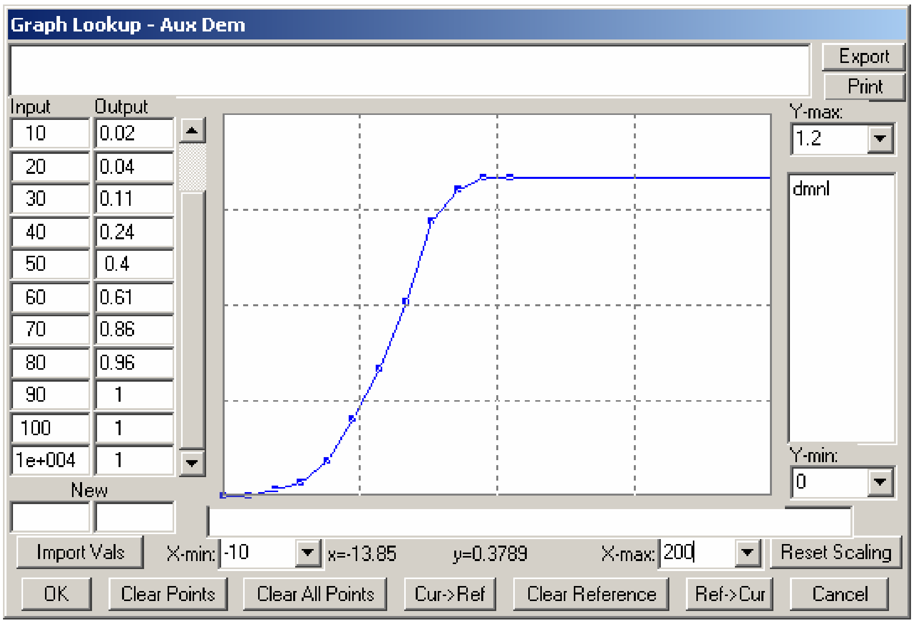

Figure 29: Variable 'Aux Dem'

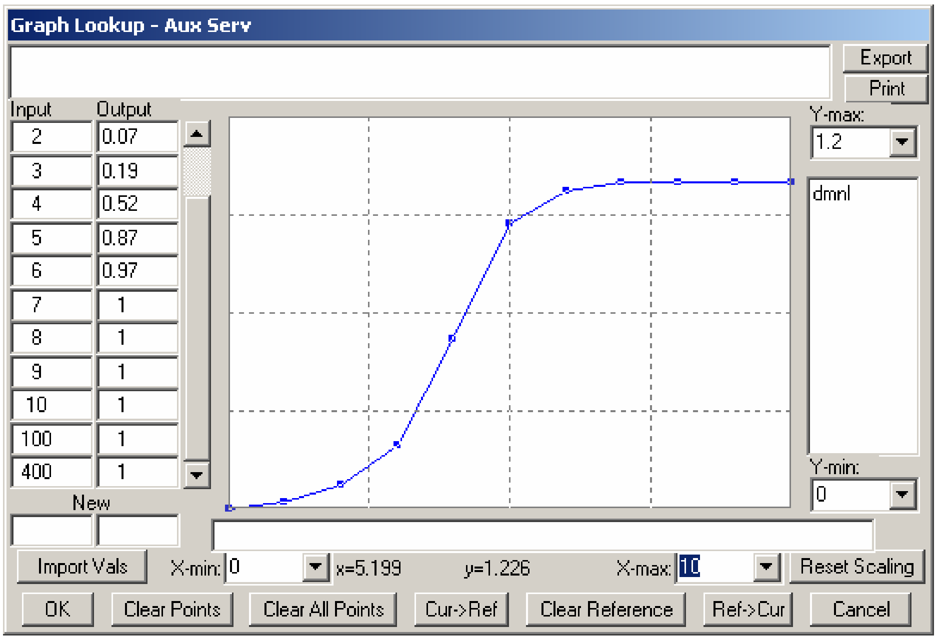

Figure 30: Variable 'Aux Serv'

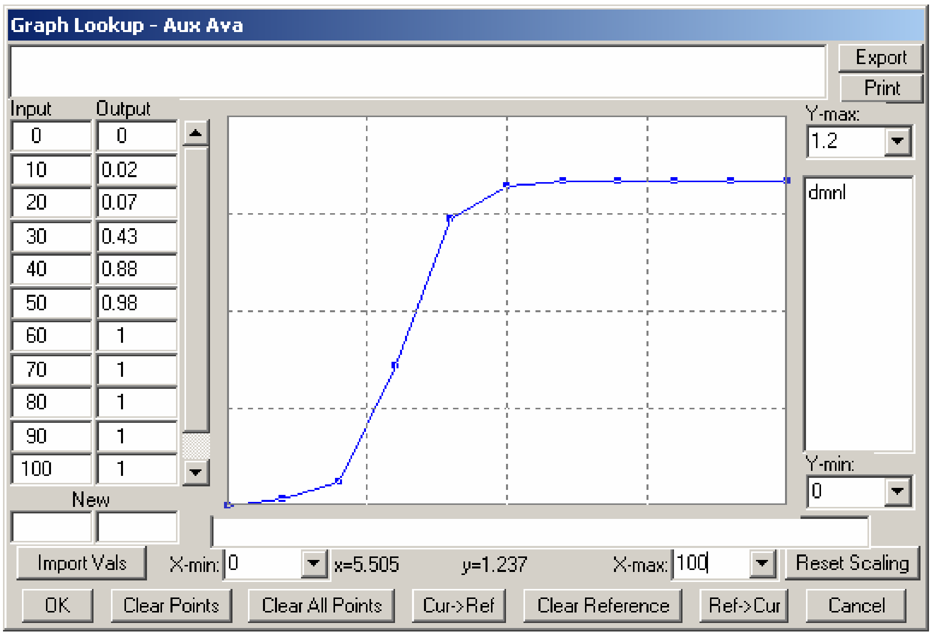

Figure 31: Variable 'Aux Ava'

Bibliography

- Davidsen PI. 1999. Graduate Programs in System Dynamics at the University of Bergen, Norway. *The Journal of the Operational Research Society* **50**(4): 311-313
- Davidsen PI. 2000. Issues in the Design and Use of System Dynamics-Based Interactive Learning Environments. *Simulation & Gaming* **31**(2): 170-177
- Follows S. 1999. Virtual Learning Environments. *T H E Journal* **27**(4): 248-252
- Forrester JW. 1958. Industrial Dynamics: A Major Breakthrough for Decision Makers. *Harvard Business Review* **26**(4): 37-66
- Forrester JW. 1969. *Urban Dynamics*: Cambridge and London
- Friedman D, Sunder S. 1994. *Experimental Methods*. Cambridge University Press: Cambridge
- Kainz D, Ossimitz G. 2002. Can Students Learn Stock-Flow Thinking? : An Empirical Investigation, *Proceedings of the 20th International Conference of the System Dynamics Society*, print/cd-rom ed. The System Dynamics Society: Palermo, Italy
- Morecroft JDW. 1999a. System Dynamics in Europe Today: A Review of Professional Infrastructure and Academic Programs. *The Journal of the Operational Research Society* **50**(4): 302-304
- Morecroft JDW. 1999b. System Dynamics in MBA Education at London Business School. *The Journal of the Operational Research Society* **50**(4): 305-308
- Moxnes E. 2000. Not Only the Tragedy of the Commons : Misperceptions of Feedback and Policies for Sustainable Development. *System Dynamics Review* **16**(4): 325-348
- Ritchie-Dunham J, Rabbino H. 2001. *Managing from Clarity: Identifying, Aligning and Leveraging Strategic Resources*. John Wiley & Sons, LTD.: Chichester, New York, Weinheim, Brisbane, Singapore, Toronto
- Smith VL. 1982. Microeconomic Systems as an Experimental Science. *American Economic Review* **72**(5): 923-955
- Sterman JD. 1989a. Misperceptions of Feedback in Dynamic Decision Making. *Organizational Behavior and Human Decision Processes* **43**(3): 301-335
- Sterman JD. 1989b. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science* **35**(3): 321-339
- Sterman JD. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill: Boston
- Sterman JD. 2002. All Models are Wrong: Reflections on Becoming a Systems Scientist. *System Dynamics Review* **18**(4): 501-531
- Sterman JD, Paich M. 1993. Boom, Bust, and Failures to Learn in Experimental Markets. MIT System Dynamics Group
- Sterman JD, Paich M. 1996. Boom, Bust, and Failures to Learn in Experimental Markets. In GP Richardson (Ed.), *Modelling for Management*, Vol. 1: 229-248. Dartmouth Publishing Company: Aldershot, UK
- Sterman JD, Stuntz LN. 2003. Learning Bathtub Dynamics: A Follow-Up. In RL Eberlein, VG Diker, RS Langer, JI Rowe (Eds.), *Proceedings of the 21st International Conference of the System Dynamics Society*, print/cd-rom ed. The System Dynamics Society: New York City, USA
- Sweeney LB, Sterman JD. 2000. Bathtub Dynamics: Initial Results of a Systems Thinking Inventory. *System Dynamics Review* **16**(4): 249-286
- Warren KD. 2004. Why has Feedback System Thinking Struggled to Influence Strategy and Policy Formulation? Suggestive Evidence, Explanations and Solutions. *System Research and Behavioral Science* **21**(4): 331-348