Structural Transparency as an Element of Business Simulators

Andreas Größler[*]

Transparent-Box Business Simulators as a Middle Course between Black-Boxes and Building Models from the Scratch

Pre-fabricated simulation tools like business simulators for individuals or computerized planning games for groups usually offer a user friendly interface. This allows inexperienced users a fast access to the simulation since they do not have to posses specific knowledge about simulation techniques. Furthermore, a learning process can be supported by various additional material like source material, external data, or instructional information creating complete, computerized interactive learning environments (ILEs). Thus, giving simulation models an easy-to-use interface increases the acceptance of the simulation tool and draws attention to it. But, in addition to that, the presence of an user interface and of additional information certainly also influences the effectiveness of a computer simulation tool to support learning.

Adding features to provide structural information about the underlying model could be a means to combine the advantages of user friendly simulators with the power of model building and analysis tools, which are supposed to give structural insight. Learners are not only able to examine the results of their decisions but the causes of these results using powerful system dynamics diagramming techniques. This introduces transparency to the former black-boxes, producing so called transparent-box business simulators (TBBS).

In the field of system dynamics the idea of providing transparency in business simulators is based on the pioneering and long-lasting work of Machuca in Sevilla (e. g., 1991, 1992, Machuca et al. 1993, Machuca and Carrillo 1996). Additional research was done by Langley in London (1995). To date also research projects at Bergen (see Davidsen and Spector 1997, Spector and Davidsen 1997) and Mannheim (Größler 1997) explore concepts and possibilities of emphasizing the relationships between structure and behavior in interactive learning environments (for a definition of terms related to computer-based simulation tools for learning purposes, see Maier and Größler 1998).

The paper reports on an experiment evaluating the relevance and effects of structural transparency. This experimental design can also be used to examine other variations of business simulators. Some hypotheses regarding the effectiveness of transparency were tested. Results show the necessity for further research and collaboration.

[*] Industrieseminar der Universität Mannheim, D - 68131 Mannheim, Germany,
Phone: (+49 621) 292-31 40 , Fax: (+49 621) 292-52 59,
E-mail: agroe@is.bwl.uni-mannheim.de

Effectiveness of Transparent-Box Business Simulators

General Hypotheses about Effectiveness of Transparency in Business Simulators

In order to measure the relevance and effectiveness of structural information (transparency condition) four hypotheses were formulated:

1a) Individuals who get structural information perform better in controlling a management simulation than those who do not get it.

1b) Individuals who get structural information in two ways (presentation about system's structure and help function containing information about system's structure) perform better in controlling a management simulation than those who only get one form of structural information.

2a) Individuals who get structural information learn more about system structure and its implications on system behavior than those who do not get it.

2b) Individuals who get structural information in two ways (presentation about system's structure and help function containing information about system's structure) learn more about system structure and its implications on system behavior than those who only get one form of structural information.

To test these hypotheses in an experiment, the independent variables must systematically be varied (see Zimbardo 1992, p. 25; for a discussion about the usefulness of experiments with complex systems see, e. g. Kluwe 1993, pp. 405–408). Thus, a 2x2 experimental design with four experimental groups is necessary: two groups of subjects who receive the presentation about the feedback relationships, two groups which do not; two groups which have access to structural help functions during playing the game, two which do not have this help.

An Experimental Design to Measure the Effectiveness of Transparency

However, in the actual experiment a slightly modified design was used. As shown in figure Figure 1, there was no third treatment group because that combination was tested in various previous experiments with insignificantly positive results (e. g. Süß et al. 1992, Putz-Osterloh 1993, Süß 1996). In fact, subjects in these studies were quite good in reproducing structural information provided, but they were not able to improve their system control competence. On the other hand, experiments conducted by Langley (1995), who used "online cognitive feedback" (i. e., a sort of help function), showed effects concerning subjects' productivity (performance / time needed). Thus, structural help was examined in two different treatment groups. In retrospect, however, it turned out to be a mistake to skip the third treatment group.

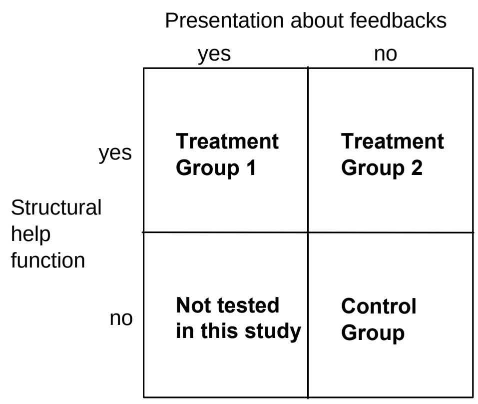

Figure 1: Groups in experiment

The experiment comprised

1. an introduction to working with *Learn!*, which was the business simulator used for the experiment (see Milling and Maier 1997, p. 336)

2. the control group and one treatment group got a lecture about ratios and ratio systems while the other group was given the presentation about the model structure,

3. three game runs with different scenarios, and

4. pre- and post-knowledge test.

This type of experimental design is known as Untreated Control Group Design with Pretest and Posttest (see Cook and Campbell 1979, p. 103):

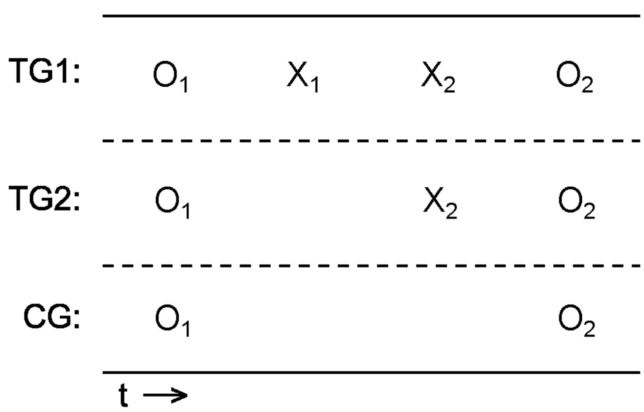

Figure 2: Design of experiment

In Figure 2 O_i stands for observation i, X_j for treatment j. Thus, O_1 symbolizes the pre-knowledge test, O_2 the post-knowledge test, X_1 the presentation about system structure, and X_2 the structural help function. The figure again shows the differences between the three experimental groups. The experiment starts at the left hand side and advances with time to the right.[1]

All groups were urged to use the possibility of getting information while playing the game (by a help function) during instruction. Structural help was available in all scenarios in the treatment groups; the control groups could access the "normal" help function of *Learn!*. The actual usage of the help function was measured by a program which recorded every call of a single help page.

At the beginning of the experiment, two goals were made clear to the subjects: to obtain a good result when controlling the simulation scenarios, and to learn as much as possible about system structure and behavior.

[1] At this point, one can also consider the double function of the used simulation tool: it serves as a treatment as well as being a means to measure the results of such a treatment. Methodological problems concerning this point are not widely discussed so far.

Figure 3 shows the time table of the experiment:

duration [min]	**control group**	**treatment group 1**	**treatment group 2**
10	biographical and business knowledge questionnaire	biographical and business knowledge questionnaire	biographical and business knowledge questionnaire
30	instructions about using simulator (including help function) and goal of simulation play	instructions about using simulator (including help function) and goal of simulation play	instructions about using simulator (including help function) and goal of simulation play
30	pre-knowledge test	pre-knowledge test	pre-knowledge test
45	useless intervention: lecture about ratios and ratio systems	intervention: lecture and exercise about feedback structure of main game areas	useless intervention: lecture about ratios and ratio systems
45	1. scenario: equal starting points	1. scenario: equal starting points (with structural help)	1. scenario: equal starting points (with structural help)
45	2. scenario: unequal starting points	2. scenario: unequal starting points (with structural help)	2. scenario: unequal starting points (with structural help)
45	3. scenario: threat of bankruptcy	3. scenario: threat of bankruptcy (with structural help)	3. scenario: threat of bankruptcy (with structural help)
30	post-knowledge test	post-knowledge test	post-knowledge test
280			

Figure 3: Timetable of experiment

The subjects in treatment group 1 got the above mentioned presentation about system structure. In addition, the subjects had to complete short exercises concerning the three main areas of importance in the pre-scenario intervention in order to turn declarative knowledge into procedural knowledge, which is supposed to be responsible for controlling success of complex systems. The groups receiving the “useless” lecture also had to fulfill short exercises. Subjects in both treatment groups played *Learn!* with structural help function.

The experiment was conducted with graduate students who could choose one of the three groups without knowing the differences between them (12 persons in the control group and in treatment group 1, 10 persons in treatment group 2, which adds up to 34 individuals). None of the subjects had any previous experience with *Learn!*. The subjects was told that the content of the lectures was crucial to playing the game in a successful way. The duration of scenario exploration had to be long enough to give subjects a chance to examine the structural help provided and to acquire a certain level of expertise. Thus, subjects had to play 15 game periods per 45 minute scenario session. The use of different scenarios should lead to more stable results because it was impossible for the subjects to improve by just repeating patterns of behavior (see Süß 1996, p. 13).

In order to measure history-related and future-related results, game results for each scenario were measured using two scores comprising development of capital resources and product quality. Subjects was told that both scores were equally important and that they had to maximize both scores. The scores were displayed during the complete game run on the screen. As a final measure the mean value of the two scores was calculated.

Learning was measured using a specific knowledge test built for *Learn!*. This test differs from widely used knowledge tests, for instance, the “Heidelberger Struktur-Lege-Technik” (Schele and Groeben 1984), in the sense that its main focus does not lie in the correct reproduction of abstract system relations but in answering about consequences of certain actions in a certain situation: subjects had to give prognoses about future system states. Thus, a qualitative instrument was used to assess the qualitative (and not easily quantifiable) goal of knowledge accumulation. As an example, subjects had to answer the question whether, in a given situation, the lowering of the product prize would increase sales revenues. An understanding of the system structure should help in fulfilling this task, but it is not enough to memorize and to reproduce, for instance, given causal loop-diagrams to accomplish this goal. Like in the simulation game, subjects have to predict results of decisions which are made. The questions comprise all decision areas of *Learn!* and are embedded in a consultancy context (i. e., subjects take over the role of a consultant showing the management of *Learn!* the results of their actions).

In addition to the advantage of not asking for memorized data, this method prevents the danger of having answers which heavily depend on situational circumstances: each question has a definite answer that can be checked for correctness. Nevertheless, the knowledge test also contains a section where subjects had to complete rudimentary causal loop diagrams. This part of the knowledge test was analyzed using Funke’s formula for quality of such diagrams (Funke 1985, p. 456).

In the very beginning of the experiment, subjects had to fill out a small questionnaire to control for potentially disturbing variables. They were asked about age, sex, field of studies, previous experiences with management simulations, and attitude towards game playing. Furthermore, a very short questionnaire was used to determine the level of business knowledge of the subjects.

Results of the Experiment

Internal consistency concerning results obtained in knowledge tests is satisfactory (Cronbach's alpha = 0.74, see Cronbach 1990, p. 202). In contrast to this, Cronbach's alpha for simulation scores is only between 0.40 and 0.58 depending whether the history-related score, the future-related score, or the mean of both is examined (see above, p. 6). Thus, concerning scores, internal consistency is rather low.

The business test showed no significant differences between the groups. Therefore, one can suppose the groups were equal concerning their pre experimental knowledge of general management issues. Also, the biographical questionnaire did not reveal any significant differences between the groups.

In the pre-knowledge test no significant differences occured, but the control group performed best. The control group also achieved the best results in the post-knowledge test; compared to treatment group 2, they did even significantly better (0.039 in an Independent Samples T-Test). So, groups differed in their performance in the knowledge tests.

Of more importance for the given issue is a change (improvement) from pre to post knowledge test. For the complete sample only a small significant improvement of performance can be observed (0.047 in a Paired Samples T-Test). If the groups are examined separately, one can find a strong significant improvement for treatment group 1 (0.001 significance in a Paired Samples T-Test), no significant changes for the other groups. These findings do not change if performance in the causal loop diagramming part of the knowledge test is calculated separately from the prognosis part of the knowledge test. If groups are compared one can find that treatment group 2 performs significantly worse than treatment group 1 in terms of knowledge test changes (0.013 in an Independent Samples T-Test). Both treatment groups taken together show no significant improvement in comparison to the control group.

The mean score does not improve from scenario 1 to scenario 3. This result was expected because the difficulty of the scenarios increased. Treatment group 1 performed best, followed by treatment group 2 and the control group. However, the only significant difference is found between treatment group 1 and the control group for the third scenario (0.009 in a Independent Samples T-Test). The two treatment groups combined do not perform significantly better in controlling the simulator than the control group.

Overall game scores and knowledge test performance show a correlation of mere 0.149. There are (at least) two explanations for this behavior: Firstly, criterium validity for the knowledge test is rather low, which means that knowledge test performance does simply not predict game performance. Then, there would be the need for a modified and improved knowledge test. Secondly, there might be only a small interdependence between those two constructs (learning and performance in game) at all. This would also lead to the conclusion, that Bakken's statement about control performance "is a reasonable indicator for dynamic understanding" is not supported by these results (Bakken 1989, p. 315). But, as was said before, this difference might be caused by the knowledge test, which does not measure what Bakken names "dynamic understanding" (he got his results from content analyses of written reports handed in from his subjects as a course asssignment).

The number of help calls shows no significant differences between the two treatment groups. Calculated over all three scenarios, the mean values for subjects calling the structural help function is 10.58 (treatment group1) respectively 12.90 (treatment group 2) times. Furthermore, there is a substantial standard deviation to be observed for help calls (12.67 respectively 14.78). There can no significant correlation be found between the number of help calls and the gaming score: only 4% of score variance is explained by different number of help calls.

Because of the very low and unbalanced use of the structural help function a modification of groups was also tested: treatment group 2 could also serve to extend the control group. This means that treatment 2 (structural help function, X_2) is removed from consideration and the experimental design shown in Figure 2. Thus, the 22 subjects from treatment group 2 and the control group now establish the new control group; treatment group 1 becomes the only treatment group.

Using these new experimental groups there were no significant differences observed concerning business test, pre-knowledge test, and post-knowledge test. However, improvement between the knowledge tests as well as game performance showed significantly better results for the treatment group (significance of 0.023 for knowledge test improvement and 0.031 for game scores in Independent Samples T-Tests). Concerning control performance these findings are in contrast to older findings reported in literature (see, e. g., Süß 1996, p. 180).

In summary, one can find that only hypotheses 2b (“Individuals who get structural information in two ways [presentation and help function] learn more about system structure and its implications on system behavior than those who only get one form of structural information”) is supported by the original experiment. However, with modified groups both remaining hypotheses (1b and 2b do not make sense with only one treatment: presentation about system structure) are supported: “Individuals who get structural information perform better in controlling a management simulation than those who do not get it” and “Individuals who get structural information learn more about system structure and its implications on system behavior than those who do not get it.”

Further Research: Experimenting with Different Forms of Structural Information

Although the above mentioned experiment yielded some interesting results it does certainly not answer all questions related to business simulators and structural information.

A methodological problem is the small number of cases. Therefore, the experiment needs to be replicated to extend the data base. This replication of the experiment will be conducted before the 1998 System Dynamics Conference. The conference presentation as well as later versions of this paper will therefore report on the results of this experiment, too. Replication could also allow other tests of the reliability of the experiment, for example, by calculating the test-retest stability (see Moser 1993, p. 93).

As in most similar experiments, the actual use of the help function was very low. Thus, effects of help functions containing structural information cannot be measured, just because people do not work with it. One can draw two conclusions from this: Firstly, the importance of the help functions has to be emphasized even more during the experiment and subjects need to have sufficient time to explore a help function. Secondly, one has to consider different forms for providing structural information, either as a kind of help function (active help?) or not.

The validity of the knowledge test is either quite low (missing correlation between test and control performance) or subjects were in fact not able to use their knowledge while controlling the complex system *Learn!*. This could be connected with the high complexity of the used simulation tool, because literature shows corresponding findings. Subjects were able to acquire knowledge which allowed them to answer the questions in the knowledge test, but this knowledge was not suitable (or “at hand”) when it was needed to control the simulator. However, the correlation of game performance with the knowledge test described here is much better than another one which was used in connection with *Learn!* before (Wittmann et al. 1996, p. 18). But it is still far from being satisfactory.

The embedding of the simulation in its environment certainly has influences on learning effectiveness. Thus, some points must be taken into account to move from a management simulation to a learning environment (constructive feedback from tutor, guides to improve performance, links to things already known, etc., see Gagné 1985). In particular, additional work is needed to clarify the role of the presentation (and the presenter, “human-human interactions”, see Davidsen and Spector 1997, p. 759) which was given to the subjects in the beginning of the experiment. Data measured in this experiment indicate that such a presentation is more effective than it was said in literature: Even a non-reproductive knowledge test was solved better by subjects who had obtained the presentation and game performance was improved as well.

Finally, the role of time to explore a system and therefore to learn about it cannot be neglected (see Größler 1997). The results of this experiment support the notion that the more complex a system is the more time is needed to understand it and to learn from it. Furthermore, the search for other means to produce significant learning effects than an enhanced help function has to be intensified.

References

Bakken, Bent. 1989. Learning in Dynamic Simulation Games; Using Performance as a Measure, in: Peter M. Milling and Erich O. K. Zahn (eds.): *Computer-Based Management of Complex Systems: Collected Papers from the 1989 International System Dynamics Conference*, Berlin: 309–316.

Cook, Thomas D. and Donald T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Chicago.

Cronbach, Lee J. 1990. *Essentials of Psychological Testing*, 5th edition, New York.

Davidsen, Pål I. and J. Michael Spector. 1997. Cognitive Complexity in System Dynamics Based Learning Environments, in: Yaman Barlas, Vedat G. Diker and Seçkin Polat (eds.): 15th International System Dynamics Conference: Systems Approach to Learning and Education into the 21st Century, Istanbul, pp. 757–760.

Funke, Joachim. 1985. Steuerung dynamischer Systeme durch den Aufbau und Anwendung subjektiver Kausalmodelle [Controlling of Dynamic Systems by Building and Using Subjective Causal Models], in: *Zeitschrift für Psychologie*, Vol. 193, no. 4: 443–465.

Gagné, Robert M. 1985. *The Conditions of Learning*, 4th edition, New York.

Größler, Andreas. 1997. Giving the Black-Box a Lid—Providing Transparency in Managment Simulations, in: Yaman Barlas, Vedat G. Diker and Seçkin Polat (eds.): *15th International System Dynamics Conference: Systems Approach to Learning and Education into the 21st Century*, Istanbul: 375–378.

Größler, Andreas. 1997. As Time Goes By—Self-Proceeding Management Simulations, in: Yaman Barlas, Vedat G. Diker and Seçkin Polat (eds.): 15th International System Dynamics Conference: Systems Approach to Learning and Education into the 21st Century, Istanbul, pp. 371–374.

Kluwe, Rainer H. 1993. Knowledge and Performance in Complex Problem Solving, in: Gerhard Strube and K. F. Wender (eds.): *The Cognitive Psychology of Knowledge*, Amsterdam: 401–423.

Langley, Paul A. 1995. Building Cognitive Feedback into a Microworld Learning Environment: Results from a Pilot Experiment, in: Toshiro Shimado and Khalid Saeed (eds.): *System Dynamics 95*, Cambridge/Tokyo: 628–637.

Machuca, José A. D. 1991. When Creating and Using Games, Are We Neglecting the Essential of System Dynamics?, in: Khalid Saeed, David Andersen and José Machuca (eds.): *System Dynamics `91*, Bangkok: 329–335.

Machuca, José A. D. 1992. Are We Losing One of the Best Features of System Dynamics?, in: *System Dynamics Review* Vol. 8, no. 2 (Summer 1992): 175–177.

Machuca, José A. D. et al. 1993. Systems Thinking Learning for Management Education. What are our Ideas and How are we Going About it in Sevilla?, in: José A. D. Machuca and Enrique Zepeda (eds.): *Proceedings of the 1993 International System Dynamics Conference: The Role of Strategic Modeling in International Competitiveness*, Cancún, Mexiko: 298–307.

Machuca, José A. D. and Miguel A. Domingo Carrillo. 1996. Transparent-Box Business Simulators versus Black-Box Business Simulators: an Initial Empirical Comparative Study, in: George P. Richardson and John D. Sterman (eds.): *Proceedings of the 1996 International System Dynamics Conference*, Cambridge: 329–332.

Maier, Frank H. and Andreas Größler. Forthcoming. A Taxonomy for Computer Simulations to Support Learning about Socio-Economic Systems, to be published in the *Proceedings of the 1998 International System Dynamics Conference*, Québec City.

Milling, Peter und Frank Maier. 1997. On the Effectiveness of Corporate Gaming Environments and Management Simulators, in: Yaman Barlas, Vedat G. Diker and Seçkin Polat (eds.): *15th International System Dynamics Conference: Systems Approach to Learning and Education into the 21st Century*, Istanbul: 335–339.

Moser, Klaus. 1993. Planung und Durchführung organisationspsychologischer Untersuchungen [Planning and Implementation of Examinations in Organizational Psychology], in: Heinz Schuler (ed.): *Lehrbuch Organisationspsychologie*, Bern et al.: 71–105.

Putz-Osterloh, Wiebke. 1993. Unterschiede im Erwerb und in der Reichweite des Wissens bei der Steuerung eines dynamischen Systems [Differences in Acquisition and Range of Knowledge in Controlling a Dynamic System], in: *Zeitschrift für experimentelle und angewandte Psychologie* Band XL, Heft 3: 386–410.

Scheele, Brigitte and Norbert Groeben. 1984. *Die Heidelberger Struktur-Lege-Technik (SLT)*, Weinheim.

Spector, J. Michael and Pål I. Davidsen. 1997. *Constructing Effective Interactive Learning Environments using System Dynamics Methods and Tools*: Interim Report, EIST Publication and Reports, No. 1, University of Bergen.

Süß, Heinz-Martin. 1996. *Intelligenz, Wissen und Problemlösen* [Intelligenz, Knowledge, and Problem Solving], Göttingen.

Süß, Heinz-Martin et al. 1992. *Wissen und Problemlösen* [Knowledge and Problem Solving], Vortrag auf dem 38. Kongreß der DGPs in Trier.

Wittmann et al. 1996. *Determinanten komplexen Problemlösens* [Determinants of Complex Problem Solving], Zwischenbericht und Verlängerungsantrag zum Projekt "Problemlösungskapazität als Funktion von Intelligenz, Wissen und Persönlichkeit" im Forschungsschwerpunktprogramm des Landes Baden-Württemberg, unpublished.

Zimbardo, Philip G. 1992. *Psychologie*, 5th edition, Berlin et al.