Integrating Stock-and-Flow Structures in Decision-support Dashboards – A Catalyst to Improve Dynamic Decision Making

Jürgen Strohhecker & Michael Leyer

Frankfurt School of Finance and Management, Frankfurt am Main, Germany

ABSTRACT

According to natural decision models, good decisions are mainly dependent on understanding the consequences of chosen options. Thus, receiving information on causal relations between options and results is supposed to be helpful. Using a capacity management simulator, we conducted laboratory experiments with two levels of complexity in which participants had to make decisions repeatedly. Results indicate that showing not only key performance indicators reflecting the consequences of chosen options on the user interface but also visualizing causal relations between them leads to better decisions. The results are stronger in the more complex situation.

Key words: Dynamic Decision Making, Mental Model, Causal Relations, Dashboard, Stock-and-flow Structure

THE CHALLENGE OF DYNAMIC DECISION MAKING

Dynamic decision making (DDM) – defined as a sequence of multiple, interdependent, and sometimes real-time decisions occurring in complex dynamic environments (Brehmer, 1992; Edwards, 1962) – is an ubiquitous challenge in society, business, and personal affairs. However, human performance in such tasks is generally poor. In more than 30 years, dynamic decision making research has accumulated ample evidence of decision making failures in dynamic complex systems. Poor results persist over a wide range of systems such as a town, a warehouse, a supply chain, fishery, and a coal-fired power plant (Dörner, 1980; Dörner *et al.*, 1994; Moxnes, 1998; Reichert & Dörner, 1988; Sterman, 1989b; Wittmann & Hattrup, 2004). Typically, these systems consist of stocks and flows as well as interrelating information links (Forrester, 1961). They are characterized by feedback and delays between cause and effect (Sterman, 1994). Human decision makers perceive these systems as opaque, incomprehensible, and hard to control (Dörner, 1996).

While a universal theory of dynamic decision making has not yet emerged (Fox *et al.*, 2013), various research efforts in the past three decades have contributed to a better understanding of the "logic of failure," as Dörner (1996) named these phenomena. Grounded in psychology, one stream of research explored the relationship between DDM performance and a broad range of personal traits, specifically intelligence, knowledge and personality. Although early studies did not find performance in micro world experiments correlating with tests of intelligence or personality (Brehmer, 1992; Dörner *et al.*, 1994), more recent research using advanced intelligence structure tests observed a significant, medium strong relationship between intelligence and performance in complex simulations (Beckmann & Guthke, 1995; Strohhecker & Größler, 2013; Süß, 1996; Wittmann & Hattrup, 2004).

Another common explanatory pattern for poor human decision making in dynamically complex situations that emerged from mostly system dynamics based research is misunderstanding of feedback and delays. Sterman (1989a, 1989b) showed that supply chain management performance suffers systematically from misperceptions of the feedback structure of the system that has to be managed. Participants use inappropriate anchoring heuristics, misperceive time lags, and use open-loop explanations of dynamics. These phenomena have been repeatedly re-observed and corroborated (Barlas & Özevin, 2004; Brehmer, 1992; Cronin & Gonzalez, 2007; Diehl & Sterman, 1995; Kleinmuntz, 1985; Moxnes, 1998; Paich & Sterman, 1993; Rahmandad *et al.*, 2009). In more general terms, this stream of research has identified shortcomings in the decision makers' knowledge or mental models about the causal structure of reality which is seen as an important element in dynamic decision making theorizing (Dörner, 1996; Dörner & Wearing, 1995; Sterman, 1994). Dörner and Wearing (1995) argue that better performing participants in a dynamic decision making experiment use more elaborated networks of causal relations in their decision making process than the bad performers. Using longitudinal case study data (Barr *et al.*, 1992, p. 15) found that "successful organizational renewal is a continuous process of first and second order changes in cognitive maps". Recently, Gary and Wood (2011, p. 569) provided empirical evidence based on laboratory experiments that “accurate mental models lead to better decision rules and higher performance”.

This paper builds on the findings of previous research, specifically on the existence of a positive relationship between causal knowledge and decision making performance. These findings are extended by focusing on the research question if and how the decision makers mental model (and as a result thereof her/his dynamic decision making performance) can be advanced by showing both causal relations between levers and objectives and the stock-and-flow structure in a decision making cockpit.

Methodologies such as system dynamics (Forrester, 1961; Sterman, 2000) or soft systems methodology (Checkland, 1981; Checkland & Scholes, 1990) have demonstrated their effectiveness in improving decision makers’ mental model and closing knowledge gaps regarding causal relations (e.g., Akkermans, 1993; Pala & Vennix, 2005; Sterman, 2010). However, learning and applying these approaches as a whole is time consuming and costly – although advances in interactive modeling and simulation tools has lowered the bar (Morecroft, 2007; Sterman, 2000). Therefore, we are focusing on one element from the system dynamics methodology that is relatively easy to integrate in decision supporting management dashboards – stock-and-flow diagrams. This type of diagrams visualizes both the causal relations between variables and their stock-and-flow structure. Management dashboards (or cockpits) that assemble and visualize the key indicators needed to steer an organization towards its objectives (Daum, 2006; Eckerson, 2011; Rasmussen *et al.*, 2009) can easily be (re-)designed to show causal linkages between levers and performance indicators and illustrate stocks and flows. This research provides empirical evidence from a laboratory dynamic decision making experiment that visualizing causal relations and stock-and-flow structures in a management cockpit indeed improves dynamic decision making performance.

The structure of this paper is as follows. In Section 2, we review the relevant literature on dynamic decision making and derive our hypotheses. Thereafter, the experimental setting and its implementation are described in Section 3. Section 4 comprises the results of the experiment which are discussed in Section 5. The paper closes (Section 6) with a general discussion of implications, limitation of the study and some suggestions for further research.

THEORY

Compared to static decision theories as, for instance, expected utility theory (e. g., Grant & Zandt, 2009; Von Neumann & Morgenstern, 1944), Bayesian decision theory (e. g., Binmore, 2009), prospect theory (Kahneman & Tversky, 1979) or regret theory (Loomes & Sugden, 1982), DDM theories place decision making in realistic settings. DDM research focuses on naturalistic decision challenges where the necessity of decision making is not at all obvious, the set of options is not predefined but has to be determined, the consequences of implementing the various decision alternatives are obscure, criteria for determining preferences have to be established, and so on. Within DDM research, the so called naturalistic decision making stream (Klein, 2008; Lipshitz *et al.*, 2006; Lipshitz *et al.*, 2001; Zsambok & Klein, 1997) seeks to understand how people decide and perform in demanding situations such as fire-fighting using field studies. Similarly, the complex problem solving research camp (Ackerman & Kanfer, 1993; Brehmer & Dörner, 1993; Frensch & Funke, 1995; Wittmann & Hattrup, 2004) focuses on the investigation of naturalistic decisions setting as, for instance, management of a fisheries business or a coal-fired power plant – typically using realistic computer simulations in laboratory research (Brehmer, 1992).

Good decision making in dynamically complex settings requires avoiding a range of typical errors that have been observed using computer-simulated micro worlds in a laboratory. More precisely, the following different tasks that make up the "whole process of action regulation" (Dörner & Schaub, 1994, p. 434) have to be conducted successfully: (1) goal elaboration, (2) hypothesis formation, (3) prognosing, (4) planning, (5) monitoring and (6) self-reflection. Bad-performing decision makers fail in clearly defining their goals and subdividing it into concrete sub goals (Dörner, 1996). Poor decisions also follow from an erroneous and/or incomplete set of hypothesis about the causal structure of the system (e.g., Dörner, 1980; Forrester, 1961). Prognosing (or forecasting) is another sub-task that is error prone – specifically when the forecast stretches far into the (simulated) future. Delays between actions and results are often misperceived (Rahmandad *et al.*, 2009; Sterman, 1989b) and nonlinear causal relations are misjudged (Dörner, 1996). With regard to planning a goal-directed course of action, Dörner and Wearing (1995) identify disregarding side- and long-term effects of actions as the main mistake, which is rooted in an incomplete and/or inaccurate mental model of the system's causal structure. Planning might further be complicated in situations of uncertainty, where action outcomes are not certain, forcing people into making error-prone probabilistic judgments (Kahneman & Tversky, 1972; Tversky & Kahneman, 1974). Typical mistakes in the monitoring stage range from simply just forgetting to deliberately neglecting to monitor previous actions (Dörner & Schaub, 1994). Similarly, the important phase of self-reflection that implies to recognise and analyse the mistakes made in past decisions is often completely abandoned (Dörner, 1996; Dörner & Schaub, 1994; Dörner & Wearing, 1995).

At least the sub tasks (2), (3) and (4) of the above mentioned six DDM phases are compromised by erroneous and vague mental models of the system's causal structure. Indeed, empirical evidence is increasing that causally more accurate mental models result in better decision making performance (Denrell *et al.*, 2004; Dörner & Schaub, 1994; Gary & Wood, 2011). How decision makers' mental models can be improved is a widely neglected issue in DDM research though. Training decision makers in system dynamics would improve their ability to build accurate mental models of a system and better understand stocks and flows (e.g., Akkermans, 1993; Pala & Vennix, 2005; Sterman, 2010). However, such training requires effort and time. Providing decision makers with decision aiding dashboards that make use of core system dynamics tools as causal loop and/or stock-and-flow diagrams would be much easier. Then, these dashboards would not only show performance indicators but also causal links in-between and the system's stock-and-flow structure. However, if and to what extent DDM performance is influenced by different types of dashboards has been neglected in DDM research so far. Obviously, micro worlds used in this research had more or less sophisticated user interfaces presenting varying amounts of information. However, the effects of different presentations on the participants' decision making performance have not been investigated.

Literature on management dashboards or cockpits is primarily practice-oriented and normative (e.g., Daum, 2006; Eckerson, 2011; Rasmussen *et al.*, 2009). However, empirical evidence on the effect of information presentation format on judgement and decision making is provided by information systems and accounting information systems research (see, e.g., the recent review of Kelton *et al.*, 2010). A substantial literature exists on the effects of graphical and tabular representations of information on decision making performance (e.g., Amer, 1991; Cardinaels, 2008; Davis, 1989; Harvey & Bolger, 1996; Schulz & Booth, 1995; Stock & Watson, 1984). It has also been examined how various task characteristics influence the relationship between the external problem representation and problem-solving performance (e.g., Amer, 1991; Benbasat & Dexter, 1985; Coll, 1992; Dennis & Carte, 1998; Speier, 2006). Additional studies focused on the effect of experience, knowledge and ability on the relation between problem representation and decision performance (e.g., Benbasat & Dexter, 1985; Cardinaels, 2008; Coll *et al.*, 1994; Libby & Luft, 1993; Speier *et al.*, 2003). An investigation of the performance impact of providing graphical information on the causal and stock-and-flow structure of the decision making problem, however, is missing.

Following the results from accounting information systems research that information presentation format matters and grounding in the findings of DDM research that more accurate mental models of the causal structure of the DDM task increase performance, we hypothesize that clarifying the causal and stock-and-flow structure in a decision aiding dashboard is beneficial:

> *Hypothesis 1: The clearer the causal and stock-and-flow structure of a dynamic decision making task is shown in a decision aiding management cockpit, the higher is the decision making performance.*

With regard to the complexity of the dynamic decision making task, Gary and Wood (2011, p. 572) propose that "more accurate mental models of the causal relationships in the business environment have a greater positive effect on performance in environments that are more complex". This proposition is mainly build on computational results (Gavetti & Levinthal, 2000; Rivkin, 2000) and has to be tested empirically in the respective context. Based on these propositions we formulate a second hypothesis as follows:

> *Hypothesis 2: The higher the dynamic complexity of the dynamic decision making task, the more decision makers' performance is improved by clarifying the causal and stock-and-flow structure in in a decision aiding management cockpit.*

EXPERIMENTAL DESIGN AND IMPLEMENTATION

Task description

We use an interactive, computer-based simulation of managing short-term capacity dynamics as experimental task in our study. The micro world is developed using the Forio Simulate platform. Participants are asked to take over the role of a team manager in a bank's settlement and clearing of securities department. In this department incoming securities orders have to be processed by employees (see Figure 1). Order processing capacity is determined by the number of employees and their productivity which is assumed constant and known. In case of capacity shortages, orders are backlogged and have to be processed the day after. Employee capacity utilization is derived from the total number of orders processed by the order processing capacity. The service level is determined as the order fill rate. Following Oliva and Sterman (2001), a first order exponential smooth structure is used to model the customers' service quality perception. Whether the changes in the perceived service levels affects demand is determined by the market structure. In a competitive environment, eroding service levels will make customers turn away and place their securities orders elsewhere – which closes both the balancing feedback loop B1 and the reinforcing feedback loop R1 shown in Figure 1. In a monopolistic situation, customers have no choice and there will be no effect of service levels on incoming orders (B1 and R1 both disappear). By closing or cutting through theses loops we can create two dynamic decision making tasks with different levels of dynamic complexity – a closed loop setting with higher dynamic complexity and an open loop version with lower dynamic complexity.

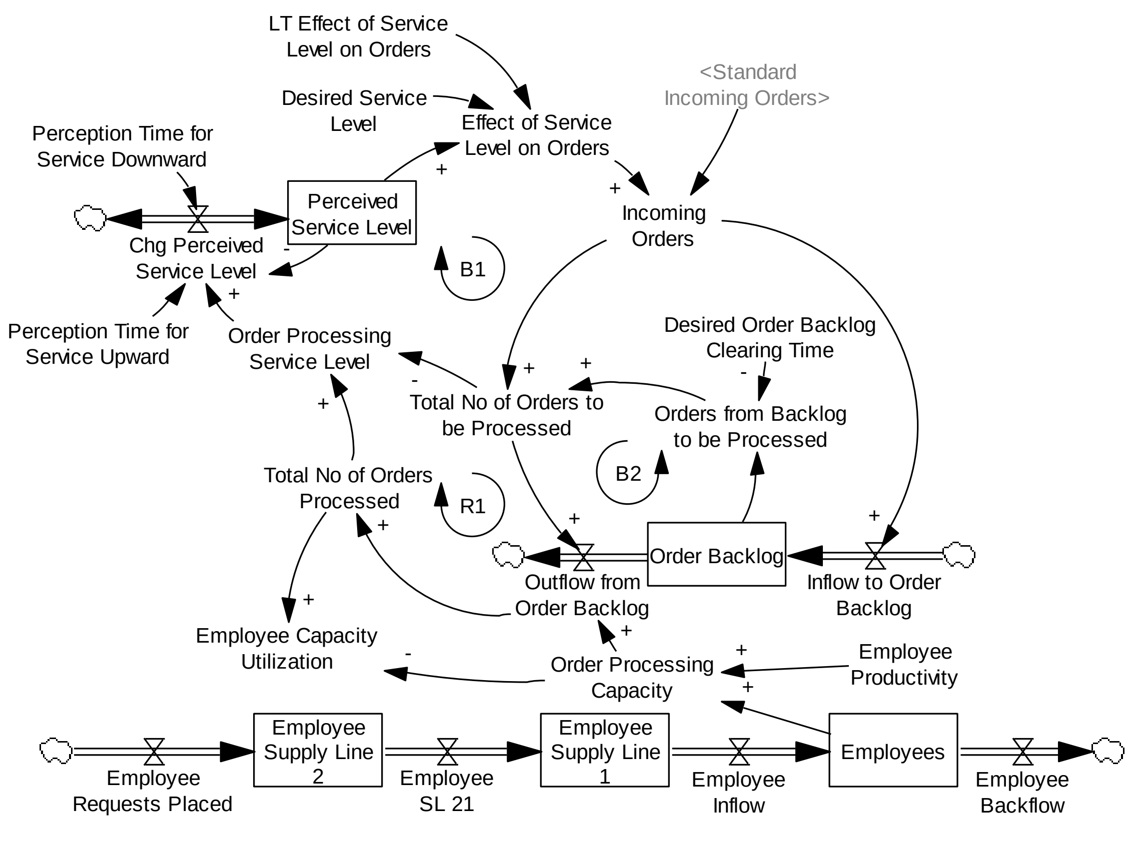

Figure 1: Stock-flow model of the dynamic decision making task

Participants have to decide on a daily basis on the number of full time equivalent employees that they request from the headquarters' employee pool. Regarding Figure 1 they decide on the variable "Employee Requests Placed". They are instructed that they will receive exactly what they requested three days later. By this, the participants have to manage an employee supply chain as shown in the lower part of Figure 1 (using pipeline delays).

The participants make their decision on the number of employees twenty days in a row. They actively move forward from day to day by pressing a button and receiving updated information on the computer screens. Starting point for decision making is day 0. In the instructions, participants receive information back to day -3 as illustrated in Table 1. For days -3 to -1 the system is held in dynamic equilibrium. Then, for day 1, a step-increase in incoming orders is announced (similar to the order pattern typically used in the beer game, e.g., Sterman, 1989b). Participants get the information that an additional fixed order volume of 800 securities orders per day is acquired changing incoming orders from 4,200 to 5,000 per day. Thus, participants are faced with the challenge to bring the system back into dynamic equilibrium as soon as possible.

	Day -3	Day -2	Day -1	Day 0	Day1
Incoming Orders [Orders/Day]	4,200	4,200	4,200	4,200	5,000
No of Employees Requested [FTE/Day]	30	30	30	Your decision	
Productive Employees [FTE]	30	30	30	30	
Order Backlog [Orders]	0	0	0	0	
Capacity Utilization [%]	100 %	100 %	100 %	100 %	
Fill Rate (Service Level] [%]	100 %	100 %	100 %	100 %	

Table 1: Information in the initial decision making situation

The participants' objectives are set slightly differently for the two complexity settings. Participants in the "open loop, low complexity" setting are asked to maximize the capacity utilization under the constraint that orders are processed the same day, that is, the fill rate equals 100 %. In the "closed loop, high complexity" environment the cumulated number of processed orders have to be added as a goal. Therefore, participants are asked to process as many orders as possible and, at the same, time maximize average capacity utilization and average fill rate (service level)!

Measures

Independent Measure: The independent variable in this experiment is the assigned treatment group.

Dependent Measures: Participants have to make the same capacity ordering decision 20 days in a row. Success is measured by (1) the number of processed orders, (2) the average service level and (3) the average capacity usage level.

The number of processed orders is the result of the amount of incoming orders which could be worked on by the demanded employee capacities. The measure is accumulated from day 1 (participants make their first decision in day 0) to day 21 (last decision is made on day 20). The average service level is calculated by dividing the number of processed orders through the number of incoming orders. The average capacity usage is the result of dividing the accumulated capacity of employees and the number of processed orders.

Setting of the experiment

For each of the two complexity settings – (1) the “open loop, low dynamic complexity” and (2) the “closed loop, high dynamic complexity” setting – we use a two-group posttest-only randomized experiment design (e.g., Trochim & Donnelly, 2007) to test our hypothesis. The treatment groups in each complexity setting had access to a dashboard in the capacity management simulator that showed not only the performance indicators but also the causal relations in-between and the stock-and-flow structure as is illustrated in Figure 3 and Figure 5. The stock-and-flow causal dashboards were designed following at large the suggestions from the system dynamics literature (e.g., Morecroft, 2007; Sterman, 2000). Stocks were shown as boxes, inflows and outflows to stocks were symbolized as bold arrows. Causal information links were symbolized as thin curved arrows. Valve and cloud symbols were omitted from the dashboard as these were suspected to cause confusion and questions. Going beyond system dynamics stock-and-flow diagrams, key performance indicators were placed in soft edge boxes to increase their salience.

	Control Group	**Treatment Group**
Open Loop, Low Dynamic Complexity	A Figure 2	B Figure 3
Closed Loop, High Dynamic Complexity	C Figure 4	D Figure 5

Table 2: Four group between-subject experimental design

The control groups received a user interface in the style of a traditional performance management dashboard (see Figure 2 and Figure 4). Of course, all measures from the treatment dashboards were also included in the control group dashboards. By this any difference in the information load were eliminated. The control group dashboards were clearly structured in three categories – KPIs shown at top, order-related indicators grouped in the middle part and capacity-related measures put at the bottom of the dashboard screen. Obviously, no causal links between the indicators or the indicator categories were shown and no indications were provided regarding their stock or flow character. Summarizing, Table 2 illustrates the experimental two-by-two design.

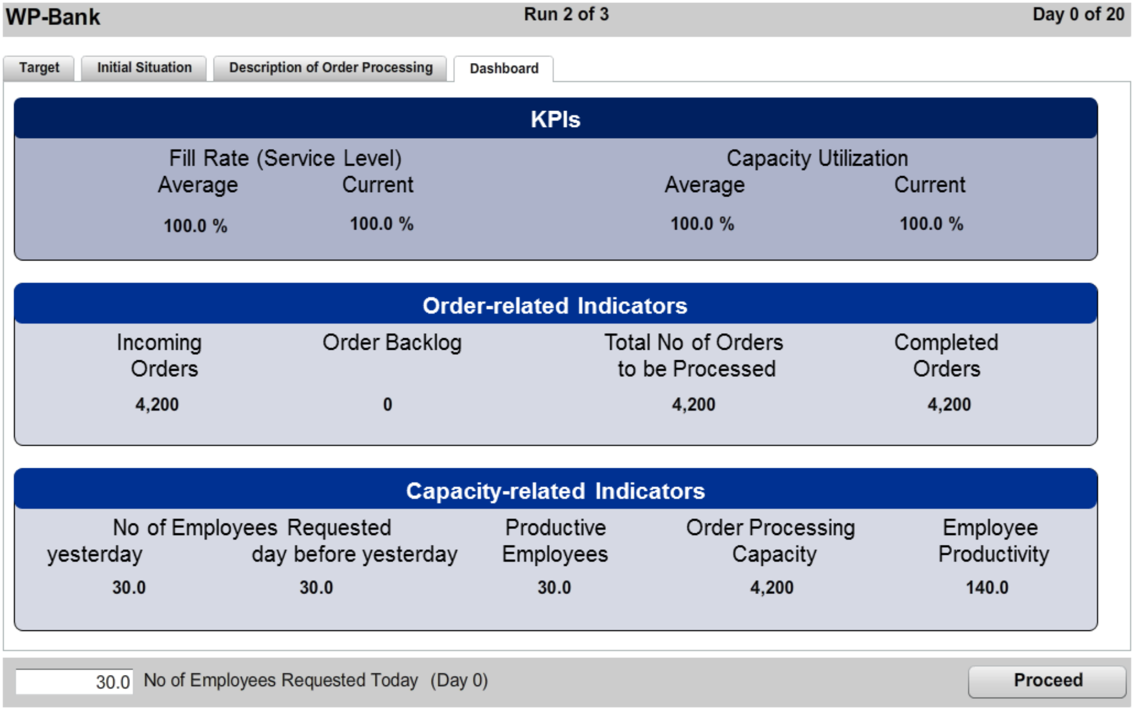

Figure 2: KPI dashboard in the open loop low complexity DDM task

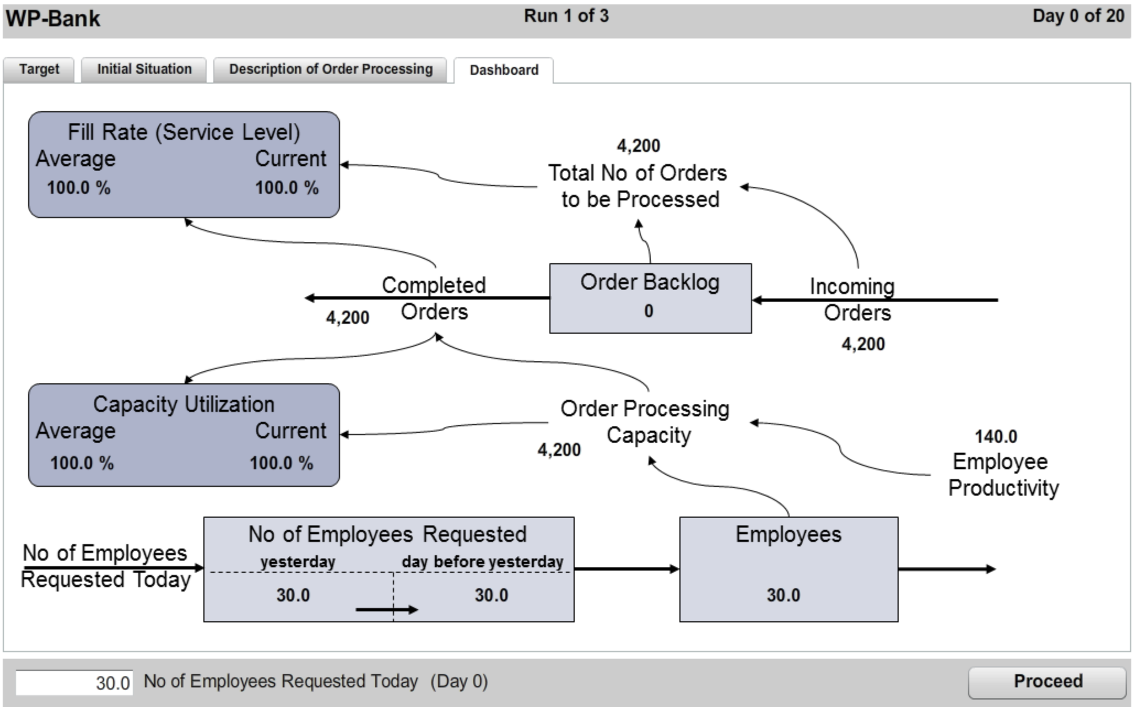

Figure 3: Stock-flow causal dashboard in the open loop low complexity DDM task

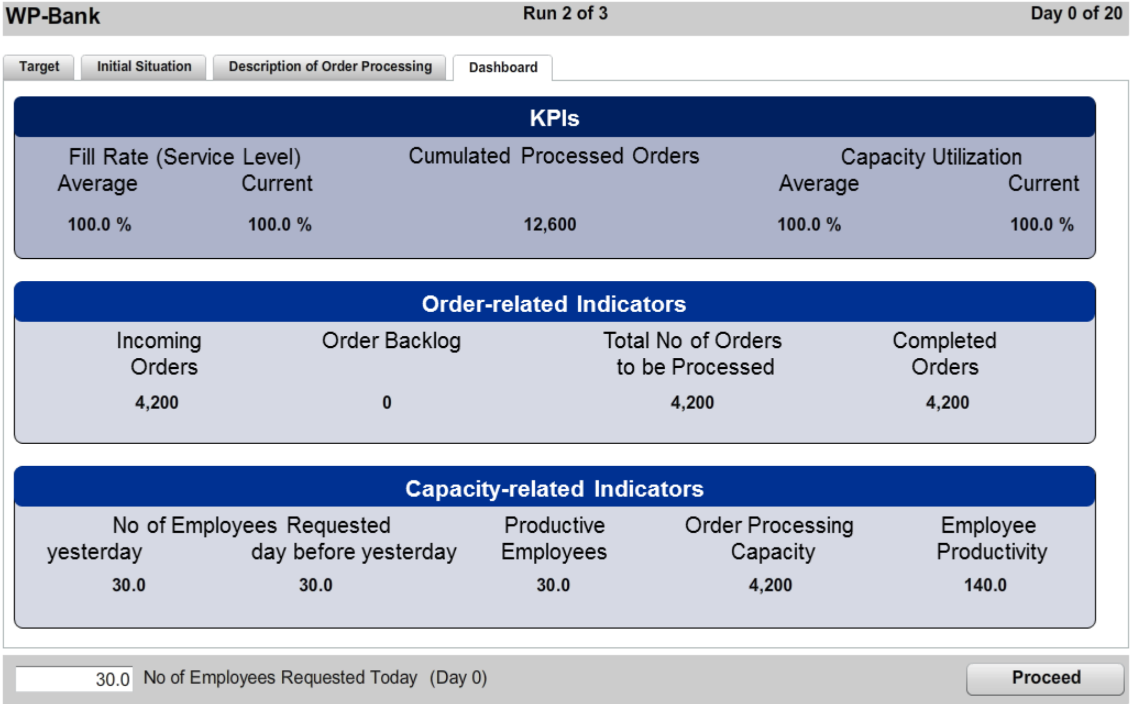

Figure 4: KPI dashboard in the closed loop high complexity DDM task

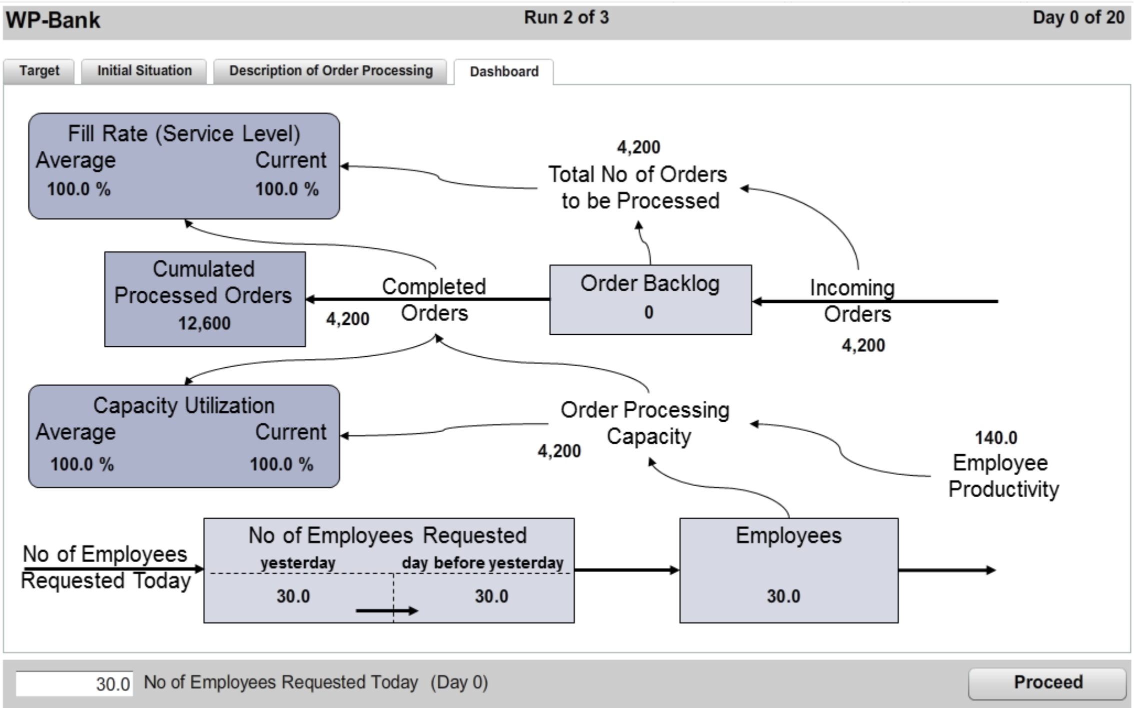

Figure 5: Stock-flow causal dashboard in the closed loop high complexity DDM task

Following suggestions from experimental economics (e.g. Friedman *et al.*, 2004; Guala, 2005), Smith's (1976, 1982) induced value theory is applied and participants are incentivized by a monetary reward. A random lottery design is chosen with chances of winning of 1:11 within each treatment group to ensure a sufficient subjective chance of winning (e.g., Cubitt *et al.*, 1998; Dugar, 2013; Vinogradov & Shadrina, 2013). In case of winning, the financial incentive was linked to this participant's performance. Payment was 20 € minimum and 50 € maximum for participating in an experiment that we expected to take about 15 minutes. An expected pay-out of 80 € to 200 € per hour is deemed an valid incentive (Tversky & Kahneman, 1992). If a participant was selected, the performance of one of the three rounds was chosen.

Participants and procedure

In total, 135 undergraduate management students from five courses participated in the experiment. It was integral part of a course on "Value Chain Management" in which, amongst other topics, capacity management was taught. In all courses, the experiment was conducted before this subject was addressed. Therefore, previous experience of the participants in capacity management was rather low with a value of 2.2 (SD: 1.1) on a five-point Likert-scale. System dynamics in general or stock-and-flow diagrams in particular were not included in the curriculum of our participants. Therefore, we can safely assume that they had no pre-experience concerning these topics.

Participants were assigned randomly to one of the four groups in advance. Due to students not showing up in class on the day of the experiment, the number of participants varies between groups. In the first experiment run, we had 81 participants who were mostly male (77.8 %), thus, 22.2 % female. 39 participants were assigned to the control group and 42 participants to the treatment group. The 54 participants in the second experiment run had a similar distribution regarding gender (74.1 % male, 25.9 % female) as well as regarding previous experience with capacity management 2.2 (SD: 1.1). Here, 25 were assigned to the control group and 29 to the treatment group.

Experiments were carried out in a large computer room. Each participant had a cubicle preventing him to see anything on screens of other participants. An invigilator assured that there was no communication among the participants and no other software than the experimental software was used. The participants started at the same time by login into the capacity management simulation with their personally assigned user data. All essential information about the procedure of the experiment was presented to the participants before the experiment started. After finishing the experiment, each participant had to fill in an additional questionnaire on paper. Every participant could spend as much time as individually needed.

Data Analysis

We apply the recommended modified Kolmogorov-Smirnov-Test of goodness of fit to test whether our data is normally distributed (Yazici & Yolacan, 2006). To test our first hypothesis, we apply a T-Test for normally distributed data and the Mann-Whitney-U-Test for non-parametric data (Ruxton & Beauchamp, 2008).

The second hypothesis is tested with the two-way ANOVA. This statistical method allows the integration of two independent variables, i.e. settings and treatments.

RESULTS FROM THE EXPERIMENT

Descriptives

Average working time on the task in the open loop setting in round one is 10.1 minutes (SD: 3.6) for the control group A and 10.3 minutes (SD: 4.1) for the treatment group B. In the second round values are 3.2 minutes (SD: 2.0) A and 2.9 minutes (SD: 1.9) B. In the closed loop setting the working time on the task in round one is 9.3 minutes (SD: 3.1) for the control group C and 8.2 minutes (SD: 3.1) for the treatment group D. Values for the second round are 2.8 minutes (SD: 1.0) C and 2.9 minutes (SD: 1.5) D. There is no statistical significant difference between the working times of treatment groups in both settings (Setting 1, Round 1: $T(77) = 0.785$, *ns*, Round 2: $T(76) = 0.602$; ns; Setting 2, Round 1: $T(98) = 1.745$, *ns*, Round 2: $T(92) = -.404$, *ns*).

Experimental Results

Mean values and standard deviations of the performance scores in the sub-samples of the different experimental conditions are reported in Table 1.

Measure	**Open loop, low complexity**				**Closed loop, high complexity**			
	Group A		***Group B***		***Group C***		***Group D***	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Average capacity usage	96.0	0.04	96.8	0.05	94.7	0.06	96.3	0.04
Number of processed orders					111783	6955	113503	2983
Average service level	93.7	0.06	92.0	0.11	93.4	0.06	94.8	0.04

Table 1: Overview of performance results

The results regarding the first hypothesis show a statistical significant influence of visualization on the average capacity usage for the low complexity setting ($U(5192.5)$, $p < .001$) as well as for the high complexity setting ($U(8064.5)$, $p < .01$). Regarding the average service level there is no statistical significant improvement in the low complexity setting ($U(6173.5)$, *ns*), but for the high complexity setting ($U(8567)$, $p < .03$). In addition, a higher number of processed orders in the treatment group is confirmed for the high complexity setting ($U(8299.5)$, $p < .01$) while this goal was not relevant in the low complexity setting. Thus, the null-hypothesis cannot be rejected and we see our hypothesis corroborated.

Hypothesis 2, stating that the effect of visualisation is higher in a more complex decision setting, can be corroborated. The average service level is significantly better in the high complexity setting ($F(1) = 6.096$, $p < .01$) while the average capacity usage is not improved significantly in the more complex setting ($F(1) = 0.841$, *ns*). The average service level is dropping in the low complexity setting (-1.8%) and it increases in the high complexity one (1.5%). Regarding the capacity usage, the non-significant increase in the low complexity setting is 0.8% compared to 1.7% in the high complexity setting.

DISCUSSION

The results provide empirical evidence that visualising a causal and stock-and-flow structure of a dynamic decision making task leads to a higher decision making performance. This result holds true independent from the complexity of the two given decision environments. Thus, previous research on the positive effect of graphical information presentation is extended with a specification how such a visualisation should be implemented. The given task requires prognosing, planning and monitoring skills as goals and hypotheses are defined in the settings. As participants have little previous knowledge regarding such decisions, they have to build a mental model to come up with good decisions. Results can be interpreted that the visualisation of causal and stock-and-flow structure helps to build a better mental model. Participants are especially better in increasing the average capacity usage on a similar service level and similar resp. increased number of processed orders. As the number of employees to be ordered is the only decision variable, average capacity usage can be influence more directly than the two other sub measures.

As the effect of visualisation increases in a more complex decision environment, our results are in line with findings of Gavetti and Levinthal (2000) and (Rivkin, 2000) in computational simulations. While these authors found non-empirical evidence that for more complex decisions a higher accuracy of mental models is required, we find empirical evidence that visualising causal and stock-and-flow logic helps to build such models better with a higher complexity. Gary and Wood (2011), also building their hypothesis on the findings of Gavetti and Levinthal as well as Rivkin, did not find such an impact for their decision environment. However, Gary and Wood also recognised that the environments used were already on a high level of complexity hardly allowing determining the increasing need of mental models in more complex situations. Thus, they could state that the mental model has to have a high level of accuracy in complex decision environments. Concluding, it seems that the mental model is more complicated and that causal and stock-and-flow diagrams are the adequate visualisation helping to build this model. Having a look at the sub measures, we can observe that the main effect is on an improved average capacity usage between setting 1 and 2. Participants seem to need the visualisation in the more complex environment to better understand the impact of their delayed capacity ordering decision and can best transfer this on improving the average capacity utilisation.

CONCLUSIONS, LIMITATIONS, AND FURTHER RESEARCH

Our research contributes to the literature by investigating the performance impact of providing causal and stock-and-flow structure of the decision making model. We can show that such a visualisation leads to a better mental model and thus better decisions. The effect is even higher in more complex decision environments. Concluding, such visualisations should be used to support decision making situations.

Two limitations have to be noted from the authors’ point of view. Participants were not very familiar with the decision making problem, thus, a comparison with participants having much knowledge of such situations was not included. Furthermore, as the working time for the third round was not recorded due to a system failure, a possible mediating effect of working time on performance could not be analysed.

Within further research three major streams are of interest. Firstly, given the two levels of complexity it would be worthwhile to analyse the effect in even more complex decision settings. A suitable extension could be the introduction of a random stochastic demand rate which would make the decision less predictable. Conducting such an analysis could reveal whether there is a linear increase in performance or if there is a tipping point from which the effect of visualisation is constant. Secondly, a closer analysis of how the mental model is influenced by the causal and stock-and-flow diagram would be helpful. Here, qualitative interviews could be conducted questioning participants afterwards and compare results with their performance. Thirdly, the learning effect over the three rounds could be further examined to understand the impact of the treatments on the learning process.

REFERENCES

Ackerman PL, Kanfer R. 1993. Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic controller success. *Journal of Applied Psychology* **78**(3): 413-432

Akkermans HA. 1993. Participative Business Modelling to Support Strategic Decision Making in Operations - A Case Study. *International Journal of Operations & Production Management* **13**(10): 34-48

Amer T. 1991. An Experimental Investigation of Multi-Cue Financial Information Display and Decision Making. *Journal of Information Systems* **5**(2): 18-34

Barlas Y, Özevin MG. 2004. Analysis of stock management gaming experiments and alternative ordering formulations. *Systems Research and Behavioral Science* **21**(4): 439-470

Barr PS, Stimpert JL, Huff AS. 1992. Cognitive change, strategic action, and organizational renewal. *Strategic Management Journal* **13**(S1): 15-36

Beckmann JF, Guthke J. 1995. Complex problem solving, intelligence, and learning ability. In PA Frensch, J Funke (Eds.), *Complex problem solving: The European perspective*: 177-200. Erlbaum: Hillsdale

Benbasat I, Dexter AS. 1985. An Experimental Evaluation of Graphical and Color-Enhanced Information Presentation. *Management Science* **31**(11): 1348-1364

Binmore K. 2009. *Rational decisions*. Princeton Univ. Press: Princeton, NJ

Brehmer B. 1992. Dynamic decision making: human control of complex systems. *Acta Psychologica* **81**: 211-241

Brehmer B, Dörner D. 1993. Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior* **9**(2-3): 171-184

Cardinaels E. 2008. The interplay between cost accounting knowledge and presentation formats in cost-based decision-making. *Accounting, Organizations and Society* **33**(6): 582-602

Charnes A, Cooper WW, Rhodes E. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* **2**(6): 429-444

Checkland P. 1981. *Systems thinking, systems practice*. Wiley: Chichester

Checkland P, Scholes J. 1990. *Soft systems methodology in action*. Wiley: Chichester

Coll JH. 1992. An experimental study of the efficacy of tables versus bar graphs with respect to type of task. *Information & Management* **23**(1): 45-51

Coll RA, Coll JH, Thakur G. 1994. Graphs and tables: a four-factor experiment. *Commun. ACM* **37**(4): 76-86

Cooper WW, Seiford LM, Tone K. 2007. *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software* (2nd ed. ed.). Springer: New York

Cronin M, Gonzalez C. 2007. Understanding the building blocks of system dynamics. *System Dynamics Review* **23**(1): 1-17

Cubitt R, Starmer C, Sugden R. 1998. On the validity of the random lottery incentive system. *Experimental Economics* **1**(2): 115-131

Daum JH. 2006. Management Cockpit War Room: Ziele, Funktionsweise und Zukunftsperspektiven eines (noch) ungewöhnlichen, aber hocheffektiven Managementinstruments. *Controlling* **18**(6): 311-318

Davis LR. 1989. Report format and the decision maker's task: An experimental investigation. *Accounting, Organizations and Society* **14**(5–6): 495-508

Dennis AR, Carte TA. 1998. Using Geographical Information Systems for Decision Making: Extending Cognitive Fit Theory to Map-Based Presentations. *Information Systems Research* **9**(2): 194-203

Denrell J, Fang C, Levinthal DA. 2004. From T-Mazes to Labyrinths: Learning from Model-Based Feedback. *Management Science* **50**(10): 1366-1378

Diehl E, Sterman JD. 1995. Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes* **62**(2): 198-215

Dörner D. 1980. On the difficulties people have in dealing with complexity. *Simulations and Games* **11**(1): 87-106

Dörner D. 1996. *The logic of failure. Strategic thinking for complex situations*. Metropolitan Books: New York, NY

Dörner D, Kreuzig HW, Reither F. 1994. *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Huber: Bern et al.

Dörner D, Schaub H. 1994. Errors in Planning and Decision-making and the Nature of Human Information Processing. *Applied Psychology* **43**(4): 433-453

Dörner D, Wearing AJ. 1995. Complex Problem Solving: Toward a (Computersimulated) Theory. In PA Frensch, J Funke (Eds.), *Complex Problem Solving: The European Perspective*: 65-99. Erlbaum: Hillsdale

Dugar S. 2013. Non-monetary Incentives and Opportunistic Behavior: Evidence from a Laboratory Public Good Game. *Economic Inquiry* **51**(2): 1374-1388

Eckerson WW. 2011. *Performance dashboards: Measuring, monitoring, and managing your business* (2nd ed. ed.). Wiley: Hoboken, N.J.

Edwards W. 1962. Dynamic decision theory and probabilistic information processing. *Human Factors* **4**(2): 59-73

Epstein MK, Henderson JC. 1989. Data Envelopment Analysis for Managerial Control and Diagnosis*. *Decision Sciences* **20**(1): 90-119

Forrester JW. 1961. *Industrial dynamics*. Productivity Press: Cambridge, MA

Fox J, Cooper RP, Glasspool DW. 2013. A canonical theory of dynamic decision-making. *Frontiers in Psychology* **4**

Frensch PA, Funke J. 1995. *Complex problem solving: The European perspective*. Erlbaum: Hillsdale

Friedman D, Cassar A, Selten R. 2004. *Economics lab: An intensive course in experimental economics*. Routledge: London

Gary MS, Wood RE. 2011. Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal* **32**(6): 569-594

Gavetti G, Levinthal D. 2000. Looking Forward and Looking Backward: Cognitive and Experiential Search. *Administrative Science Quarterly* **45**(1): 113-137

Grant S, Zandt TV. 2009. Expected Utility Theory. In P Anand, P Pattanaik, C Puppe (Eds.), *Handbook of Rational and Social Choice*. Oxford University Press: New York, NY

Guala F. 2005. *The methodology of experimental economics*. Cambridge University Press: Cambridge

Harvey N, Bolger F. 1996. Graphs versus tables: Effects of data presentation format on judgemental forecasting. *International Journal of Forecasting* **12**(1): 119-137

Kahneman D, Tversky A. 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology* **3**(3): 430-454

Kahneman D, Tversky A. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **47**(2): 263-291

Kelton AS, Pennington RR, Tuttle BM. 2010. The Effects of Information Presentation Format on Judgment and Decision Making: A Review of the Information Systems Research. *Journal of Information Systems* **24**(2): 79-105

Klein G. 2008. Naturalistic decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society* **50**(3): 456-460

Kleinmuntz DN. 1985. Cognitive heuristics and feedback in a dynamic decision environment. *Management Science* **31**(6): 680-702

Libby R, Luft J. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society* **18**(5): 425-450

Lipshitz R, Klein G, Carroll JS. 2006. Introduction to the special issue. Naturalistic decision making and organizational decision making: Exploring the intersections. *Organization Studies* **27**(7): 917-923

Lipshitz R, Klein G, Orasanu J, Salas E. 2001. Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making* **14**(5): 331-352

Loomes G, Sugden R. 1982. Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *The Economic Journal* **92**(368): 805-824

Morecroft JDW. 2007. *Strategic Modelling and Business Dynamics: A Feedback Systems Approach* Wiley Chichester

Moxnes E. 1998. Not only the tragedy of the commons: Misperceptions of bioeconomics. *Management Science* **44**(9): 1234-1248

Oliva R, Sterman JD. 2001. Cutting Corners and Working Overtime: Quality Erosion in the Service Industry. *Management Science* **47**(7): 894

Paich M, Sterman JD. 1993. Boom, bust, and failures to learn in experimental markets. *Management Science* **39**(12): 1439-1458

Pala Ö, Vennix JAM. 2005. Effect of System Dynamics Education on Systems Thinking Inventory Task Performance. *System Dynamics Review* **21**(2): 147-172

Rahmandad H, Repenning N, Sterman J. 2009. Effects of feedback delay on learning. *System Dynamics Review* **25**(4): 309-338

Rasmussen N, Chen CY, Bansal M. 2009. *Business dashboards: A visual catalog for design and deployment*. Wiley: Hoboken, N.J.

Reichert U, Dörner D. 1988. Heurismen beim Umgang mit einem "einfachen" dynamischen System. *Sprache und Kognition* **7**(1): 12-24

Rivkin JW. 2000. Imitation of Complex Strategies. *Management Science* **46**(6): 824-844

Ruxton GD, Beauchamp G. 2008. Some suggestions about appropriate use of the Kruskal–Wallis test. *Animal Behaviour* **76**(3): 1083-1087

Schulz AKD, Booth P. 1995. The Effects of Presentation Format on the Effectiveness and Efficiency of Auditors' Analytical Review Judgments. *Accounting & Finance* **35**(1): 107-131

Smith VL. 1976. Experimental economics: Induced value theory. *American Economic Review* **66**(2): 274

Smith VL. 1982. Microeconomic system as an experimental science. *American Economic Review* **72**(5): 923

Speier C. 2006. The influence of information presentation formats on complex task decision-making performance. *International Journal of Human-Computer Studies* **64**(11): 1115-1131

Speier C, Vessey I, Valacich JS. 2003. The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance. *Decision Sciences* **34**(4): 771-797

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. 1994. Learning in and about complex systems. *System Dynamics Review* **10**(2-3): 291–330

Sterman JD. 2000. *Business dynamics: Systems thinking and modeling for a complex world*. Irwin/McGraw-Hill: Boston, MA

Sterman JD. 2010. Does formal system dynamics training improve people's understanding of accumulation? *System Dynamics Review* **26**(4): 316-334

Stock D, Watson CJ. 1984. Human Judgment Accuracy, Multidimensional Graphics, and Humans Versus Models. *Journal of Accounting Research* **22**(1): 192-206

Strohhecker J, Größler A. 2013. Do personal traits influence inventory management performance?—The case of intelligence, personality, interest and knowledge. *International Journal of Production Economics* **142**(1): 37-50

Süß H-M. 1996. *Intelligenz, Wissen und Problemlösen. Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen*. Hogrefe: Göttingen et al.

Trochim WMK, Donnelly JP. 2007. *The research methods knowledge base* (3 ed.). Atomic Dog Publishing

Tversky A, Kahneman D. 1974. Judgment under uncertainty: Heuristics and biases. *Science* **185**(4157): 1124-1131

Tversky A, Kahneman D. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5**(4): 297-323

Vinogradov D, Shadrina E. 2013. Non-monetary incentives in online experiments. *Economics Letters* **119**(3): 306-310

Von Neumann Jv, Morgenstern O. 1944. *Theory of games and economic behavior*. Princeton Univ. Press: Princeton

Wittmann WW, Hattrup K. 2004. The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science* **21**(4): 393-409

Yazici B, Yolacan S. 2006. A comparison of various tests of normality. *Journal of Statistical Computation and Simulation* **77**(2): 175-183

Zsambok CE, Klein GA. 1997. *Naturalistic decision making*. L. Erlbaum Associates: Mahwah, N.J.