Paper presented at the 33rd International Conference of the System Dynamics Society, July 19-23, 2015, Cambridge, MA

The more you see the less you "get":

On the importance of a higher-level perspective

for understanding dynamic systems

Helen Fischer[1*], Florian Kapmeier[2], Mihaela Tabacaru[3], and Birgit Kopainsky[3],

[1] University of Heidelberg, Department of Psychology, Hauptstr. 47-51, 69117 Heidelberg, Germany

[2] ESB Business School, Reutlingen University
Alteburgstraße 150, 72762 Reutlingen, Germany

[3] System Dynamics Group, Department of Geography, University of Bergen,
P.O. Box 7800, 5020 Bergen, Norway

*corresponding author: helen.fischer@psychologie.uni-heidelberg.de

Version 1.1

August 2015

Abstract

This paper describes the design and outcomes of an experimental study that addresses stock-and-flow-failure from a cognitive perspective. It is based on the assumption that holistic (global) and analytic (local) processing are important cognitive mechanisms underlying the ability to infer the behavior of dynamic systems. In a stock-and-flow task that is structurally equivalent to the department store task, we varied the format in which participants are primed to think about an environmental system, in particular whether they are primed to concentrate on lower-level (local) or higher-level (global) system elements. 148 psychology, geography and business students participated in our study. Students' answers support our hypothesis that global processing increases participants' ability to infer the overall system behavior. The beneficial influence of global presentation is even stronger when data are presented numerically rather than in the form of a graph. Our results suggest presenting complex dynamic systems in a way that facilitates global processing. This is particularly important as policy-designers and decision-makers deal with complex issues in their everyday and professional life.

Introduction

Booth Sweeney & Sterman, (2000) show that people have difficulties understanding the relationship between a system's structure and its behavior over time – even when the system structure is regarded as fairly simple. This phenomenon is called stock-and-flow-(SF-) failure (Cronin et al., 2009) and has been reproduced by scholars world-wide (i.e., Brunstein et al., 2010; Cronin & Gonzalez, 2007; Cronin, et al., 2009; Kainz & Ossimitz, 2002; Kapmeier, 2004; Kapmeier et al., 2014; Kapmeier & Zahn, 2001; Moxnes & Jensen, 2009; Ossimitz, 2002; Sterman & Booth Sweeney, 2002; Gonzalez & Wong, 2012). Reasons for SF-failure can be categorized in systems thinking skills, domain-specific experience and knowledge, and visualization (Kapmeier, et al., 2014).

First, referring to lacking systems thinking skills, scholars have examined whether people correctly understand the dynamics of stocks and flows. As already identified in the original study (Booth Sweeney & Sterman, 2000), participants very often assume a positive correlation between the inflow, or the input variable, and the stock, the output variable – or, how we are going to argue later in this paper, a positive correlation between a lower-level element and a higher-level element. This assumed correlation is called pattern matching (Booth Sweeney & Sterman, 2000) or correlation heuristic (Cronin, et al., 2009).

Second, scholars have identified an impact of previous knowledge on SF-performance (Booth Sweeney & Sterman, 2000; Brunstein, et al., 2010; Cronin & Gonzalez, 2007; Cronin, et al., 2009; Kapmeier, 2004; Kapmeier, et al., 2014). This stream of literature indicates that participants use background knowledge, for example from their education, or their previous professional experience to solve a SF-task.

Third, scholars have analyzed the impact of visualization on the understanding of the interplay between stocks and flows, or lower-level and higher-level variables. According to Kainz

& Ossimitz, (2002) and Sterman, (2002), one of the arguments explaining the low SF-performance refers to the proposition that people do better when working with numerical data than when reading graphical representations of the same data. Although some scholars (Cronin & Gonzalez, 2007; Schwarz et al., 2013; Sedlmeier et al., 2014) have tested different ways for visualization and not observed improvements in SF-performance, Veldhuis & Korzilius, (2012) have found that "the visualization dimension has a positive effect on performance in various systems thinking inventory tasks and a negative effect on the likelihood that the participant selects a response typical for correlation heuristic reasoning" (p. 1).

The main focus in this paper is on the first explanation for SF-failure, lack of systems thinking skills. We vary the format in which participants are primed to think about a dynamic system, in particular whether they are primed to concentrate on lower-level or higher-level system elements. However, we also include findings on the two remaining categories of SF-failure, visualization and previous knowledge in the design of an integrated experimental study.

Systems as interrelations between lower-level and higher-level elements

We propose that SF-failure is linked to the structure of dynamic systems, specifically to a focus on lower-level elements. Dynamic systems can be seen as hierarchical: through the interrelations between lower-level elements, increasingly higher-level elements arise. In that sense, velocity, for example, is a higher-order variable that relates time and distance. The hierarchy of systems can thus be seen as a product of relations between more and more abstract entities. Most importantly, for human understanding of dynamic systems, people may possess properties on a higher, macro level, that none of the elements on lower, micro levels possess and that are unpredictable from isolated lower-level system parts (Wilensky & Resnick, 1999). For illustration, take the system of a fishpond consisting of lower-level extraction and reproduction rate and the higher-level stock of fish. Since the properties of the stock of fish arise through the interplay of extraction and reproduction, the stock can be interpreted as a higher-order variable rather than extraction or reproduction. Consequently, the stock may decrease, for example, even though reproduction and extraction each are increasing. Or, in system-dynamics-terminology, a system can be analyzed on different levels of detail. Coyle, (1996), for instance, describes how a system is often initially observed on the level of symptoms that are of concern (see Figure 1). This happens usually on a high level of abstraction. Yet, analyzing the underlying causes of the symptoms happens on a level of higher detail, or, with a detailed stock-and-flow-diagram that requires detailed and specific knowledge and understanding. Insights into the problem-solving are then presented on a higher level of abstraction, as this is usually easier for people to understand. Coyle stresses that, when traveling through the cone, there is conceptual consistency; at the same time, there can be different names for the variables on the different levels. Sterman (2000), for example, describes how decision-makers at General Motors were able to decide on a specific strategic question on car leasing only after the complex issue was presented to them in a simple picture of a bathtub with inflows and outflows – a comprehensive system dynamics analysis would have been too detailed for the decision-makers to understand. Similarly, Sterman's (2009) explanation of global warming in Science is also on a highly aggregated level, providing fundamental insight – whereas Fiddaman's (1997) thesis, on which the insight is based on, provides much fine detail and equations on a computer coding level.

Wilensky & Resnick, (1999) idea of higher-level variables and lower-level variables and Coyle's (1996) cone could be related to the original Bathtub task and the Cash flow task by Booth Sweeney and Sterman (2000) as the higher-level variable is the water level in the bathtub or the cash in a bank account – or the symptom of concern - , whereas lower-level variables – or the underlying, more detailed variables – are water flows in and out of the bathtub and receipts to and expenditures from a bank account.

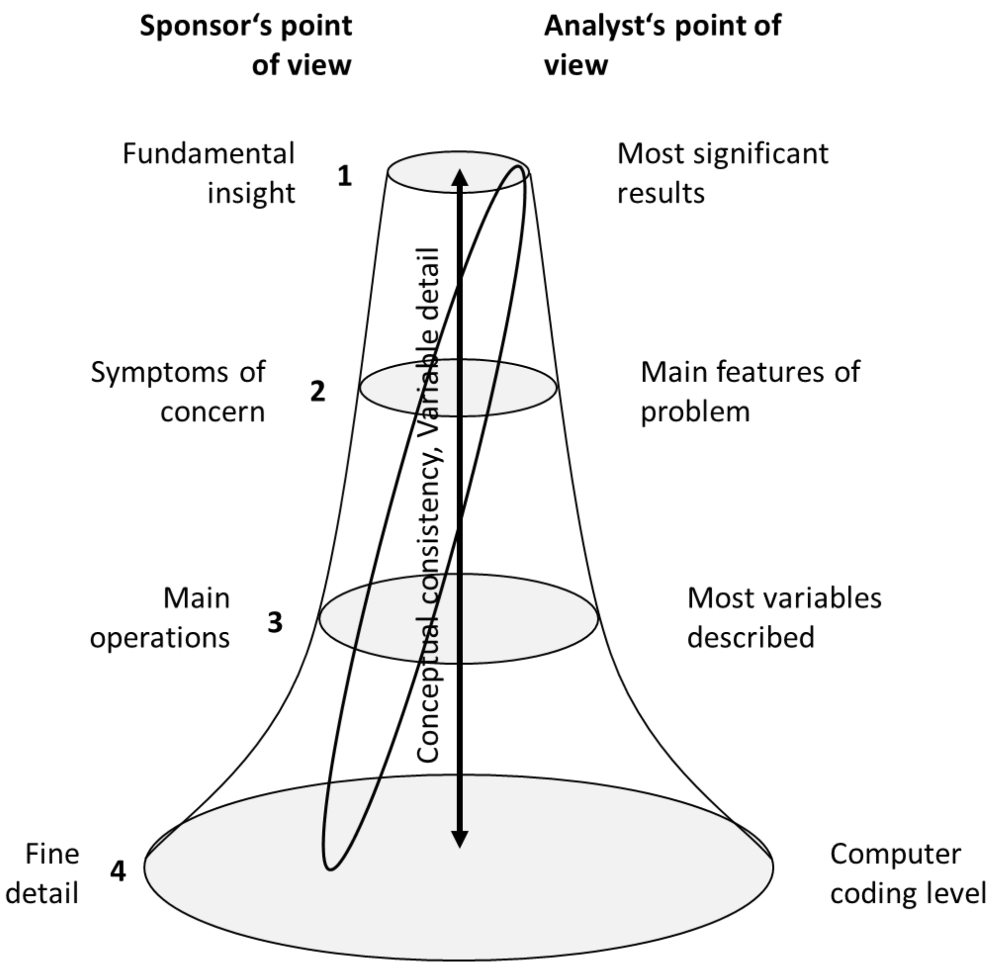


Figure 1: The cone represents conceptual consistency throughout different levels of details in simulation models (source: Coyle, 1996: 346)

We propose that it is therefore necessary to adopt higher-level thinking when trying to infer the behavior of higher-level system elements. That is, it is necessary to engage in (an adequate level of) abstraction: Instead of focusing on low-level, local system elements in isolation, it is necessary to adopt a more global perspective, taking into account the interrelations between system elements.

Higher-level properties cannot be inferred from observing lower-level elements in isolation; they are *irreducible* to isolated lower-level elements. Since lower-level elements are usually those entities that are readily observable (Burgoon et al., 2013), however, we may be tempted to simply reason over lower-level system elements and erroneously conclude that higher-level elements should possess similar properties. It therefore seems warranted to speculate that one of the main reasons for failing to understand dynamic complexity stems from the fact that what is easily observable in a dynamic system, may well lead us astray.

In previous research, Fischer & Gonzalez (2015) showed that such as the gestalt of a hierarchical figure can only be recognized when attending to how the local letters are spatially related, the gestalt of a dynamic system can only be recognized when attending to how its *elements are structurally related*. Similarly to how Hämäläinen et al. (2013) argued that the format dynamic systems were typically presented in (e.g., Cronin, et al., 2009) might have triggered erroneous reasoning strategies is general, we argue specifically that the format might have triggered local processing. This might be the case because the questions were worded such that a focus on isolated, lower-level system elements was induced. For clarification, take the example of the department store task (Sterman, 2002). In the Department Store task, questions focus on specific points in time and specific, isolated numbers of people (e.g., minute 8; 17 people entering). Thus, participants might get the impression that they need to work on lower-level system elements (such as isolated, single points in time), instead on *relations between* elements (such as the over-time relations between people entering and leaving). To correctly infer the overall system behavior, that is, the behavior of the stock, it is vital that participants work on higher-level system elements. In other words, participants need to engage in global processing.

Building on these arguments and results, we aim at investigating how abstraction (global as opposed to local processing) as a cognitive process affects people's understanding of dynamic systems to gain deeper insight into the cognitive processes underlying dynamic systems understanding. We investigate this by systematically varying the format a dynamic system is presented in: A local format inducing a focus on isolated, lower-level system elements versus a global format inducing a more abstract focus on interrelated, higher-level system elements. That is, we vary the format in which a dynamic system is presented in such a way that abstract processing of the system is made more or less likely. By doing so, we test the following:

(a) Does a format that induces abstract (global) system processing increase people's ability to infer the overall system behavior (such as the development of the number of tress) as opposed to a system that induces local system processing?

(b) Does a format that induces abstract (global) system processing reduce people's tendency to believe the output of a system should simply be linearly correlated with its isolated input? That is, does abstract system processing reduce people's tendency to use the correlation heuristic?

Methods

Task and materials

We tested our research questions in a laboratory experiment in which participants had to solve a SF-task about an environmental system. In doing so, we are addressing the three reasons of SF-failure mentioned above. Firstly, systems thinking, the first reason for SF-failure, is covered by the task design, in particular, the question format inducing abstract (global) system processing. The task is structurally equivalent to the department store task developed by Sterman (2002) and which was used by Fischer and Gonzalez' (in press) previous study on the impact of global-local processing. Like in the department store task, participants were presented with information about the development of an inflow to and an outflow out of a stock over time.

Secondly, we modify the cover story to test previous results (Fischer & Gonzalez, in press) and to account for potential influences arising from the application domain of the task. By doing so, we account for domain-specific knowledge, the second reason for SF-failure. With planting and cutting trees in the rain forest, we refer to a contemporary, everyday issue in natural resource management. The particular system described consists of a single stock, (trees), with an inflow (planting trees) and an outflow (harvesting trees). In particular, the cover story for the SF-task relates to trees planted and harvested in the Brazilian Amazon region, thus changing an imaginary stock of trees:

In a fictional place in the Brazilian Amazon region, the evolution of trees for the period between 1990 and 2012 is analyzed. Trees are harvested and at the same time, a reforestation program plants new trees.

Lastly, we cover visualization as the third reason for SF-failure by displaying the SF-task in different ways. The combination of different ways of visualizing the SF-task and formulating questions results in four experimental conditions (see Figure 2). In all conditions, participants are asked to answer four questions. The first two questions refer to the inflow and the outflow while the subsequent two questions test whether participants are able to infer the behavior of the stock over time, based on the behavior of the flows. In order to answer the questions, participants are required to analyze data that are presented in two different ways.

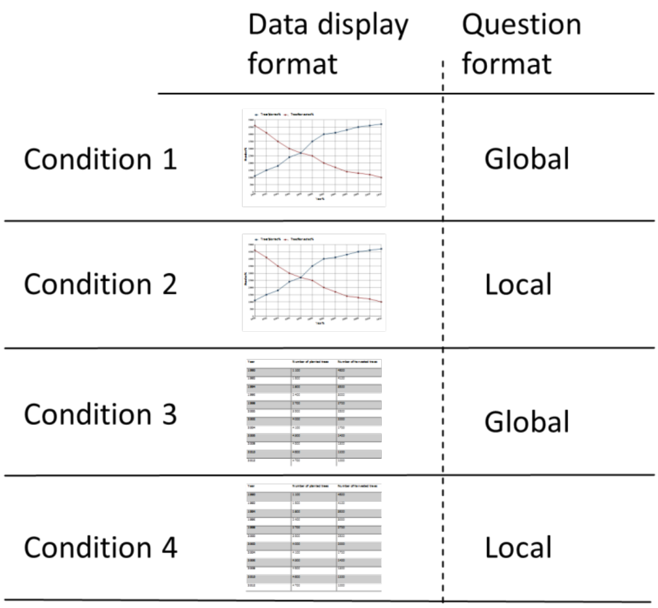
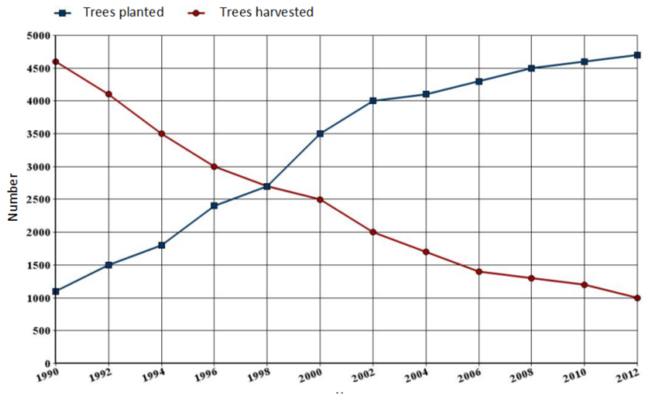


Figure 2: Experimental conditions: The conditions vary in data display (left) and the question format (right)

Data are displayed either as a graph over time or numerically as a table (see Figure 2). We varied the format of displaying the data given the inconclusive findings in previous studies regarding the effectiveness of graphical versus numerical formats of visualizing a SF-task. Note that the number of trees planted increases from 1,100 in 1990 until 4,700 in 2012, whereas the number of trees harvested decreases from 4,700 in 1990 until 1,000 in 2012. Both numbers are identical in the year 1998.

Graphical representation



Numerical representation

Year	Number of planted trees	Number of harvested trees
1990	1100	4600
1992	1500	4100
1994	1800	3500
1996	2400	3000
1998	2700	2700
2000	3500	2500
2002	4000	2000
2004	4100	1700
2006	4300	1400
2008	4500	1300
2010	4600	1200
2012	4700	1000

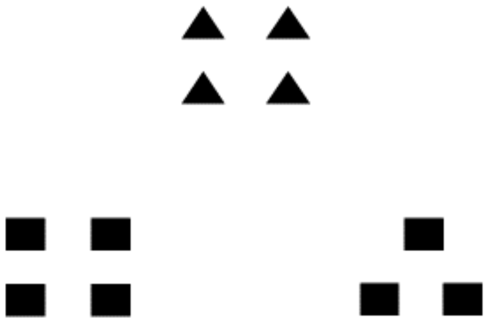
Figure 3: Data are displayed as graphs over time (left) or as numerical tables (right).

There are four SF-task questions to each dataset. They are formulated in such a way that they either induce global or local processing (see Figure 4). Questions 1 and 2 asked whether more trees were planted than harvested in the time periods between 1990 and 1998, or between 1998 and 2012, respectively. Questions 3 and 4 referred to the development of the stock of trees during the same time periods. Questions formulated in a format that induces global processing refrained from asking about the value of an inflow or an outflow at a specific point in time. Instead, they focus on the ratio between the inflow and the outflow for the two main time periods in the task (Questions 1 and 2). Similarly, the global question format concentrates on the development of the stock of trees over time instead of asking about one specific year as the local question format does (Questions 3 and 4).

Inducing global processing	Inducing local processing
Please answer the following questions: 1) What is the ratio between planted trees and harvested trees between 1990 and 1998? a) More trees are planted than harvested. ☐ b) More trees are harvested than planted. ☐ c) The same number of trees are planted and harvested. ☐ 2) What is the ratio between planted trees and harvested trees between 1998 and 2012? a) More trees are planted than harvested. ☐ b) More trees are harvested than planted. ☐ c) The same number of trees are planted and harvested. ☐ 3) How would you describe the evolution of the number of trees between 1990 and 1998? a) Increasing ☐ b) Decreasing ☐ c) Stable ☐ 4) How would you describe the evolution of the number of trees between 1998 and 2012? a) Increasing ☐ b) Decreasing ☐ c) Stable ☐	Please answer the following questions: 1) In which year were most trees harvested? Year __________ . 2) In which year were most trees planted? Year __________ . 3) In which year was the number of trees highest? Year __________ . 4) In which year was the number of trees lowest?

Figure 4: Question format: There are four questions for inducing both, global (left-hand-side) and local (right-hand-side) processing. Formulation of the questions differ

SF-accuracy is the dependent variable in all formats. Previous research has shown that individual processing styles affect solution rates to SF-problems (Fischer & Gonzalez, 2013). Individual, more global processing is connected to higher solution rates in SF-tasks than using the original department store task questions (8% local vs. 24% global, Fischer & Gonzalez, 2013) . To measure whether solution rates in our SF-task were due to the format in which the task was presented or whether they were caused by individual global-local processing styles, we used the Kimchi-Palmer-Figures task (Kimchi & Palmer, 1982). The task shows triangles and squares consisting of smaller triangles and squares. For each of the 16 figures, participants indicate whether the figure on the top (e.g., a global triangle made of local squares) appears to them more similar to a sample figure that matched its global or its local form. In Figure 5, The sample figure on the left hand side is the global match of the figure on the top and the sample figure on the right is the local match. We rated each participant in a range from 0 (completely local processing style over the 16 figures) to 1 (completely global processing style).



Is the figure at the **top** more similar to the **left** figure at the bottom or the **right** figure at the bottom?

Left Right

Figure 5: First Kimchi-Palmer figure in the study to measure individual global-local processing styles (e.g., Kimchi & Palmer, 1982)

Participants

A total of 182 participants with a mean age of 21.3 years (SD = 3.4) took part in the study. The study was administered to 64 psychology students, 49 geography students, and 68 management students at universities in Germany. The questionnaire was handed out to the geography, psychology, and management students in January and February 2015. Students are enrolled in BA program in Geography and Psychology programs at the University of Heidelberg. Management students are enrolled in the 5^{th} semester of the BSc program in International Business at the ESB Business School at Reutlingen University. Participation was voluntary and anonymous. Performance in the study could not have any impact on participants' course grades. Participants were told in the beginning that they could withdraw from the study at any time without any penalty and that in this case, their data would be destroyed. None of the participants made use of this option.

Procedure

Participants were informed that the study takes approximately 10 minutes. They were randomly assigned to one of the four conditions. After studying instructions, participants first completed the Kimchi-Palmer-Figures task that measures individual global-local processing style. Participants then answered the SF-task presented in their respective format.

Results

We computed the mean solution for each task and for each participant separately for the questions on flows, or the lower-level system elements (Q1 and Q2), and the questions about the stocks, or higher-level system element (Q3 and Q4) (Table 1). The results concerning participants' ability to infer the overall system behavior, that is results concerning Q3 and Q4 are of particular interest.

	Graphical data display		Numerical data display	
	Local processing	Global processing	Local processing	Global processing
Lower system elements Q1 & Q2	.93	.68	.98	.84
Higher system elements Q3 & Q4	.53	.61	.48	.79

Table 1: Mean solution rates for lower-level system elements (Q1 and Q2) and higher-level system element (Q3 and Q4)

In line with our hypotheses, participants' ability to infer the behavior of the stock was highly influenced by whether the question format induces local or global processing. This was the case for both, the graphical and the numerical presentation. Specifically, in the graphical presentation, participants achieved mean solutions of .51 in the local, compared to .70 in the global format, t(178)

=3.6, $p<.001$ for Q3 and Q4. Similarly, in the numerical presentation, participants achieved mean solutions of .79 in the local, compared to .85 in the global format, $t(84)=4.6$, $p<.001$.

Interestingly, results showed that the questions about the lower-level system elements of flows (Q1 and Q2) were answered more correctly in the local compared to the global presentation, and, again, this was the case for both the graphical, and the numerical presentation. Specifically, in the graphical presentation, participants achieved mean solutions of .76 in the global, compared to .96 in the local format, $t(178)=4.5$, $p<.001$. Similarly, in the numerical presentation, participants achieved mean solutions of .85 in the global, compared to .98 in the local format, $t(81)=4.6$, $p=.009$. Analogously to the above results concerning higher-level system elements, participants were better able to answer the lower-level system elements when the task was presented such that local processing of the system elements was induced.

Participants were thus better able to infer the overall system behavior (Q3 and Q4) when the questions were presented such that the system is processed globally rather than locally, and this beneficial influence of global presentation was stronger in the numerical presentation than the graphical. We found no difference in mean solutions between the graphical (M=.53, SD=.31) and the numerical display (M=.48, SD=.28) in the local question format for Q3 and Q4, $t(84)=.71$, $p=.46$. However, in the global format, a significant difference emerged between both displays, with solutions in the numerical (M=.79, SD=.32) being higher than in the graphical display (M=.61, SD=.42), $t(90)=2.3$, $p=.026$.

Analysis of the Kimchi-Palmer-questions revealed that individual processing styles had no impact on solution rates in any of the four experimental conditions. The higher solution rates are thus due to the global question format and within the global question format, due to the numerical data display and not to individual tendencies towards more global processing of information.

Discussion

At the outset of this paper, we argued that the behavior of higher-level system elements, the stock, emerges through the interrelations between lower-level elements, the stock's in-and outflows. We therefore tested whether people's understanding of the overall behavior of a dynamic system is increased when questions on the system are formatted such that they induce global processing of the system (i.e., processing of interrelations between elements) instead of local processing (i.e., processing of isolated elements). We found that

(a) in line with our expectations, in the global question format, participants achieved higher mean solutions in the question concerning higher-level system elements (i.e., the stock) than in the local question format,

(b) in the local question format, participants achieved higher mean solutions in the question concerning lower-level system elements (i.e., the flows) than in the global question format, and

(c) whereas in the local format no effect of data display (numeric vs. graphical) was found, in the global format, a significant effect of data display emerged, with participants achieving higher solution rates in the numerical compared to the graphical display.

In sum, our results demonstrate that, to achieve system understanding, it is crucial that the way the system is presented is in line with the kind of understanding that the audience needs to acquire: In order for people to understand issues about lower-level system elements, local presentation of the system is crucial. If, however, people need to infer the behavior of the system as a whole, then global presentation of the system is essential.

Our results go beyond previous results in two critical respects. First, they deliver an explanation for the previously identified stock-flow failure in that we show that stock-flow failure is increased, or system understanding is decreased, when questions about the system focus on isolated system elements and thereby induce local processing of the system.

Second, our results deliver an explanation of why scholars in previous research have not found effects of data display on SF-performance, albeit it was argued that different displays should exert an effect on system understanding (Cronin, et al., 2009). Specifically, in the example of the Department Store task, Cronin, et al., (2009) argue that the numerical display should be easier to understand since the “line graph may conflict with participants’ conception of the discrete event of a person entering or leaving a store” (p. 120). While this reasoning is highly plausible, our results suggest that the effect did not occurr because the questions were formulated locally by asking for specific and isolated system-elements. We found that when questions are framed globally, data display did exert an effect. Moreover, the direction of the effect was in line with Cronin et al.’s (2009), expectations: Participants performed better in the numerical (depicting discrete events) than the graphical display.

As the global presentation of the system in our specific SF-task not only led to higher solution rates of the task but also enabled the numerical data display to have a significant impact, the question arises whether something similar could hold for the remaining category of SF-failure, namely the impact of previous knowledge. Future research will test the effectiveness of the global presentation and the numerical data display with the same diversity of participants’ background and in SF-tasks that are structurally identical to the one used in this study but situated in different application domains.

Conclusion

We argue that, due to the hierarchical structure of dynamic systems, cognitive processing of these systems should take place on different cognitive levels as well. In line with this expectation, our results show that a local focus on lower-level system elements is beneficial for answering questions about inflows and outflows, whereas a global focus on higher-level system elements is beneficial for answering questions about the stock. Moreover, when the questions on the system are formatted in such a way that they highlight interrelations between system elements (i.e., global processing), as opposed to highlighting the system elements in isolation, we found an effect of the display format such as it was anticipated previously by Cronin et al. (2009), but was not found empirically in previous research. In sum, our results underline the importance of a match between the hierarchical level that the format of a dynamic system focuses on, and the hierarchical level of the system that problem-solvers and policy-designers need to understand.

The ability to infer the behavior of the system as a whole is relevant in a multitude of systems, with the world's climate system being just one example. Sterman et al., (2012) demonstrate that the general public as well as policy makers have difficulties understanding the dynamics between CO_2 emissions and CO_2 accumulation in the atmosphere, the impact of time delays, accumulations, feedback, and non-linearities. Publicly available information on this topic is published by the Intergovernmental Panel on Climate Change (IPCC) and includes Assessment Reports and the Summary Report for Policy Makers. Even the latter (IPCC, 2014), that addresses the wide public untrained in climatology and physics, provides a multitude of detailed information about the various aspects of climate change. It includes many specifics about scientific concepts and their underlying mathematical details, including confusing units of measure like '$GtCO_2$-eq' and 'Wm^{-2}', for example. It also details likely human and possible natural causes of climate change, their impacts on global mean temperature and sea level rise and possible future warming scenarios. The Summary Report also lays out possibilities for adaptation and mitigation. Even though it illustrates how some of the drivers of climate change interact with each other, the overall picture is missing. The report focuses on great detail, without explaining the high-level insights with the highest impacts and their interrelations.

The implications of our findings in the context of the world's climate system are, for example, that the effectiveness of communication of results by the IPCC could be improved. One way for doing so is the approach that ClimateInteractive (cf. http://www.climateinteractive.org/) and the MIT LearningEdge (cf. https://mitsloan.mit.edu/LearningEdge/simulations/Pages/Overview.aspx) use to simplify communication of climate change with the C-ROADS (Climate Rapid Overview And Decision Support) model, a system-dynamics-based management flight simulator. One of the main purposes of system-dynamics-based management flight simulators (Sterman, 1989, 2014a, 2014b) is to provide users with the possibility to explore the behavior of a system as a whole, as it results from the interplay of lower-level elements. C-ROADS (Sterman et al., 2013) helps users understand the relationship between greenhouse gas emissions, atmospheric CO_2, global mean temperature, sea level rise, and ocean ph-level, for example (http://www.climateinteractive.org/tools/). When there is interest, users may travel along the cone shown in Figure 1 deeper into details of mathematical equations of the simulator, the lower-level system elements. Yet, on a higher level of abstraction, C-ROADS sheds light on the fundamental insights of climate change. Based on the findings from our study, it is important that management flight simulators are designed in such a way that they facilitate global rather than local processing of the information provided in the user interface. The same holds for briefing and debriefing of these management flight simulators.

References

Booth Sweeney, L., & Sterman, J. D. (2000). Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review, 16*(4), 249–286.

Brunstein, A., Gonzalez, C., & Kanter, S. (2010). Effects of domain experience in the stock–flow failure. *System Dynamics Review, 26*(4), 347–354.

Burgoon, E. M., Henderson, M. D., & Markman, A. B. (2013). There Are Many Ways to See the Forest for the Trees: A Tour Guide for Abstraction. *Perspectives on Psychological Science, 8*(5), 501-520. doi: 10.1177/1745691613497964

Coyle, R. G. (1996). The Practice of System Dynamics: Milestones, Lessons and Ideas from 30 Years Experience. *System Dynamics Review, 14*(4), 343-365.

Cronin, M. A., & Gonzalez, C. (2007). Understanding the building blocks of dynamic systems. *System Dynamics Review, 23*(1), 1–17. doi: 10.1002/sdr.356

Cronin, M. A., Gonzalez, C., & Sterman, J. D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organizational Behavior and Human Decision Processes, 108*(1), 116–130.

Fiddaman, T. S. (1997). *Feedback Complexity in Integrated Climate-Economy Models.* Thesis (Ph.D.), Massachusetts Institute of Technology, Cambridge, Mass.

Fischer, H., & Gonzalez, C. (2013). *Seeing the forest for the trees predicts accumulation decisions*. Paper presented at the Paper presented at the The Annual Meeting of the Cognitive Science Society, Berlin, Germany.

Fischer, H., & Gonzalez, C. (2015). Making sense of dynamic systems: How our understanding of stocks and flows depends on a global perspective. *Cognitive Science,* 1-17.

Gonzalez, C., & Wong, H.-y. (2012). Understanding stocks and flows through analogy. *System Dynamics Review, 28*(1), 3–27. doi: 10.1002/sdr.470

Hämäläinen, R. P., Luoma, J., & Saarinen, E (2013). On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. *European Journal of Operations Research*, *228*(3), 623-634.

IPCC. (2014). Climate Change 2014. Synthesis Report. Geneva: WMO, UNEP.

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 in Palermo, Italy.*

Kapmeier, F. (2004). Findings from four years of bathtub dynamics at higher management education institutions in Stuttgart. *Proceedings of the 22nd International Confernce of the System Dynamics Society*.

Kapmeier, F., Happach, R. M., & Tilebein, M. (2014). *Bathtub Dynamics Revisited: Does Educational Backgroud Matter?* Paper presented at the System Dynamics Conference, Delft, NL.

Kapmeier, F., & Zahn, E. O. K. (2001). Bathtub Dynamics: Results of a Systems Thinking Inventory at the Universität Stuttgart, Germany. Retrieved from http://www.esb-business-school.de/business-school/organisation/professoren-und-dozenten/kapmeier.html

Kimchi, R., & Palmer, S. E. (1982). Form and texture in hierarchically constructed patterns. *Journal of Experimental Psychology: Human Perception and Performance, 8*(4), 521-535.

Moxnes, E., & Jensen, L. (2009). Drunker than intended: Misperception and information treatments. *Drug and Alcohol Dependence, 105*(1-2), 63–70.

Ossimitz, G. (2002). Stock-flow-thinking and reading stock-flow-related graphs: an empirical investigation in dynamics thinking abilities. *Proceedings of the 20th International Conference of the System Dynamics Society in Palermo, Italy.*

Schwarz, M. A., Epperlein, S., Brockhaus, F., & Sedlmeier, P. (2013). Effects of illustrations, specific contexts, and instructions: Further attempts to improve stock-flow task performance. *Proceedings of the 31st International Confernce of the System Dynamics Society in Cambridge, MA, USA*.

Sedlmeier, P., Brockhaus, F., & Schwarz, M. (2014). Visual integration with stock-flow models: How far can intuition carry us? In T. Wassong, D. Frischemeier, P. R. Fischer, R. Hochmuth & P. Bender (Eds.), *Mit Werkzeugen Mathematik und Stochastik lernen – Using Tools for Learning Mathematics and Statistics* (pp. 57–70). Wiesbaden: Springer Fachmedien.

Sterman, J. D. (1989). MODELING MANAGERIAL BEHAVIOR: MISPERCEPTIONS OF FEEDBACK IN A DYNAMIC DECISION MAKING EXPERIMENT. [Article]. *Management Science, 35*(3), 321-339.

Sterman, J. D. (2000). *Business Dynamics: System Thinking and Modeling for a Complex World*. Boston, Madison u.a.: Irwin Mc Graw-Hill.

Sterman, J. D. (2002). All models are worng: reflections on becoming a systems scientist. *System Dynamics Review, 18*(4), 501–531.

Sterman, J. D. (2009). The Carbon Bathtub. *National Geographic*(12).

Sterman, J. D. (2014a). Interactive web-based simulations for strategy and sustainability: The MIT Sloan LearningEdge management flight simulators, Part I. [Other]. *System Dynamics Review, 30*, 89-121. doi: 10.1002/sdr.1513

Sterman, J. D. (2014b). Interactive web-based simulations for strategy and sustainability: The MIT Sloan LearningEdge management flight simulators, Part II. [Other]. *System Dynamics Review, 30*, 206-231. doi: 10.1002/sdr.1519

Sterman, J. D., & Booth Sweeney, L. (2002). Cloudy skies: assessing public understanding of global warming. *System Dynamics Review, 18*(2), 207–240. doi: 10.1002/sdr.242

Sterman, J. D., Fiddaman, T., Franck, T., Jones, A., McCauley, S., Rice, P., . . . Siegel, L. (2012). Climate interactive: the C-ROADS climate policy model. [Article]. *System Dynamics Review, 28*(3), 295-305. doi: 10.1002/sdr.1474

Sterman, J. D., Fiddaman, T., Franck, T., Jones, A., McCauley, S., Rice, P., . . . Siegel, L. (2013). Management flight simulators to support climate negotiations. *Environmental Modelling & Software, 44*(0), 122-135. doi: http://dx.doi.org/10.1016/j.envsoft.2012.06.004

Veldhuis, G. A., & Korzilius, H. (2012). Seeing with the mind - The role of spatial ability in inferring dynamics behaviour from graphs and stock and flow diagrams. *Proceedings of the 30th International Confernce of the System Dynamics Society in St. Gallen Switzerland*.

Wilensky, U., & Resnick, M. (1999). Thinking in Levels: A Dynamic Systems Approach to Making Sense of the World. *Journal of Science Education and Technology, 8*(1), 3-19. doi: 10.1023/A:1009421303064