

Effects of an Educational Scenario Exercise on Participants' Competencies of Systemic Thinking

Simon Burandt

Institute for Environmental and Sustainability Communication,
Leuphana University Lüneburg, POB 2440, D-21314 Lüneburg, Germany

Abstract: Problem statement: Education for Sustainable Development (ESD) aims to shape key competencies of individuals and therefore needs methods to enable learners to acquire these competencies. Systemic thinking can be regarded as a meta-competency in ESD, because it contains many important aspects found in most key competencies of ESD. Scenario analysis is described as a learning environment that fosters the acquisition of systemic thinking and other important competencies, but empirical proof of this assumption is rarely found in the literature. This article presents such an empirical study and develops a specific instrument to investigate the effects on participants' competencies taking part in ESD seminars in which scenario analysis was used as methodology. **Approach:** A study was conducted of four educational seminars, using a pre/post design with two treatment groups which took part in a scenario analysis seminar and two control groups. Altogether 72 university students from different disciplines (semesters 1-6) were involved. In a questionnaire, constructs like domain specific knowledge and the perception of the future as well as systemic thinking were operationalized quantitatively in order to achieve a practicable and quick measurement. Similarity Judgement Rating was used to elicit participants' knowledge structures about climate change in order to gain concept-maps for comparison with reference models. Paired t-test, mean values, factor and cluster analysis and correlation were used. **Results:** No significant changes in the structural knowledge, perception and formal knowledge of the groups could be observed although some developments were noted, e.g. in the perception of the future. The fact that treatment groups showed little advancement in factual knowledge could be read as a hint that generally only weak (measurable) effects had taken place. **Conclusion:** Some indications were found that the measurement instrument works in principle, but that its application in the thematic domain of climate change seems to be problematic due to a relatively high level of general knowledge and systemic interrelations of concepts (of climate change) that cannot be precisely described. The hypothesis of educational effects of participation was not confuted but should be investigated in a different thematic context.

Key words: Systemic thinking, measurement of competencies, Scenario Analysis (SA), HESD, similarity judgment, pathfinder network, Education for Sustainable Development (ESD)

INTRODUCTION

In the literature educational effects are described on participants in an (educational) scenario analysis (SA), like competence building through reflection of feedback-loops and interrelations, stimulation of imaginative and explorative thinking in alternatives and being able to cope with complexity and uncertainties (Godet, 2000; van Notten *et al.*, 2003; Swart *et al.*, 2004). Therefore the use of scenario analysis is discussed as a teaching method in environmental education and Education for Sustainable Development (ESD).

So far there have only been a few attempts to test the educational effects attributed to scenario analysis. This article describes the development of a method and

the results of its application in order to investigate the hypothesis that participation in an educational scenario analysis leads to more sophisticated systemic thinking and also affects participants' perception of the future and domain specific knowledge.

For this purpose, a learning setting (higher education for sustainable development) was developed in the form of a project seminar that used a basically qualitative and explorative scenario exercise to arrange a didactic setting. All goals of the seminar were derived from the relevant literature. The seminar focused among other driving forces on the development of regional future scenarios with a special emphasis on climate change. It aimed at answering the question of "what could the environmental conditions for tourism look like in the Black Forest in 2050 when

climate change is taken into account” (Burandt and Barth, 2010).

The hypothesis mentioned above was studied especially by means of participants’ cross-linked knowledge structure of the domain of climate change. Climate change was chosen because it was part of the seminar’s framework of thought and also is a very complex topic demanding well linked knowledge structures to understand it. Therefore a new method for the measurement of the competency of systemic thinking was developed. One requirement on this method was for it to enable a quantitative collection of data that uses as little of the participants’ time as possible.

The article first gives an overview of the background of systemic thinking and integrates it into the discourse about competencies for education for sustainable development. This is followed by an introduction to different methods of measuring systemic thinking before the theoretical framework for its measurement in this study is deduced by using Similarity Judgment Ratings. The second part of the article describes the actual study, the research instrument, results and conclusions of the study.

Systems, cross-linked thinking and structural knowledge:

What is systemic thinking?: The concept of “systems thinking” or “systemic thinking” has evolved over the last 50 years and is now used in many disciplines; this is why it has acquired diverse meanings, ranging from a set of skills to a proper discipline. From an educational perspective, all approaches agree more or less on the existence of thinking skills that help people to better understand interdependencies and processes in systems e.g. in order to achieve improved decision-making or to foresee the outcome of an action. The different approaches deliver tools and methods to cope better with complex situations or to facilitate the acquirement of useful thinking skills.

The quantitatively oriented branch of systems thinking emerged from Forrester’s “industrial dynamics” and later “system dynamics” (Forrester, 1987). These concepts follow a strict quantitative paradigm and are often linked with the use of simulation software. Later, Richmond introduced the concept of “systems thinking” and enabled a broader use of simulation software by using flow charts and a graphic user interface instead of a simulation language only. Richmond describes systems thinking as a set of skills indispensable for the (competent) use of simulation software, including: dynamic thinking, closed-loop thinking, generic thinking, structural

thinking, operational thinking, continuum thinking and scientific thinking (Richmond, 1993).

The qualitative use of the concept was, in the German speaking countries, introduced and established by Vester (leitmotif of cross linked thinking) (Vester, 1989). Thus, system oriented management approaches refer to Vester’s concept and have developed a methodology to “model” and analyze systems without a computer, e.g., with flowcharts (Gomez and Probst, 1995). It is possible to identify a system’s core drivers, feedback loops and certain aspects of system dynamics in order to deduce possibilities of managing the system that is being looked at. Senge’s “systems thinking”, though originating in the quantitative branch, has been developed into the idea of (qualitative) organizational learning (Ossimitz, 2000). Senge describes systems thinking as the most important “fifth discipline” in organizational learning that integrates four other disciplines. In its quintessence, systems thinking is viewed as a skill which enables one to see feedback and processes of change instead of snapshots; and to see “interrelationships rather than linear cause effect chains” (Senge, 1993).

In cognitive psychology Dörner introduced the approach of complex problem solving, which refers to Vester’s idea of cross linked thinking (Dörner and Wearing, 1995). He initiated a series of experiments from which evolved a whole new field of research. These experiments were intended to measure how people perform in complex problem situations (e.g. in a complex computer simulation with variables and a lot of feedback loops: “Tanaland”) where systems thinking skills had to be used. Research was also done on whether special (educational) trainings can improve performance and skills. As some conclusions from large studies have shown, “correct” behavior is dependent on the situation. This demonstrates that individual systems thinking is context dependent and cross-references to different contexts are difficult to establish. Finally, it was concluded that systemic thinking is a bundle of abilities that cannot be described and investigated like a single ability (Dörner and Wearing, 1995).

It is obvious that a common definition or understanding of systems thinking is difficult to obtain, as is also stated by Sweeney and Sterman (Sweeney and Sterman, 2000): “There are as many lists of systems thinking skills as there are schools of systems thinking”. Ossimitz (Ossimitz, 2000; 1997) gives a general definition of systemic thinking that attempts to integrate different system approaches from the literature. Systemic thinking embraces four interrelated dimensions:

- Interrelated thinking: A thinking in interrelated, systemic structures
- Thinking in models: Explicitly comprehended modeling
- Dynamic thinking: A thinking in dynamic processes (delays, feedback loops, oscillations).
- Steering systems: The ability for practical system management and system control

This is a very comprehensive definition that does not focus on a special “systemic thinking school”. However, there is considerable evidence that thinking and knowing does not necessarily lead to appropriate action (Kollmuss and Agyeman, 2002). Furthermore, thinking in models does explicitly integrate a constructivist perspective, but is also understood as a skill at using software tools to represent (mental) models. As “dynamic thinking in cross-linked structures is always thinking in models” (Seel, 1991), people still could fail in representing models in an unfamiliar way (Sweeney and Sterman, 2000). So, Ossimitz definition (especially from the viewpoint of ESD) seems too limiting. In this article systemic thinking is understood accordingly to comprise only the first three dimensions: The ability to recognize, describe and model complex parts of reality as systems; the ability to identify drivers, elements and their interrelations and, third, the consideration of dynamic processes (dimension of time) for the further development of (mental) models (Although using different arguments, Rieß and Mischo (2008) arrived at a similar understanding of systemic thinking).

Systemic thinking - a competency in education (for SD)?: Internationally, the OECD project “Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)” established a conceptual framework that embraces “Key Competencies for a Successful Life and a Well-Functioning Society”. Independently from the ESD discourse, a normative framework was developed consisting of three categories into which certain key competencies can be classified: (1) interacting effectively in socially heterogeneous groups, (2) using tools interactively and (3) acting autonomously (Rychen, 2009).

Education for sustainable development embraces attempts to educate, enable and empower people to contribute and participate actively in the sustainable development of our society (de Haan 2006). Therefore the German discourse led to the central goal for ESD “to offer possibilities to acquire shaping competence” (“Gestaltungskompetenz”) (de Haan 2006).

Shaping competence covers a set of corresponding key competencies. There is substantial agreement about these key competencies even if the exact definitions and distinctions are still under discussion. Barth classified key competencies of shaping competence into the DeSeCo framework, offered a theoretical background and showed that they are internationally linkable (Barth, 2009).

In neither discourse can “systemic thinking” be found explicitly among the discussed key competencies. But they consist in any case of a set of subsidiary competencies that in their interplay account for the full competence. As an example, the importance of systemic thinking on this subsidiary level is illustrated in the following.

The competency to plan processes and sequences of action refers directly to individual action but also implies a certain degree of systemic understanding, or the ability to identify steps and to correlate them. The importance of systemic thinking becomes more obvious in both the competency to think anticipatorily, to cope with uncertainties and to develop prognoses and the competency for dealing with uncertainties and thinking proactively. Coping with complex systems requires the ability to identify elements as part of a system as well as to adopt a holistic view of interrelations and dynamics (Burandt and Barth, 2010). The necessity to think proactively generally involves anticipating (unintentional) effects and accounting for the possibilities of risk, both of which necessitate a high degree of systemic and dynamic thinking. For the competency to collaborate interdisciplinarily (and transdisciplinarily) and the competency for using, shaping, handling and sharing different sets of information and knowledge, it is necessary to have a systemic understanding of the specific knowledge of e.g. one’s own discipline and to be able to transfer it to new contexts. Systemic knowledge in the form of knowledge about structures, processes or interrelations also has to be linked to new contexts or problem areas and it has to be applied and communicated. In addition, the DeSeCo project stresses the general importance of systemic and cross-linked thinking because it contributes to (self) reflection, which is a core aspect of most key competencies (Rychen, 2009).

In summary, systemic thinking is an essential part of most key competencies. Therefore it is understood in the sense of a meta-competency that is part of, or facilitates the use of, specific competencies (Weinert, 2004).

Measuring systemic thinking: Like measuring the success of (educational) interventions, systemic

thinking has always been a focus of interest in the field of educational assessment. Today, this whole area is linked to the different discourses about competencies. The assessment and measurement of competencies is a very complex topic needing “sound models of competence structures, competence levels and competence development” (Klieme *et al.*, 2008).

Pioneers from cognition psychology in the assessment of systems thinking in educational contexts are especially because of the Hilu-scenario task. Among other things, pupils had to represent a text in the form of a chart that describes the complex life of the Hilu-tribe and also had to answer questions about future states/scenarios (of this dynamic system). Over the years the survey methods became more complex and their scale increased. Later Niedderer *et al.* (1991) used a very sophisticated method consisting of eight different types of task and Ossimitz operationalized his definition of systemic thinking by assigning seven sub aspects (skills) to the 4 dimensions, e.g. the identification of cross-linkings whereby these aspects contribute to some extent to all dimensions. Among other tasks he used a refined version of the Hilu-scenario in his survey instrument that he assessed as suitable (Ossimitz, 2000).

Sweeney and Sterman (2000) summarize a great amount of (international) research which deals with the efficacy of interventions designed to develop systemic thinking. Their conclusion is that most research questions have so far remained unanswered. They identify (like Ossimitz) a couple of specific “systems thinking skills” and introduce a new type of tests to explore students’ baseline systems thinking abilities. One of these is the “bathtub task”: here a relatively simple task about in- and outflowing water in a bathtub without any feedback loops had to be visualized in a graph (diagram). This type of test captures systems thinking through some technical, quantitatively oriented skills and was repeated in a lot of similar studies in other countries. Sweeney and Sterman found that there is only a weak relationship between education and performance (even in the case of MIT students who are likely to be very familiar with higher mathematics). They assumed that there might be a difference between (everyday) understanding of systems and the presentation of the problem in the form of a graph (Sweeney and Sterman, 2000).

Although these types of tests have delivered viable results, the capturing of systemic thinking in the form of representing a linear quantitative development (water in a bathtub) without (complex) interrelations seems to be too narrowly focused to map it in the sense of education for sustainable development. As introduced

before, the meta-competency systems thinking requires a more general understanding. Sustainability implies complex, ill-structured, real-world problems where relations of system elements can often not be exactly quantified. A tool would be desirable that captures quantitatively all qualitative dimensions of systems thinking, but this would lead to very complex survey methods that would be difficult to apply in practice-Niedderer *et al.* (1991) e.g. needed about 8½ and Ossimitz (2000) up to 5 hours for their respective surveys. Rieß and Mischo (2008) have developed and validated a questionnaire on the comprehension of systemic thinking in a context relevant to sustainability. They do not provide information about the time needed to answer all questions, but used open questions as did Sweeney (Sweeney and Sterman, 2000) too. These tasks generally require an elaborate procedure for the analysis of the data because e.g. charts or diagrams have to be interpreted and translated into codes by human raters.

So, a wide range of different methods has been evolved over the last decades, each with their own advantages and disadvantages. None of them is able to “measure” the construct of systems thinking independently from context so that results cannot be generalized. Finally, the right balance has to be found for each study between practicability and complexity regarding the coverage of the theoretical construct.

Measuring systemic thinking in ESD: Following Ossimitz (2000) some sub-aspects can be assigned to the four different dimensions of systemic thinking – in the sense of a meta-competency of key competencies for sustainable development. To keep the survey method as lean as possible, the construct of systems thinking will be represented here mainly by one aspect: the identification of connections and relations of drivers in a complex system of a given domain, because cross-linked thinking is an important aspect that is closely interwoven with all dimensions of systemic thinking (Ossimitz, 2000).

Cross-linked thinking demands both a certain degree of knowledge of the system one is looking at and a specific use (transfer) of this knowledge. Cognitive psychology gives different constructions of knowledge that can be useful for cross-linked thinking. Beside declarative (Ryle and White, 1972), procedural (Schank and Abelson, 1979) and tacit knowledge (Polanyi, 1997), an intermediate type of knowledge appears most useful “that mediates the translation of declarative into procedural knowledge and facilitates the application of procedural knowledge. Structural knowledge is the knowledge of how concepts within a

domain are interrelated” (Diekhoff, 1983). According to this the following assumption can therefore be made: systemic thinking as a meta-competency for ESD can be represented largely by structural knowledge because qualitative relationships among concepts play a particularly important role in contexts relevant to sustainability.

Structural knowledge is a widely accepted construct of cognitive structure. It is based on the theory of semantic networks (Collins and Quillian, 1969), the most important feature of which is that human memory is organized semantically. Memory structures are composed of nodes and ordered relationships or links connecting them. Three important aspects build up the rationale for structural knowledge (Jonassen *et al.*, 1993):

- Structure is inherent in all knowledge (Mandler, 2004)
- Learners assimilate structural knowledge (Shavelson, 1972)
- Experts’ structural knowledge differs from that of novices (Chi *et al.*, 1981)

Generally, when one works with structural knowledge, two assumptions have to be made. The first is the concept of semantic similarity: it refers to the spreading activation theory (Collins and Loftus, 1975) which basically says that the more closely two concepts are linked, or the more common properties two concepts have, the more similar they are processed in semantic networks. The second assumption that has very often to be made is that semantic similarity in the form of semantic space of concepts in memory can be represented in terms of geometric space.

The literature offers several ways to analyze structural knowledge. Generally it is necessary to (1) elicit the structural knowledge (directly or indirectly) in

order to (2) represent and analyze the underlying structure (Jonassen *et al.*, 1993). The elicitation of the structure of knowledge can be done by similarity ratings or similarity judgment tests (SJTs). For this, after a set of related concepts that define a subject’s domain have been identified, the respondent is asked to assess the degree of relationship/proximity of each pair of concepts (e.g., on a Likert scale). It is the “most direct method for rating or comparing the semantic similarity between concepts in an individual’s cognitive structure” (Jonassen *et al.*, 1993). The results of the rating can be transformed into a (proximity) matrix that delivers the data basis for the representation of the knowledge structure. This is an indirect method to elicit the knowledge structure (Stanners *et al.*, 1983; Goldsmith *et al.*, 1991).

In a next step, the underlying knowledge structure of the matrix can be represented and analyzed. It is for example possible to transform the matrix into cognitive maps (Schvaneveldt *et al.*, 1989; Jonassen *et al.*, 1993; Shavelson *et al.*, 2005). The software package “Pathfinder KNOT” delivers several possibilities of creating cognitive maps with nodes and links, e.g. in the form of data nets. The Pathfinder algorithm is in addition also able to extract the system’s latent structure by identifying the closest connections of concepts. Cognitive maps can be analyzed qualitatively or quantitatively. To process cognitive maps automatically it is possible either to correlate their rough data or to compare maps with a reference system. KNOT supports the comparison of networks, e.g. by the index of correspondence “csim” that basically puts into relation the number of interrelations with the number of shared interrelations (corrected by probability). Figure 1 shows a data matrix (of concept pairs) and the resulting Pathfinder network.

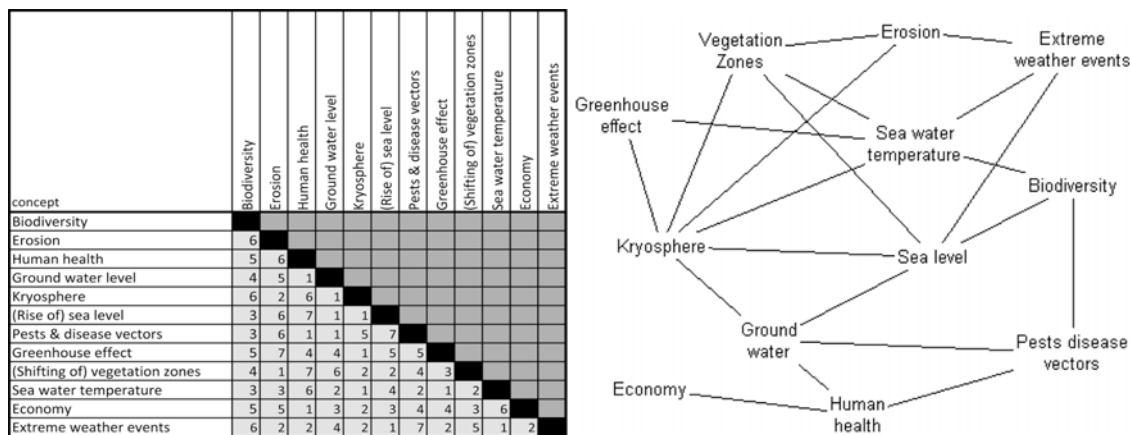


Fig. 1: Similarity matrix and pathfinder network ($\infty, n1$) of a first semester student @t1

Therefore it is suggested that a comparison of students' networks with that of an expert at different stages of an educational setting can show changes in individual knowledge structures and indicate an advancement in systemic thinking.

MATERIALS AND METHODS

The primary purpose of this study was to test the hypothesis that participation in an educational scenario analysis has an effect on the competency of systemic thinking and, secondly, an impact on participants' knowledge and their perception of the future.

These hypotheses were tested in a pre/post design with four university seminars: Two Treatment Groups (TG) and two Control Groups (CG). Table 1 gives an overview of the samples. The CGs participated in a "normal" seminar while the TGs took part in the educational scenario exercise. The study was conducted in the winter semesters of 2008/2009 and 2009/2010; course length was 14 weeks.

In the first sessions students had to fill in a questionnaire (t1) and at the end of the seminars (t2) the same questionnaire had to be filled in again in an online version. It took students about 35-45 minutes each to answer all questions. As the data in the questionnaires were collected anonymously, participants were asked to create their own personal code so that t1 and t2 could be compared individually. At t1, each seminar was attended by about 30 students from various disciplines, such as environmental sciences, cultural sciences, law, economics and education in their (obligatory) complementary part of their disciplinary studies at [name of university] (semesters 1-5). Due to drop-outs and to wrong code entries the number of data pairs was reduced so that not all respondents could be taken into the analysis (see N paired datasets).

The questionnaire consisted of two parts: one was to measure systemic thinking and the other was to investigate changes in constructs of perception and knowledge. As the TGs' class work concerned climate change, that was also the topic which was chosen as the questionnaire's thematic domain.

To elicit structural knowledge a SJT about climate change was developed. Table 2 shows 12 concepts from the domain of "climate change" which were derived covering ecological, societal and economic aspects. These concepts were paired in all possible combinations (12 x (12-1) / 2) and ordered randomly in the questionnaire to avoid context dependent effects. The 66 concept pairs had to be rated for their proximity (or semantic similarity, which could include causal relationships) on a scale from one (very close) to seven (no relation).

For the analysis of students' responses the list of pairs was converted into a similarity (or distance) matrix and transformed by the Pathfinder algorithm into concept maps in order to represent the underlying knowledge structure. To analyze the quality of and differences in individual systemic thinking both the structural knowledge networks and data matrices were compared to reference systems by two indices c_{sim} and R_{STJ} (correlation of data matrix) calculated with KNOT. The main reference system was created as the median of the responses of three national and international experts on climate change. As additional references were used the group average and the TG's-teacher's concept map.

Attitude to climate change, perception of climate change and perception of the future were tested by different item batteries (Likert scale 1-7). Constructs were validated by factor analysis and tested for reliability (Cronbachs Alpha range 0.56-0.74 for t1 and t2).

Table 1: Composition of investigated groups

Group	CG01	CG02	TG01	TG02
Seminar (title)	"ESD: challenges and chances"	"ESD in different contexts"	"Coping with a complex future	by using scenario analysis"
Setting	Introductory meetings, 1x one-day session, presentation and evaluation meetings	14 weekly sessions	14 weekly sessions	Initial meeting, preparation phase, 2x 2-day weekend sessions, collaboration phase in-between
Didactic	self directed, collaborative, practice oriented	Self directed, collaborative, problem oriented	Self directed, collaborative, problem oriented	
Methods	presentations, group work, discussions, mutual evaluation	Self organized research, pronounced participatory approach	Scenario analysis (as framework)	
Workload	150 h	150 h	150 h	150 h
Type	Project seminar		Project seminar	
Participants (t1)	30	32	30	27
N (paired datasets)	15	21	19	17
Female	8	20	9	12
Male	7	1	10	5
Mean age	23.6	21.4	22	23.1

Table 2: concepts of climate change

Biodiversity	Pests & disease vectors
Erosion	Greenhouse effect / average temperature
Human health	(Shifting of) vegetation zones
Ground water level	(Rise of) sea water temperature
Kryosphere	Economy
(Rise of) sea level	Extreme weather events

Students were asked to self-assess their knowledge about climate change by assigning school grades to it. Knowledge also was tested by two item batteries:

- Assessment of importance of 18 given “drivers” for climate change (Likert scale 1-4)
- “Facts’ about climate change” consisted of a list of 15 true/false statements, arranged in order of increasing difficulty, that also contained some of the public “myths” about climate change in order to investigate whether their knowledge would advance to the level of expert knowledge. A third option (“don’t know”) was given to minimize guesses

The data were analyzed by descriptive and multivariate statistics (percentage, mean, paired t-tests). Factor and cluster analysis were also applied to the SJT data in order to gain additional access to data as well as to identify different types of responding structures/areas of improvement within the 12 concepts of climate change and latent responding structures among participants. Other aspects that were calculated included bivariate correlation (Pearson and Spearman) of knowledge, different constructs and the similarity of structural knowledge to changes in systemic thinking (csim and R_{SJT}) and demographic aspects.

RESULTS

Changes occurred in all four groups during the period of the seminars in both the latent (csim) and the direct structure (R_{SJT}) of participants’ knowledge. These changes in relation to the different reference systems are shown in Table 3. Taking the structural knowledge of the “3-experts average” model as reference point, both TGs’ mean values of csim and R_{SJT} are relatively high at the beginning (t1), but csim for example decreases at t2 (0.021 and 0.015), while CG01 showed an increase of -0.026 and CG02 a *loss* of 0.043 in correspondence with the reference model. In the TGs especially no common effect could be observed. Tests with additional reference models, like the teacher’s one, delivered similar results.

A significant measurable effect was that participants in CG02 veered away from a shared

knowledge structure of climate change (“Group Average” csim), whereas in both TGs the participants approximated their structural knowledge minimally (<0).

The data of the similarity matrices were investigated to identify structures and parts in which significant differences between the groups could be observed. No usable factors were yielded by a factor analysis of the 66 concept-pair variables of the difference t1-t2 (showing areas of strong and slight change) and of the absolute values |t1-reference model|-|t2-reference model| (showing areas of improvement in relation to the reference model). Concepts of major change were identified: within the top one third of concept-pairs with the highest number of changes (t1-t2), the dominant concepts were groundwater (7x), diseases (6x), erosion (6x) and kryosphere (5x). But no connections to different groups could be established.

A cluster analysis of participants’ responding structures showed that those students who showed an above-average improvement in their knowledge structure during the course also evinced an above average correspondence with the reference model at t1. Additionally, there were no usable results of clusters of participants (by analyzing the strength of changes in the structural knowledge between t1 and t2). An analysis of only those areas which showed major changes (see above) did not deliver any clusters or groups that had describable differences.

A paired t-test of perception, attitude and knowledge did not deliver significant results either. Therefore only descriptive data about attitude and perception (Table 4) and relevant knowledge (Table 5) have been included in this article. These results do not indicate differences between the TGs and the CGs during t1 and t2. TG01 achieved a higher score for seeing/accepting the unpredictability of the future after the seminar. Both TGs reached a higher score in their attitude to act proactively in face of an uncertain future, but they did also deliver higher values than the CGs at t1 (Table 4).

Concerning knowledge, all groups tended generally to consider all given impacts as relevant and to have difficulties in distinguishing between important and unimportant impacts; only about 56 % of the maximum score of “wrong/not important impact factors” (like waste separation) was reached (Table 5); no major differences between t1 and t2 were noted. In “Facts’ about climate change”, TG02 reached a higher percentage (3.9%) while TG01 lost and CG02 won a bit. CG01 did lose about 6.5% of correct answers.

Table 3: descriptive data and results of paired t-test of pre/post comparison of participants' cognitive maps with different reference systems

Group	Reference model & index	t1, pre			t2, post		t1-t2		T	df	p (2 way)
		N	M	SD	M	SD	M	SD			
CG01	"3 Experts" csim ^a	15	0.0991	0.1024	0.1259	0.0964	-0.0268	0.1194	-0.8690	14	0.3995
CG02	"3 Experts" csim	21	0.1257	0.0826	0.0828	0.0801	0.0429	0.0796	2.4713	20	0.0220
TG01	"3 Experts" csim	19	0.1451	0.0730	0.1302	0.0977	0.0149	0.1037	0.6278	18	0.5380
TG02	"3 Experts" csim	17	0.1296	0.0894	0.1082	0.0720	0.0213	0.0745	1.1805	16	0.2551
CG01	"3 Experts" R _{SJT} ^b	15	0.2562	0.1593	0.2665	0.1472	-0.0103	0.1614	-0.2467	14	0.8087
CG02	"3 Experts" R _{SJT}	21	0.2472	0.1432	0.2178	0.1256	0.0294	0.1081	1,4660	20	0.2269
TG01	"3 Experts" R _{SJT}	19	0.2594	0.1184	0.2654	0.1216	-0.0060	0.1495	-0.1756	18	0.8626
TG02	"3 Experts" R _{SJT}	17	0.2844	0.1500	0.2449	0.1711	0.0395	0.0891	1.8273	16	0.0864
CG01	"Group (CG01) Average" csim ^a	15	0.1607	0.1001	0.1363	0.0791	0.0245	0.1102	0.8600	14	0.4043
CG02	"Group (CG02) Average" csim	21	0.2177	0.0855	0.0769	0.0706	0.1408	0.0979	6.5940	20	0.0000***
TG01	"Group (TG01) Average" csim	19	0.1828	0.0674	0.1921	0.1038	-0.0093	0.1325	-0.3066	18	0.7620
TG02	"Group (TG02) Average" csim	17	0.1366	0.0764	0.1377	0.0983	-0.0011	0.1220	-0.0375	16	0.9705

***: Significant at <0.001 level; a: index of correspondence corrected by probability – csim (range 0 to 1, similarity of cognitive-maps with reference system); b: correlation coefficient R_{SJT} (accordance of data matrices with reference system); M = mean, SD= standard deviation

Table 4: Descriptive data of participants' attitude and perception of the future

Group		t1, pre			t2, post	
		N	M	SD	M	SD
CG01	unpredictability of future	15	6.60	2.06	6.80	2.46
CG02		21	7.48	2.40	7.14	1.96
TG01		19	6.47	2.22	7.21	2.18
TG02		17	7.00	2.32	6.94	1.30
CG01	personal pro-activity	15	12.87	2.23	12.67	3.09
CG02		21	12.33	3.04	11.86	1.93
TG01		19	13.37	1.80	13.68	2.19
TG02		17	13.35	2.06	13.88	1.97
CG01	understanding of sustainable future approaches (lower values = social and ecological aspects gain more importance)	15	11.13	2.56	10.27	2.40
CG02		21	11.43	2.23	11.52	3.04
TG01		19	11.79	2.42	11.89	2.58
TG02		17	11.29	1.57	10.65	2.09
CG01	Attitude: Chance to contain (or stop) climate change (lower value = more possible)	15	10.67	2.58	10.80	2.70
CG02		21	11.38	2.58	11.33	2.1
TG01		19	11.16	1.95	11.32	1.73
TG02		17	11.29	2.49	11.35	2.64
CG01	Attitude: human impact on climate change (lower value= humans have more impact)	15	7.20	2.76	6.73	2.43
CG02		21	7.10	2.72	7.38	2.48
TG01		19	8.26	3.23	8.32	3.11
TG02		17	6.06	2.11	6.18	2.21

Table 5: descriptive data of participants' knowledge

group		t1, pre			t2, post	
		N	% max ^a	SD	% max	SD
CG01	true impact on climate change	12	82.2	0.065	81.6	0.070
CG02	true impact on climate change	21	81.9	0.084	80.5	0.077
TG01	true impact on climate change	19	74.6	0.103	76.7	0.118
TG02	true impact on climate change	17	86.9	0.107	85.0	0.114
CG01	false impacts on climate change	12	57.7	0.136	55.1	0.118
CG02	false impacts on climate change	21	50.7	0.110	54.6	0.111
TG01	false impacts on climate change	19	59.6	0.096	58.5	0.106
TG02	false impacts on climate change	17	52.1	0.114	56.3	0.153
CG01	impacts on climate change (all)	12	72.7		71.3	
CG02	impacts on climate change (all)	21	69.8		70.4	
TG01	impacts on climate change (all)	19	68.8		69.6	
TG02	impacts on climate change (all)	17	73.4		73.9	
CG01	"facts" about climate change	11	63.5	0.159	56.9	0.148
CG02	"facts" about climate change	20	52.8	0.211	53.6	0.184
TG01	"facts" about climate change	18	57.3	0.123	56.9	0.089
TG02	"facts" about climate change	17	58.2	0.117	62.1	0.116

a: (mean) percentage of maximum score; "facts": (mean) percentage of correct answers

Students were asked to assign school grades to their own knowledge about climate change. This item correlates with “impacts on climate change (all)” (0.338, $p < 0.01$) and is also a predictor for “true facts about climate change” (0.247, $p < 0.05$). This self assessment correlates very clearly (0.605, $p < 0.01$) with the csim index (3-experts-model at t1) but not at t2. The similarity of the knowledge structure (csim) with the teacher’s model (not shown in table) correlates negatively with the number of wrong answers given in the true/false statement test (-0.207, $p < 0.05$). No demographic data showed a correlation with the constructs and parameters used.

DISCUSSION

Although there were some changes visible in the data between t1 and t2, no significant results could be found to verify the hypothesis that perception, knowledge and knowledge structure were affected by participation in the educational scenario analysis. The following discusses the individual aspects in more detail.

Formal Knowledge, perception and attitude: Considering the mean values of all four groups it is noticeable that the two treatment groups behave differently: for instance, while the expert knowledge of TG02 improved slightly it decreased in TG01 (generally the standard deviation is relatively high), but in other cases TG01 performed better.

The age, number of semesters and course of studies did not correlate significantly with results in the SJT and knowledge tests. These results suggest that a) cross-linked thinking is a competency that is independent of courses of studies and b) that while knowledge about climate change is general knowledge it is nevertheless more difficult than expected to advance partly (wrong) general knowledge to expert level. To illustrate: a frequent topic of discussion in the TGs was climate simulations that calculate mean values over time periods of 30 years. This means that it is very difficult to draw conclusions from the results of climate simulations about individual weather events in a specific year. One item (knowledge true/false statement) stated that it will never snow in the Black Forest if the global mean temperature rises by 3°C. This statement is incorrect (as is shown e.g., by the freak snowfall in the Sahara in 2005), but with few exceptions this question was answered incorrectly at t1 as well as at t2 in all groups. In both TGs enough information on and discussion of this topic was offered so that a transfer of this knowledge would have been

possible, yet it did not happen. Assuming that the question was understood correctly, some support can be found in this result for the supposition that participants in the TG learned less than expected.

Perception and attitudes did not correlate significantly with variables nor could significant differences be observed. However, a positive change in personal proactivity was noted, but then the values of the TGs were higher at all stages than those of the CGs. These higher values can also be interpreted as showing that students’ interest in and awareness of this topic made them choose the seminars about scenario analysis and future planning.

Structural knowledge: The results for this aspect can be interpreted in three different ways: (a) the educational scenario analysis did not have a measurable effect on participants’ knowledge structure; or (b) the teacher or the design of the educational scenario analysis was insufficient; or (c) SJT and comparison of indices (with reference models) did not capture relevant knowledge structure or delivered unreliable results.

The falsification of the main hypothesis would contradict all the literature that describes educational effects of participation in a scenario analysis, although the attributed effects have not been proved and measured quantitatively. But nor did the knowledge tests show any significant advancement in knowledge. The CGs did not work on climate change so that no significant change in formal knowledge could be expected. Due to the great relevance of climate change in the media, a significant rise in this performance could have indicated strong outside influence on (all) the seminars, but there was none. Additionally, the CGs and TG01 participated in at least one further lecture or seminar treating the topic of climate change. In the TGs, climate change was discussed as driver and impact factor in connection with the guiding question of the seminars of how tourism in the Black Forest could develop in future. While there was no explicit factual input from the teacher, there were students’ presentations and, as mentioned at the beginning of this chapter, relevant facts and information was offered during the seminars to advance individual performance in the knowledge test. As no significant increase in formal knowledge about climate change was visible, this did not contribute to a confirmation of the main hypothesis. These considerations lead to possibility (b).

Insufficiency of the teacher or the design of the learning setting cannot be investigated by the design of this study. For one, the TGs students’ final reports (assessments) indicated a general understanding of the methodology and most students did indeed report new

insights gained into this complex knowledge topic. Although no qualitative analyses of the seminars' assessments were carried out for this study, some learning effects of participation in the seminar can nonetheless be assumed from them. Evaluation of the seminars by the regular evaluation of the Leuphana University of Lüneburg indicated a (self reported) success of the seminar. For another, the design of the learning setting for the TGs varied (Table 1) while the guiding question and steps within the seminars were exactly the same. Differing influences of the setting of the seminars on the results could not be found.

Measuring systemic thinking and knowledge structure is a difficult task, as was indicated in the theory section of this article. Similarity judgment ratings are used in practice to capture cognitive concept maps. Concept maps deliver insightful qualitative results, but statistical processing is difficult. Although both indexes (sim and R_{SJT}) delivered contradictory prognostic validity in previous studies, csim had greater validity (Goldsmith *et al.*, 1991) or lower validity (Großschedl, 2010) in relation to R_{SJT} . Nevertheless the indexes have delivered valid results for the comparison of conceptual knowledge (Großschedl, 2010) and its structure (Shavelson *et al.*, 2005; Beatty and Gerace, 2002) gained by SJTs. How can the results of SJT comparison in this study be appraised?

First, it can be asserted that participants' self assessed climate change knowledge at t1 correlates highly with the correspondence of structural knowledge to 3-experts-model at t1 which indicates some validity of the measurement instrument. From a constructivist perspective it could be argued that a comparison with an expert's model need not show any approximation as long as the expert did not teach the lessons. Therefore the teacher's cognitive map was analyzed in addition but showed similar results like the external experts' one.

Nor were usable results delivered by the factor-analysis carried out to elicit certain areas of perturbation (differences t1-t2 of the absolute values (of the difference in single pair rating from the teacher's model) or areas of strong reframing in the knowledge structure (individual change from t1-t2). The individual differences from a common shared knowledge structure (group average model) decreased in the TGs although not significantly while the models differed more strongly (significantly in one case, Table 3) in the CGs. From this can be assumed a more common reframing of TG's mental models.

Generally it can be debated whether the 12 concepts chosen for the SJT focus too exclusively on certain aspects so that even in the case of enhanced

systemic thinking hardly any changes could be made visible or measurable. Another explanation also suggests itself - that the general knowledge of climate change, one of the most hotly debated topics at present, was initially relatively high (at t1) so that, again, changes could hardly be measured because the concepts were already strongly linked at t1. Other studies using concept maps for quantitative measurements were confronted with only a few direct links between the concepts at t1 (e.g. only one connection (Wasmann-Frahm, 2005)) which offers broader possibilities for respondents to improve. In contrast, SJT offers the possibility to assign strength to the connection (here from 1-7) whereas in concept maps it is the kind of connection that is assessed (and whether there is one or not). Interestingly, neither the prescribed factor nor the cluster analysis led to usable results in identifying structures. Even if the whole construct of structural knowledge did not change significantly during the treatment, it was reasonable to expect that at least some of its parts would undergo a change. It seems that if the treatment has an effect on participants that there is no common responding scheme.

Finally, climate change as a domain for measuring and comparing structural knowledge could be a problematic choice because even in science many interrelations of concepts are under discussion and cannot be described precisely. Although the three expert ratings used had a correlation >0.8 , a fourth expert could not be included in the model because his rating differed too much from that of the other three. Probably the domain of climate change or other complex sustainability problems are too difficult to be used for exact quantitative comparisons of structural knowledge.

CONCLUSION

The hypothesis that participation in an educational scenario analysis has effects on participants' knowledge, future perception and competency of cross linked thinking could not be verified by the data - no final conclusion could be drawn from it. Pedagogical research is often confronted with weak effects.

On the one hand, some hints were given that while the effects of scenario analysis attested in the literature may be describable qualitatively, a quantitative effect is difficult to measure. The design of the learning setting should also be rethought in order to foster the acquisition of formal knowledge.

On the other hand, a methodology was developed both to measure the competency of systemic thinking by investigating the knowledge structure quantitatively

by SJT and by comparison with experts' reference models and also to investigate a way to enable a large scale investigation of competence development. This instrument should be tested with a different set of concepts which are less linked to each other per se. The domain of climate change was probably not an ideal choice for eliciting structural knowledge.

In addition, the research instrument developed in this article should be tested in a different context and with a different thematic domain. Finally, a quantitative, empirically founded proof of educational effects of scenario analysis has yet to be given.

REFERENCES

- Barth, M., 2009. Assessment of Key Competencies-a Conceptual Framework. In: World in Transition- Sustainability Perspectives for Higher Education, Adomßent M. *et al.* (Eds.). VAS-Verlag Für Akademische Schriften, ISBN: 978-3-88864-464-1, pp: 93-100.
- Beatty, I.D. and W.J. Gerace, 2002. Probing physics students' conceptual knowledge structures through term association. *Am. J. Phys.*, 70: 750-758. DOI: 10.1119/1.1482067
- Burandt, S. and M. Barth, 2010. Learning settings to face climate change. *J. Cleaner Product*, 18: 659-665. DOI: 10.1016/j.jclepro.2009.09.010
- Chi, M.T.H. *et al.*, 1981. Categorization and representation of physics problems by experts and novices. *Cognitive Sci.*, 5: 121-152. DOI: 10.1207/s15516709cog0502_2
- Collins, A.M. and E.F. Loftus, 1975. A spreading-activation theory of semantic processing. *Psychol. Rev.*, 82: 407-428. DOI: 10.1037/0033-295X.82.6.407
- Collins, A.M. and M.R. Quillian, 1969. Retrieval time from semantic memory. *J. Verbal Learn. Verbal Behav.*, 18: 240-247. DOI: 10.1016/S0022-5371(69)80069-1
- de Haan, G., 2006. The BLK '21' programme in Germany: A 'Gestaltungskompetenz'-based model for education for sustainable development. *Environ. Educ. Res.*, 12: 19-32. DOI: 10.1080/13504620500526362
- Diekhoff, G.M., 1983. Testing through relationship judgements. *J. Educ. Psychol.*, 75: 227-233. DOI: 10.1037/0022-0663.75.2.227
- Dörner, D. and A. Wearing, 1995. Complex Problem Solving: Toward a (Computer Simulated) Theory. In: Problem Solving: The European Perspective, Frensch, P. and J. Funke (Eds.). Psychology Press, Australia, ISBN: 0805817832, pp: 65-99.
- Forrester, J.W., 1987. Lessons from system dynamics modeling. *Syst. Dynamics Rev.*, 3: 136-149. DOI: 10.1002/sdr.4260030205
- Godet, M., 2000. The art of scenarios and strategic planning: Tools and pitfalls. *Technol. Forecast. Soc. Change*, 65: 3-22. DOI: 10.1016/S0040-1625(99)00120-1
- Goldsmith, T.E. *et al.*, 1991. Assessing structural knowledge. *J. Educ. Psychol.*, 83: 88-96. DOI: 10.1037/0022-0663.83.1.88
- Gomez, P. and G.J.B. Probst, 1995. Die Praxis des Ganzheitlichen Problemlösens: Vernetzt Denken, Unternehmerisch Handeln, Persönlich Überzeugen. 2nd Edn., Haupt Verlag AG, Austria, ISBN: 3258051879, pp: 301.
- Großschedl, J., 2010. Einfluss ausgewählter instruktionaler Maßnahmen auf Struktur und Niveau zellbiologischen Wissens. 1st Edn., Logos Verlag, Germany, ISBN: 978-3-8325-2565-1, pp: 227.
- Jonassen, D.H. *et al.*, 1993. Structural Knowledge: Techniques for Representing, Conveying and Acquiring Structural Knowledge. 1st Edn., Routledge, USA., ISBN: 0-8058-1009-9, pp: 280.
- Klieme, E., J. Hartig and D. Leutner, 2008. Assessment of Competencies in Educational Contexts. 1st Edn., Hogrefe and Huber Publishers, USA., ISBN: 0889372977, pp: 356.
- Kollmuss, A. and J. Agyeman, 2002. Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behaviour? *Environ. Educ. Res.*, 8: 239-260. DOI: 10.1080/13504620220145401
- Mandler, J.M., 2004. The Foundations of Mind. Origins of Conceptual Thought. 1st Edn., Oxford University Press, Pakistan, ISBN: 0195172000, pp: 359.
- Niedderer, H. *et al.*, 1991. The role of computer-aided modelling in learning physics. *J. Comput. Assis. Learn.*, 7: 84-95. DOI: 10.1111/j.1365-2729.1991.tb00231.x
- Ossimitz, G., 1997. The development of systems thinking skills using system dynamics modeling tools. University of Klagenfurt. http://www.uni-klu.ac.at/gossimit/sdyn/gdm_eng.htm
- Ossimitz, G., 2000. Entwicklung Systemischen Denkens. Theoretische Konzepte und Empirische Untersuchungen. 1st Edn., Profil Verlag, Germany, 2000, ISBN: 3-8902-9494-X, pp: 256.
- Polanyi, M., 1997. The Tacit Dimension. In: Knowledge in Organizations, Prusak, L. (Ed.). Butterworth Heinemann, U.K., pp: 135-146.
- Richmond, B., 1993. Systems thinking: Critical thinking, skills for the 1990s and beyond. *Syst. Dynamics Rev.*, 9: 113-133. DOI: 10.1002/sdr.4260090203

- Rieß, W. and C. Mischo, 2008. Entwicklung und erste Validierung eines Fragebogens zur Erfassung des Systemischen Denkens in Nachhaltigkeitsrelevanten Kontexten. In: Kompetenzen der Bildung für Nachhaltige Entwicklung. Operationalisierung, Messung, Rahmenbedingungen, Befunde, G. de Haan and I. Bormann (Eds.). VS Verlag, Austria, ISBN: 3-5311-5529-6, pp: 215-232.
- Rychen, D.S., 2009. Key Competencies: Overall Goals for Competence Development: An International and Interdisciplinary Perspective. In: International Handbook of Education for the Changing World of Work, Maclean, R. and D. Wilson (Eds.). Springer-Verlag, New York, DOI: 10.1007/978-1-4020-5281-1_169, pp. 2571-2583.
- Ryle, G. and A.R. White, 1972. Collected papers. Philos. Books, 13: 29-32. DOI: 10.1111/j.1468-0149.1972.tb03781.x
- Schank, R. and R. Abelson, 1979. Scripts, plans, goals and understanding: An inquiry into human knowledge structures. Am. J. Psychol., 92: 176-178. DOI: 10.2307/1421499
- Schvaneveldt, R.W., F.T. Durso and D.W. Dearholt, 1989. Network structures in proximity data. Psychol. Learn. Motiv., 24: 249-284. DOI: 10.1016/S0079-7421(08)60539-3
- Seel, N.M., 1991. Weltwissen und Mentale Modelle. 1st Edn., Hogrefe, Canada, ISBN: 3-8017-0489-0, pp: 245.
- Senge, P.M., 1993. The fifth discipline. Measur. Bus. Excell., 1: 46-51. DOI: 10.1108/eb025496
- Shavelson, R.J., M.A. Ruiz-Primo and E.W. Wiley, 2005. Windows into the mind. Higher Educ., 49: 413-430. DOI: 10.1007/s10734-004-9448-9
- Shavelson, R.J., 1972. Some aspects of the correspondence between content structure and cognitive structure in physics instruction. J. Educ. Psychol., 63: 225-234. DOI: 10.1037/h0032652
- Stanners, R.F., L.T. Brown, J.M. Price and M. Holmes, 1983. Concept comparisons, essay examinations, and conceptual knowledge. J. Educ. Psychol., 75: 857-864. DOI: 10.1037/0022-0663.75.6.857
- Swart, R.J., P. Raskin and J. Robinson, 2004. The problem of the future: Sustainability science and scenario analysis. Global Environ. Change, 14: 137-146. DOI: 10.1016/j.gloenvcha.2003.10.002
- Sweeney, L.B. and J.D. Sterman, 2000. Bathtub dynamics: Initial results of a systems thinking inventory. Syst. Dynamics Rev., 16: 249-286. DOI: 10.1002/sdr.198
- Van Notten, P.W.F., J. Rotmans, M.B.A. van Asselt and D.S. Rothman, 2003. An updated scenario typology. Futures, 35: 423-443. DOI: 10.1016/S0016-3287(02)00090-3
- Vester, F., 1989. Leitmotiv Vernetztes Denken: Für Einen Besseren Umgang Mit Der Welt. 2nd Edn., Heyne, Spain, ISBN: 3453028651, pp: 287.
- Weinert, F.E., 2004. Competencies and key competencies: Educational perspective. Int. Encycl. Soc. Behav. Sci., 2433-2436. DOI: 10.1016/B0-08-043076-7/02384-6