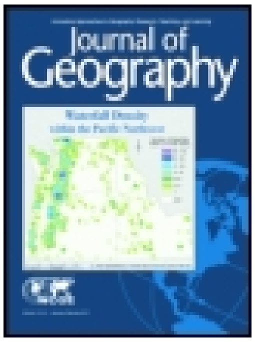

Journal of Geography

ISSN: 0022-1341 (Print) 1752-6868 (Online) Journal homepage: https://www.tandfonline.com/loi/rjog20

The Use of Causal Diagrams to Foster Systems Thinking in Geography Education: Results of an Intervention Study

Marjolein Cox, Jan Elen & An Steegen

The Use of Causal Diagrams to Foster Systems Thinking in Geography Education: Results of an Intervention Study

Marjolein Cox, Jan Elen, and An Steegen

Abstract

Increasing interconnectedness of people and goods enhances the complexity of geographical problems. For students to understand geography, systems thinking—and in this context, the use of causal diagrams—is a promising approach. A quasi-experimental design is used in which the systems thinking ability of students working with causal diagrams is compared to a control group where students did not work with causal diagrams. Pre- and posttests were taken by 448 students in the experimental group and 168 students in the control group. The results indicate that students in the experimental group outperform students in the control group.

Key Words: *systems thinking, secondary geography education, geographical relational thinking*

Marjolein Cox is a teaching assistant in the Teacher Training Program at KU Leuven, Leuven, Belgium and a geography teacher at Sint-Jozef Instituut Kontich, Kontich, Belgium.

Jan Elen is a professor in the Centre for Instructional Psychology and Technology at KU Leuven, Leuven, Belgium.

An Steegen is an assistant professor in the Department of Earth and Environmental Sciences at KU Leuven, Leuven, Belgium.

Introduction

Globalization processes induce a growing interconnectedness between people, goods, and events in different regions. Understanding this rising global complexity and the interaction with the natural environment is a goal in geography courses in secondary education (International Geographical Union 2016). Acquiring insight into global and local interconnections also contributes to the ability to take more sustainable decisions and is therefore part of Education for Sustainable Development (Riess and Mischo 2010).

Systems thinking is a promising approach to understand these interconnections (Assaraf and Orion 2005). This has led to the inclusion of systems thinking in learning outcomes in many locations (e.g., The Next Generation Science Standards for primary and secondary education in the United States; the German educational standards for the subject of geography; and the latest geography curricula in Flanders, Belgium) (Yoon and Hmelo-Silver 2017; Rempfler and Uphues 2012; Katholiek Onderwijs Vlaanderen 2017).

Several studies indicate a rather poor level of students' systems thinking ability (Cox, Elen, and Steegen 2017; Favier and van der Schee 2014; Karkdijk, van der Schee, and Admiraal 2013). Therefore, research should examine how adequate systems-oriented teaching and learning can be designed. This study focuses on the effect of a teaching strategy in which relations within systems are explicitly described and visualized by the use of causal diagrams. In addition to measuring students' systems thinking abilities, content knowledge is measured. As a focus on systems thinking might take more time than traditional teaching strategies, it is therefore important that systems thinking also helps students to better understand geography content as well.

This article investigates the following research questions:

1. What is the effect of the use of causal diagrams in geography lessons on students' systems thinking ability in upper secondary education?
2. What is the effect of the use of causal diagrams in geography lessons on content knowledge?

The operational definition of systems thinking we use points at students' ability to identify variables, to recognize relations, to assign the nature of these relations, to describe relations in words, and to explain influences in a system if there is an interference within the system. The effect of the intervention on these aspects is discussed.

Theoretical Background

Systems Thinking and Its Relation to Teaching and Learning Geography

Many authors agree that systems thinking is about identifying and understanding different components of a system and the relations in between them. Arnold and Wade (2015) compared eight definitions. Four elements often reoccurred, namely *interconnections or interrelationships*, *wholes rather than parts*, *feedback loops*, and *dynamic behavior*. They developed a new definition in which systems thinking is a system itself: "Systems thinking is a set of synergetic analytic skills used to improve the capability of identifying and understanding systems, predicting their behaviors, and devising modifications to them in order to produce desired effects. These skills work together as a system" (Arnold and Wade 2015, 675).

Rempfler and Uphues (2012) distinguished four dimensions in their systems competency model for geography education. The first dimension, *system organization*, is about the identification, modeling and description of the system's organization and its essential elements. This corresponds to what Stave and Hopper (2007) describe as recognizing interconnections, or to what Senge et al. (2000) describes as seeing relationships. The second dimension of Rempfler and Uphues (2012) is *systems behavior*, thus understanding insight the functions and behaviors of a system. This dimension is also recognized by Stave and Hopper (2007) and by Sweeney and Sterman (2000). The other dimensions listed by Rempfler and Uphues (2012) are a *system-adequate intention to act* and *system-adequate action*. In contrast with the first dimensions, these dimensions refer to knowledge application rather than to knowledge acquisition (Mehren et al. 2018).

Systems thinking or relational thinking is inherently part of geographical thinking and hence of geography education. To explain the meaning of geographical thinking, sets of key concepts are used. Maude (2017), for example, uses key concepts such as place, space, environment, and interconnections. Jackson (2006) distinguishes four sets of key concepts to optimize the understanding of connections between places and scales: space and place, scale and connection, proximity and distance, and relational thinking. He argues that these concepts provide a language to think geographically and "a powerful way of seeing the world and making connections between scales, from the global to the local" (Jackson 2006, 199). Lambert (2011, 2017) suggests a three-part framework to explain what thinking geographically entails. Students need geographical vocabulary, grammar, and inquiry. In particular, the grammar of geography refers to relational thinking underpinning geographical thought. It is about a relational understanding of ourselves in the environment and our society. Underlying this relational thinking are key concepts such as place, space, and environment to organize information, to identify a question, or to guide an investigation. Given its importance, geography education aims at promoting this relational thinking as an essential grammatical component of geographical thinking. Teaching geography encourages students to develop the three parts of geographical knowledge and enables them to think geographically.

Relational thinking is thus an important element in geographical thinking (Lambert and Morgan 2010). In relational geographical thinking it is important to examine and to understand interconnections between geographical phenomena from a holistic perspective (Maude 2017). The focus on interconnections between human, environmental, and physical factors at different scales is required to understand issues such as climate change, international migration, or globalization (Fögele 2017). In an attempt to elaborate a theoretical framework, Favier (2017) arranges geographical relational thinking along a scale of geographical complexity. It ranges from thinking about a linear cause–effect relation to thinking in entire systems with many variables, causes, consequences, and feedback loops. The latter connects "to the idea of geographical systems thinking, which is a holistic approach that focuses on how constituent parts of a geographical system are related to each other, how such a geographical system responds to changes, and how geographic systems work within the context of larger geographical systems" (Favier 2017, 95). Hereby geographical relational thinking, as part of geographical thinking, and systems thinking are explicitly connected.

In short, systems thinking can be seen as a form of complex geographical relational thinking. The acquired knowledge enables students to look at the world in a different way and from a more comprehensive background. Therefore, a systems thinking approach might help to broaden and deepen students' worldview and is valuable in geography education.

Interventions to Foster Systems Thinking

A first group of intervention studies focuses on different types of teaching methods. Kali, Orion, and Eylon (2003) tried to foster forty middle school students understanding of the rock cycle system. They pointed to the importance of integrating an activity at the end of the learning program to foster students' understanding of the dynamic and cyclic aspects of the rock cycle. Assaraf and Orion (2005) studied the effect of a study program about the hydrological cycle on fifty middle school students' systems thinking abilities. The program was designed to improve systems thinking and consisted of activities such as constructing concept maps, drawing, and summarizing outdoor experiences. Data collected via several qualitative and quantitative assessment tools indicated meaningful progress in students' systems thinking abilities in the context of the hydrological cycle. The authors also report long-lasting effects of the acquired learning strategy on the students content knowledge six years later (Assaraf and Orion 2010). They suggest paying attention to metacognitive learning patterns in the learning experience. These patterns include awareness about one's own thought processes. Students who figure out, for example, what they don't know or ask to themselves, "What does this tell me?" are trying to assign meaning to the content and are able to maintain an overview on the system (Assaraf and Orion 2010). Karkdijk, van der Schee, and Admiraal (2013) examined the effect of teaching mysteries on students' geographical relational thinking skills. Secondary school students in the intervention solved mystery exercises on different topics and presented them in a concept map. Mystery exercises are challenging problems that students can solve with a card-sorting activity. The learning content in mysteries was linked to the causes and consequences of blood minerals and mobile phones in the Democratic Republic of Congo and to the causes and consequences of slum dwellers in Rio de Janeiro. Students in the experimental group reported more correct geographical relationships in the posttest than students who attended the regular lessons.

A second kind of intervention studies compares the use of a computer simulation as a main element in the instructional design of lessons to foster systems thinking. Riess and Mischo (2010) analyzed the effect of three teaching methods. In the first condition, 115 sixth-grade secondary students school played a computer-simulated forest game to experience the dynamics of the forest as a cultivated ecosystem. The second condition, followed by 112 students, consisted of eleven lessons about ecology of forests, food relationships, and biodiversity. The teaching methods that were used included short lectures, the Socratic method of teaching, and creating cause-and-effect diagrams. In the third condition, 113 students followed a combination of the computer simulation in the first condition and the lessons in the second condition. Only the students in the third condition showed a significant increase in their systems thinking abilities. These findings are consistent with the study of Hmelo-Silver et al. (2017), in which sixty-five middle school students, who were taught a combination of a conceptual representation and modeling practices about ecosystems, deepened their understanding of natural systems in comparison to the forty-seven students who were engaged in traditional instruction.

One last group of intervention studies to foster systems thinking skills uses computer simulations as main element in the instructional design. The results of these interventions reveal mixed results (Pala and Vennix 2005; Sweeney and Sterman 2000; Rates, Mulvey, and Feldon 2016; Smetana and Bell 2012). Each of these authors concluded that computer simulations can be effective. However, the effectiveness depends upon elements such as the incorporation of high-quality support structures and encouragement of student reflection.

Despite differences in context, teaching methods, the content taught, and the assessment tools used in the intervention studies, all findings indicate possibilities to foster systems thinking. However, interventions in geography courses are still rather rare and, according to the knowledge of the authors, intervention studies to foster systems thinking in topics such as international migration, world food problems, or global resource consumption are currently lacking. These topics are very complex because of different but related causes and consequences. A systems thinking approach is suited to increase understanding about these topics.

Causal Diagrams as a Tool to Foster Systems Thinking

Causal diagrams are representations in which the relations between variables are visualized by arrows. A plus or minus sign is added to the arrows to indicate whether it is a positive or negative relation. A relation is positive if an increase of variable A leads to an increase of variable B, or a decrease of variable A leads to a decrease of variable B. A relation is negative if an increase of variable A leads to a decrease of variable B, or a decrease of variable A leads to an increase of variable B (Bala, Arshad, and Noh 2017; Öllinger et al. 2015). Causal diagrams are a variation on concept maps, and both are considered to be effective tools to foster relational thinking (Assaraf and Orion 2005; Novak and Cañas 2008; Mehren, Rempfler, and Ullrich-Riedhammer 2015; Wehry et al. 2012). Whereas all kinds of relations are combined in concept maps, only causal relations are included in causal maps or causal diagrams. Causal diagrams are closely related to actual simulation models because of the use of only one kind of arrow and the use of more abstract plus and minus signs, compared to more descriptive linking words used in concept maps. Causal diagrams can be considered as conceptual models and might therefore be used as a first step in understanding complex geographical systems. In higher education, this conceptual knowledge can then be extended by simulation modeling.

To conclude, given the broad definition of systems thinking and the possibilities of tools, we developed an operational definition of systems thinking as follows: systems thinking is a cognitive skill that enables (1) the ability to construct a causal diagram based on the information of a given source, which means (1a) identifying the relevant variables in the information, (1b) recognizing the relations between the different variables, and (1c) assigning the nature of the relationship (+ or −); (2) the ability to describe relations between variables in words; and (3) the ability to explain the influence within a system if there is an interference.

METHOD

This study examines (1) the effect of using causal diagrams in geography lessons on students' systems thinking ability, and (2) the effect of using causal diagrams in geography lessons on the acquired content knowledge in upper secondary education. A quasi-experimental research design was used.

Research Design

Students in the experimental group ($n = 448$) were taught a lesson series explicitly based on systems thinking via the use of causal diagrams. Students in the control group ($n = 168$) were taught the same content by their own teacher. In these traditional classes, causal diagrams were not used and the researchers did not intervene. All students took a pretest and posttest to measure their systems thinking abilities. Content validation of both tests was performed by expert panels while the reliability was statistically analyzed. A few weeks after the intervention, students also answered three questions on their exam, measuring their achievement level of the curriculum objectives taught in the lesson series. These objectives focused more on content knowledge than on systems thinking.

Twelve teachers in the experimental group participated in a professional learning community. In the first three meetings the design of the lesson series was discussed, suggestions and adjustments considered, and exercises tested. This prepared the teachers to teach the lessons as designed by the researchers. During the lesson series, the teachers also kept a diary in which they wrote down how the students cooperated, what changes if any were made, and what positive or negative issues they encountered while teaching. In addition, two more meetings were held to discuss and evaluate the lesson series.

For the six teachers in the control group, no group meetings were organized, but an individual conversation between the researcher and each teacher took place. The researcher also read the course materials to understand the lessons and their focus on systems thinking elements.

Furthermore, a selection of lessons taught to both the experimental and control groups was observed by the researcher. At the end, at least one lesson of each teacher was observed, and all observations together covered the whole lesson series.

Several ethical considerations were taken into account. First, the lessons were designed in cooperation with experts in geography education such as geography professors and lecturers at college, and in cooperation with teachers in the professional learning community. All these persons reflected on the appropriate level of complexity in the diagrams, the progression throughout the lesson series, the instruction language, and the coverage of the attainment goals. Furthermore, the lessons were tested in a pilot study. Second, students and teachers of both groups signed a consent form. In this form the content of the experiment was explained, as well as the possibility to end their participation at any moment. Third, the test responses were treated anonymously. After the intervention studies we informed the teachers about the results of the research and they received the teaching materials.

Participants

Participating teachers reacted positively to a call that was spread via an association of geography teachers in Flanders and via a mailing list of geography teachers who participated earlier in a professional development activity at the university. In total, eighteen teachers from seventeen schools in Flanders participated in the study with one or more class groups. The groups were enrolled in different degree programs, but they all had the same attainment goals for these topics in geography. This originally led to 735 participating students in grades 11 or 12 (ages 16–18). Because of a lack of time in some classes and to illnesses of individual students, the researchers finally collected data from 616 students in forty-five class groups for all tests (pretest, posttest, and exam questions). There were 168 students in the control group and 448 in the experimental group. An equal amount of female (309) and male (307) students was present.

Intervention and Assessment Tools

Students in the experimental group were trained in seven lessons of fifty minutes each to retrieve information from different sources and to use this information to construct causal diagrams. Discussions were allowed, and they were crucial in grasping the problems as a whole. Students also used these diagrams to examine the effect of certain interventions into the systems.

Topics in the lesson series are food supply, globalization, resources, Earth's carrying capacity, and international migration. These topics are all part of one theme called "carrying capacity of the Earth and global shifts" in the Flemish attainment goals for geography. In the first lesson, students experienced the idea of interrelated elements in a system using an educational game on food supply. In this game all students play an element in the global food supply system and try to find out how they are connected to another element. In the second and third lessons, students construct causal diagrams about global food supply. They receive variables and information in texts, graphs, and maps. The students read this information, discuss the relationships between the variables in small groups, and construct the diagram. Some groups focus on possibilities of agricultural area expansion and intensification. Other groups work on the impact of population growth and prosperity on food demand. The teacher stimulates the students to actively express their reasoning behind relations. The diagrams developed by the students and a synthesis diagram in which the relations between agricultural expansion, intensification, population growth, prosperity, and food demand are shown and are discussed in the whole class group.

In the fourth and fifth lesson, students cover the topics of globalization and resources. The students start by watching a video fragment and identify variables. Afterward they complete a prestructured causal diagram on globalization and construct a diagram from scratch about coltan mining. The prestructured diagram already shows some relations between variables and organizes the relations and variables to be filled in by the students.

In this way it functions as a starting point for the students to support their reasoning. In the sixth lesson, students learn about the concepts of sustainability, such as the P concept (people, planet, profit, and participation). Students apply this concept while discussing the sustainability of their lifestyles. They are challenged to come up with actions to foster a sustainable policy and lifestyle. In the last lesson, students learn about international migration flows and draw small diagrams in pairs about push and pull factors. They place these diagrams into a synthesis diagram and discuss them in class. Cox, Steegen, and Elen (2018) describe the lesson series in detail.

In contrast to students in the experimental group, students in the control group were taught in a more traditional way. The teacher was in front of the classroom and presented the learning content using a PowerPoint presentation including maps, figures, and images. Thereby, the teachers asked questions to the students, stimulated their reasoning, and asked students to take notes. Students did not draw diagrams and did not work in groups, and complex relations behind several causes and consequences were not emphasized.

Two tests were developed to measure the systems thinking abilities of the students. The items in both tests were based on the operational definition of systems thinking. Although both tests look rather similar, the posttest is more difficult. The full description of the design and validation of the pretest can be read in Cox, Elen, and Steegen (2017). The posttest is described in Cox, Elen, and Steegen (2018). Content validation of both tests was performed by expert panels while the reliability was statistically analyzed. The pretest contains four items. In Items 1 and 2, students receive text describing the relation between climate change and refugee flows to Europe. First, students draw a diagram with provided variables; in the second item, they describe why social unrest in Syria can contribute to economic growth in Western Europe. In Items 3 and 4, students read text on the relationship between air travel and prosperity in countries. Students identify influencing variables and draw a causal diagram explaining this relation.

The posttest consists of six items. Items 1 and 2 are shown in Figure 1. A model answer of these items is shown in Figure 2. In these items students read a text about the social and cultural effects of second homes in the south of France. They identify relevant variables (item 1) and construct a causal diagram (item 2) based on the question: 'Which factors cause the high amount of second homes and what are the consequences for the local community?' In Items 3 and 4, a causal diagram is given and students describe a selection of relations between the variables in the diagram, e.g.,: 'Describe in words why the purchase of a flat screen television made in China can contribute to the acidification of oceans.' In Items 5 and 6, students add variables to the diagram based on different information sources. Students' responses are scored by comparing their answers with the model answer of the experts.

The exam questions represent the curriculum objectives but do not have a focus on systems thinking. In Item 1, students look for arguments in a graph to support statements on the influence of immigrants on the financial status of the social security system. Item 2 deals with the relocation of a company due to globalization processes. Students read a newspaper article and assess five provided statements. In Item 3, a cartoon about the world food problem is shown (Figure 3) and students have to describe one political action in Europe to decrease deforestation in Latin America.

Student answers are compared with the model answer, but different answers are taken into account. To check for interrater reliability, forty randomly chosen exams were scored by two raters and the Cohen's kappa was calculated (0.72 for Item 1, 0.86 for Item 2, and 0.69 for Item 3). These coefficients suggest high reliability among raters (Landis and Koch 1977).

In the analysis a multilevel approach was conducted using IBM SPSS statistical software. This allowed us to take the different class groups and different teachers into account when comparing the average score of the experimental and control group.

Results: Quantitative Analysis of the Effects of the Intervention Study

To measure the effect of the intervention study, the mean scores on the pretest, posttest, and on the exam questions are compared. Students in the control group had a mean score of 0.56 out of 1 on the pretest, whereas students of the experimental group scored 0.53. The effect of the group (control vs. experimental) is not significant, meaning that there is no difference between both groups regarding prior knowledge (Table 1).

Students in the experimental group have a mean score of 0.47 on the posttest and outperform the students in the control group who have a mean score of 0.42. In the third model, the group has a significant effect on the score (Table 2). This means that the intervention itself had a positive impact and that the score on the posttest also can be explained by the score on the pretest. A student scoring higher on the pretest is likely to score higher on the posttest as well. Furthermore, 80% of the variance in posttest scores is situated within the class groups, which means that the effect of being in a different class group or being taught by a different teacher explains respectively only 7.2% and 12% of the variance in posttest scores.

The mean score in the experimental group is 0.62 and 0.58 in the control group for the exam questions. The effect of the group is not significant (Table 3), but the scores on the pretest and the posttest have a significant positive effect on the exam scores. Similar to the posttest scores, the variance is largely situated within class groups (89%). Only 2.7% of the variance is situated

Item 1 & 2

Read the text below. Construct on the next page a diagram with variables from the text. The diagram should be an answer on this question:

What causes the number of second homes and what are the consequences for the local community?

Social and cultural effects of second homes.

Second homes are houses that are often only temporary used by the owner. An example is a little house in the South of France bought by a family to spend a few weeks of the summer holidays each year. The increased communication and transport possibilities stimulate the purchase of a second home as this makes it easier to buy such a house and to travel to the house more often. On the other hand, the purchase of such a house might be a reflection of the increased longing for peace in a rapidly changing world. This peace and quietness often exist in a rural village with empty houses, due to earlier emigration of local people. In addition, people should be prosperous to buy a second home. An increasing prosperity is therefore also contributing to the number of second homes.

The presence of several second homes in a municipality causes positive and negative consequences for the local community. The financial status of the municipality increases due to a higher tax revenue. This is because a second home is higher taxed. Furthermore, the tourists living in these second homes bring higher purchasing power for the local retail, at least in the tourist season. But, real estate prices are rising with this higher demand for houses in the region. This causes the effect that local people are not able to pay the rent and purchase prices anymore. Some of these local people decide to emigrate, which leads to a lower contribution to the local community life. The development of this local community life is already difficult because the owners of these second homes are only occupied temporary (often only in the tourist season). So, the owners of the second homes take only part in the community life during this time period. Another consequence of the emigration of local people is that more houses become empty again.

The policy of the municipality is crucial. Each decision should be thought through. E.g. if the municipality decides to invest the increased tax revenue into tourism, the number of tourists is likely to increase. However, it is possible that so many tourists come to the region that a decrease of the quietness and authenticity in the village is caused. Therefore, it becomes less attractive to invest in second homes. On the other hand, the increasing amount of tourists coming to the village means a higher purchase power for the local retailers. After all, tourists love to go out for dinner and consume in other ways. The local retailers will experience an increase in their income and will be less likely to emigrate to another municipality. The prosperity of the retailers increases and maybe they are willing to buy a second home themselves elsewhere. The consequences of the decisions are thus often complex.

Source: based on Schmude and Aevermann (2015).

Figure 1. Item 1 and 2 of the posttest. Students read the text, identify relevant variables, and construct a causal diagram (Source: Cox, Elen, and Steegen, 2018).

between class groups and 8.5% between different teachers.

Discussion: Qualitative Analysis of the Effects of the Intervention Study

To better understand which aspects of systems thinking the intervention had an influence, the student responses on the posttest are discussed. This discussion is structured based on the different aspects of systems thinking in the operational definition.

Identifying Relevant Variables

On average, students in the experimental group identified 9.5 variables, thus slightly more than students in the control group, who identified 9.2 variables. This difference is situated on several levels. Variables that are more frequently identified by students in the experimental group are often variables that were already part of some causal diagrams in the intervention study such as the variables *communication and transport* and *prosperity*. Although these were also used in the control group, they

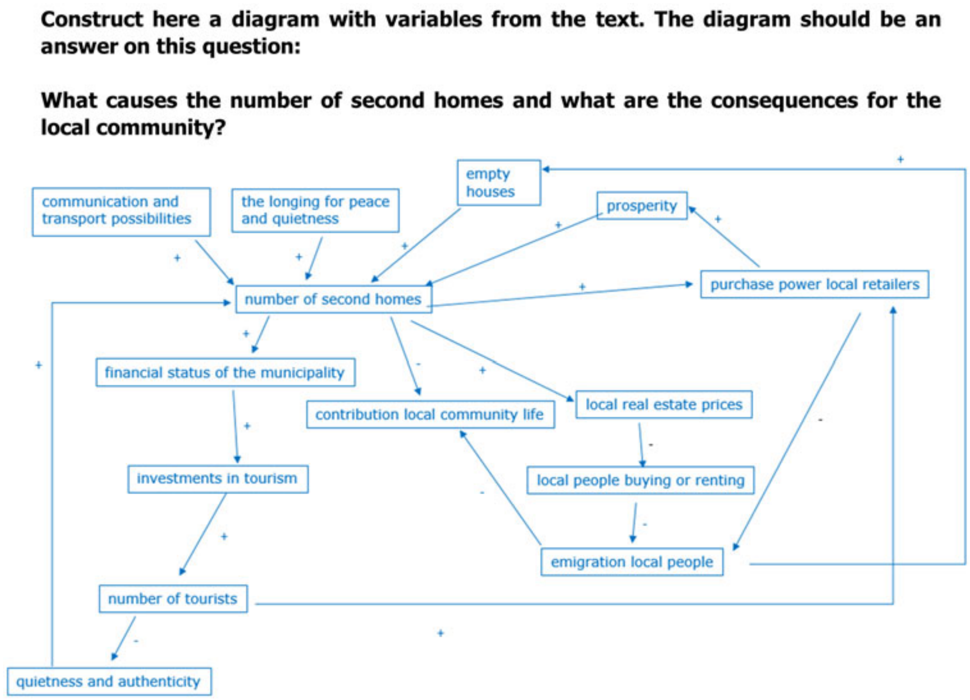

Figure 2. A model answer of Items 1 and 2 of the posttest. This is only one possible causal diagram. Students can draw different correct diagrams (Source: Cox, Elen, and Steegen, 2018). (Color figure available online.)

seem to be better identified because of their explicit use in causal diagrams. Furthermore, some variables were found by almost all students, and some were hardly detected by any student in both groups. This can partly be explained by the frequency of these variables in the text sources, as well as by their occurrence in the research question.

Recognizing Relations between Different Variables

Recognizing relations is specifically tested in Item 2 of the posttest. Students in the control group drew 10 relations on average; this is slightly less than students in the experimental group, who drew 11.25 relations. For both groups about 63% of the drawn relations are correct, but only 37.5% of all correct relations are drawn by the students in the experimental group and 33.5% in the control group. This might be due to the explicit demand to draw these causal diagrams during the intervention in class. Teachers in the professional learning communities mentioned already that they had to stimulate students in some class groups to think about and to draw relations. Some students were indeed unfamiliar with these kinds of teaching activities and were afraid to draw incorrect relations. This hesitation disappeared during the lesson series for students in the experimental group. Moreover, teachers mentioned at the end that students reasoned more spontaneously and focused on understanding relations. This contrasts with students in the control group, who were often only implicitly taught about different relations and were not stimulated to take these interconnections into account. The combination of relations was in this group not visualized in a diagram.

Frequently missing relations in diagrams of both groups were often relationships between variables who were not identified by students in the information sources. Complementarily, incorrect relations are often relations in which an in-between variable was not identified. This indicates that although the reasoning of the students might be correct, this is not rewarded in the test score. Therefore, it might be assumed that the test scores are an underestimation of the students' real systems thinking ability.

Assigning the Nature of the Relations

In general, a majority of the students were able to assign the correct sign to the relation. Students in the control group consistently drew fewer correct signs than those in the experimental group. In more detail, students find it really difficult when two related variables decrease. Only a few students drew the correct plus sign in that case. Also, when the first variable in the relation decreases and the second one increases, it is hard for students to assign the correct minus sign.

Students in the control group assigned the correct sign on average 15.6% less than those in the experimental group. Students in the control group also used "other

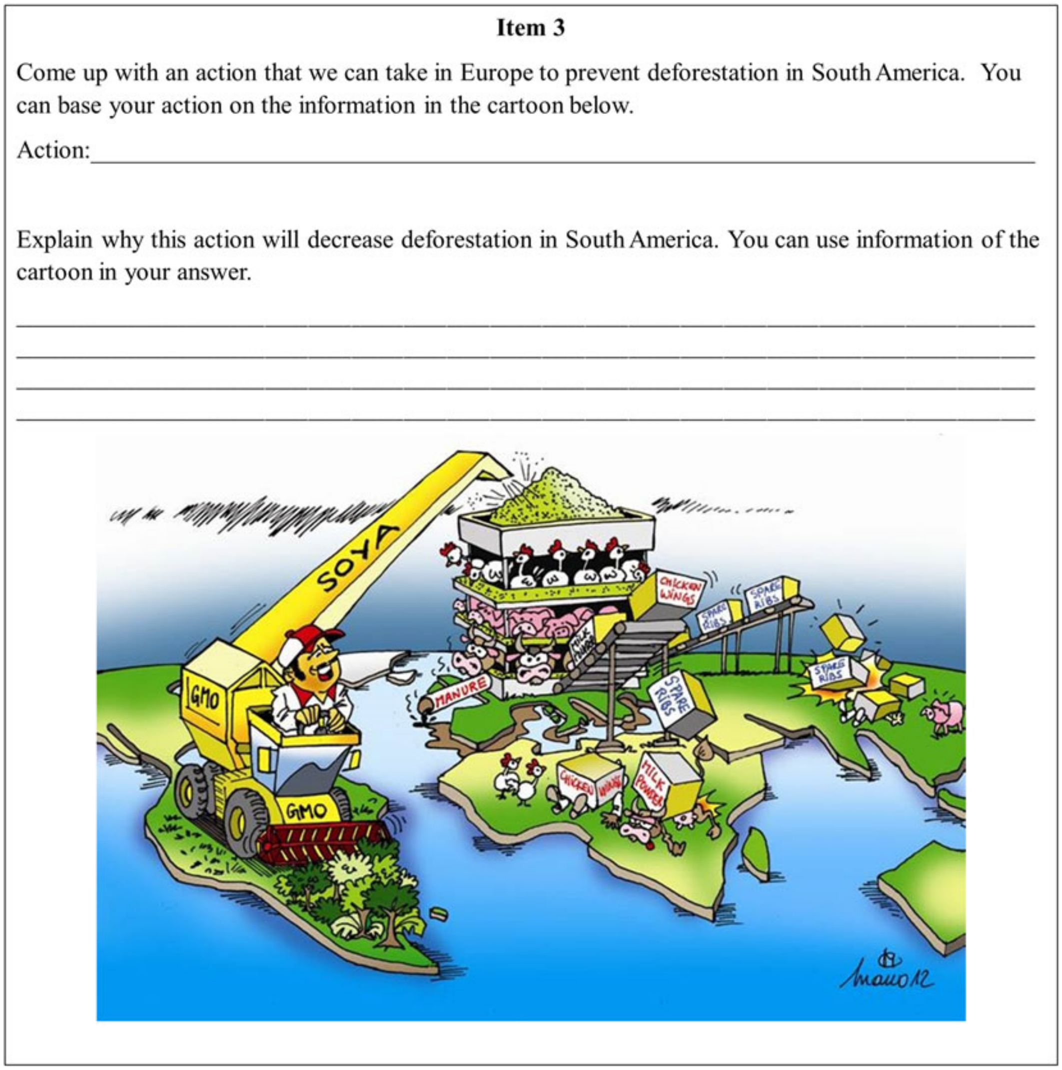

Figure 3. The third exam question, measuring one of the curriculum objectives, where geographical relating is required to achieve the objective. (Friends of the Earth Europe et al. [2012]. Used with permission.) (Color figure available online.)

signs" more often, such as a combination of a plus and a minus sign, an equality sign, or some words. Despite these differences a rather large group of students was able to assign the correct sign. Only in three of nineteen correct relations did a majority of the students use an incorrect sign. One example is a relation between two decreasing variables: only 20.19% of the students in the experimental group and 2.63% of the students in the control group who drew this relation added the correct plus sign. We assume that all students understood this particular relation in his context but that they were unable to assign the correct relation in a more abstract diagram. However, assigning the correct sign is crucial to understand the dynamics in the system and to understand the behavior of a variable when another variable in the system changes.

These results are in line with our expectations. It is rather counterintuitive to assign a plus sign if the two variables in the relations decrease. Some teachers in the learning communities also mentioned difficulties with assigning the nature of relations, as some students associated a relation with a plus sign with a positive effect and a minus sign with a negative effect. For example, if more deforestation leads to more greenhouse gasses in the atmosphere, a plus should be used, but as the effect is negative for our planet, students tended to draw a minus sign. These results point to the need for a clear instruction phase, on one hand, and a rather explicit formulation of relations in the used sources, on the other. Doing so, even more progress is possible.

Describing Relations between Variables in Words

Students in the experimental group more often use the correct variables in their description compared to

Table 1. Multilevel analysis of the scores on the pretest.

	Notation	Parameter Estimates and Standard Errors: Model 1 (Null Model)	Model 2
Fixed effects			
Intercept	γ_{000}	0.55 (0.032)*	0.56 (0.034)*
Group (control)	γ_{010}		−0.016 (0.037)
Random effects			
Variance between teachers	$\sigma^2_{v_{00k}}$	0.014 (0.0059)	0.015 (0.0060)
Intraclass correlation	ρ_v	0.33	0.35
Variance between class groups	$\sigma^2_{u_{0jk}}$	0.0021 (0.0011)	0.0022 (0.0011)
Intraclass correlation	ρ_u	0.050	0.051
Variance between students	$\sigma^2_{r_{ijk}}$	0.026 (0.0015)	0.026 (0.0015)
Intraclass correlation	ρ_r	0.62	0.60

Note. Equation for the second model: $scorepretest_{ijk} = \gamma_{000} + \gamma_{010} * Group + r_{ijk} + u_{0jk} + v_{00k}$ with $r_{ijk} \sim N\left(0, \sigma^2_{r_{ijk}}\right)$ and $u_{0jk} \sim N\left(0, \sigma^2_{u_{0jk}}\right)$ and $v_{ook} \sim N(0, \sigma^2_{v_{00k}})$.
*$p < .05$.

Table 2. Multilevel analysis for the scores on the posttest.

	Notation	Parameter Estimates and Standard Errors: Model 1 (Null Model)	Model 2	Model 3
Fixed effects				
Intercept	γ_{000}	0.45 (0.018)*	0.34 (0.018)*	0.35 (0.017)*
Score pretest	γ_{100}		0.22 (0.022)*	0.23 (0.022)*
Group (control)	γ_{010}			−0.057 (0.018)*
Random effects				
Variance between teachers	$\sigma^2_{v_{00k}}$	0.0042 (0.0019)	0.0018 (0.00098)	0.0013 (0.00076)
Intraclass correlation	ρ_v	0.27	0.16	0.12
Variance between class groups	$\sigma^2_{u_{0jk}}$	0.0018 (0.00067)	0.0010 (0.00045)	0.00075 (0.00037)
Intraclass correlation	ρ_u	0.12	0.089	0.072
Variance between students	$\sigma^2_{r_{ijk}}$	0.0094 (0.00055)	0.0084 (0.00050)	0.0084 (0.00050)
Intraclass correlation	ρ_r	0.61	0.75	0.80

Note. Equation for the third model: $scoreposttest_{ijk} = \gamma_{000} + \gamma_{100} * scorepretest + \gamma_{010} * group + r_{ijk} + u_{0jk} + v_{00k}$ with $r_{ijk} \sim N\left(0, \sigma^2_{r_{ijk}}\right)$ and $u_{0jk} \sim N\left(0, \sigma^2_{u_{0jk}}\right)$ and $v_{ook} \sim N(0, \sigma^2_{v_{00k}})$.
*$p < .05$.

Table 3. Multilevel analysis for the scores on the exam questions.

	Notation	Parameter Estimates and Standard Errors: Model 1 (Null Model)	Model 2	Model 3
Fixed effects				
Intercept	γ_{000}	0.62 (0.023)*	0.40 (0.035)*	0.40 (0.036)*
Score pretest	γ_{100}		0.14 (0.043)*	0.15 (0.043)*
Score posttest	γ_{200}		0.30 (0.070)*	0.30 (0.071)*
Group (control)	γ_{010}			−0.013 (0.026)
Random effects				
Variance between teachers	$\sigma^2_{v_{00k}}$	0.0070 (0.0030)	0.0026 (0.0015)	0.0026 (0.0016)
Intraclass correlation	ρ_v	0.20	0.086	0.085
Variance between class groups	$\sigma^2_{u_{0jk}}$	0.00070 (0.00071)	0.00072 (0.00068)	0.00082 (0.00071)
Intraclass correlation	ρ_u	0.020	0.024	0.027
Variance between students	$\sigma^2_{r_{ijk}}$	0.028 (0.0017)	0.027 (0.0016)	0.027 (0.0016)
Intraclass correlation	ρ_r	0.65	0.89	0.89

Note. Equation of the third model: $Scorequestionsexam_{ijk} = \gamma_{000} + \gamma_{100} * scorepretest + \gamma_{200} * scoreposttest + \gamma_{010} * Group + r_{ijk} + u_{0jk} + v_{00k}$ with $r_{ijk} \sim N\left(0, \sigma^2_{r_{ijk}}\right)$ and $u_{0jk} \sim N\left(0, \sigma^2_{u_{0jk}}\right)$ and $v_{ook} \sim N(0, \sigma^2_{v_{00k}})$.
*$p < .05$.

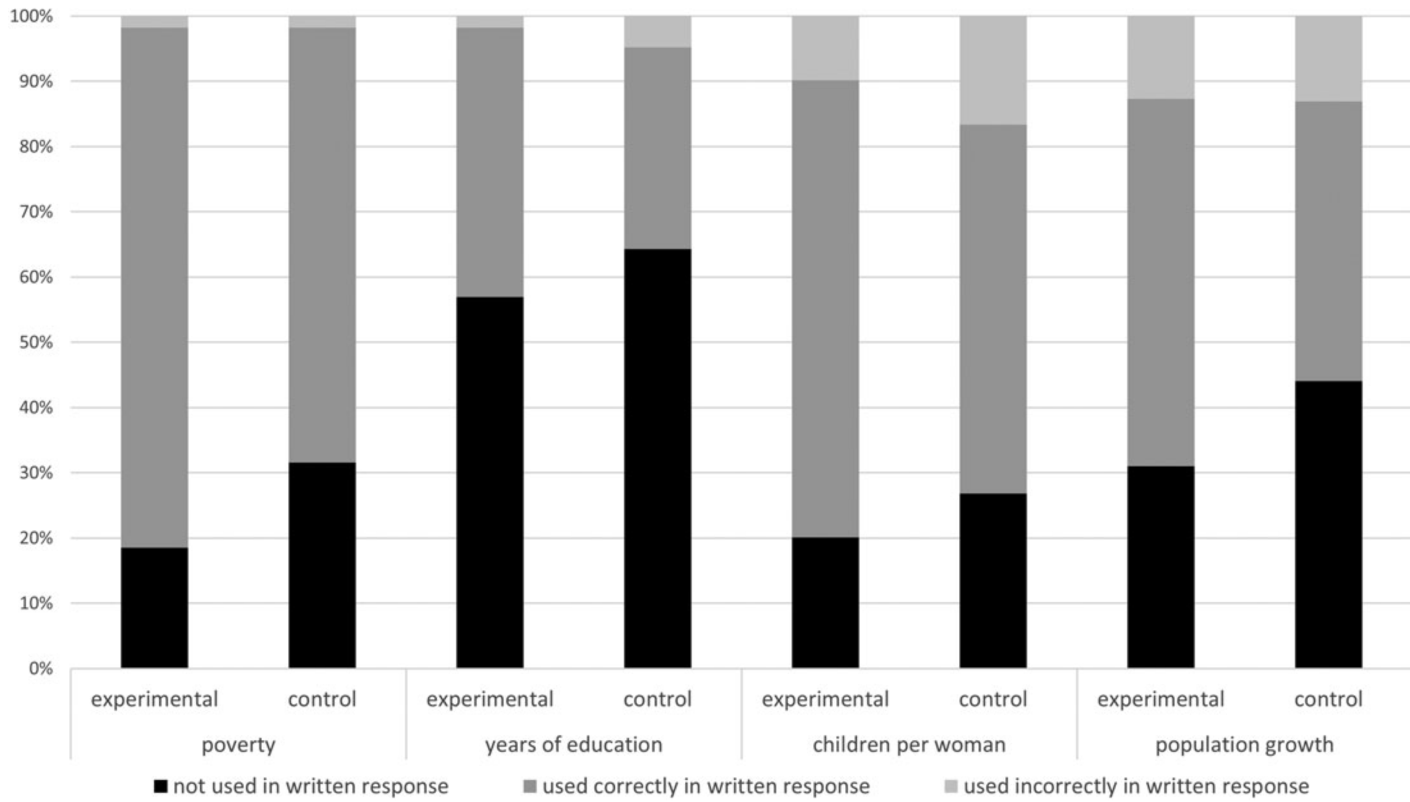

Figure 4. The percentage of students in both groups who used the relevant variables in their written response of item 4 of the posttest. It is indicated whether the variable is used in a correct way or not.

students in the control group. As an example, the results of Item 4 in the posttest are shown (Figure 4). All variables are used less by the students in the control group but, at the top, the variable *years of education* is used relatively less by both groups. This variable can be skipped while explaining the relation between the variables poverty and population growth (Figure 5), but this gives only a partial answer to the question of Item 4, which asks for the influence of improving poverty on population growth. Based on the given diagram the variables *children per woman* and *years of education* are both related to the population growth, so both variables must be present in the students' answer. Some variables are also used incorrectly in students' answers, probably because students have to take a dynamic change in the variables into account to correctly interpret the diagram.

Explaining Influences in a System if There Is an Interference

When asking students to add one or more variables in the causal diagram based on sources like graphs and maps, students in the experimental group more often identified these extra variables (Table 4).

To conclude, students in the experimental group outperformed students in the control group for different aspects of systems thinking. In general, they identified more variables, drew more correct relations, and assigned the correct sign to relations more often. They were also better able to read and explain causal diagrams. This corresponds to earlier research by Kali, Orion, and Eylon (2003); Assaraf and Orion (2005); and Karkdijk, van der Schee, and Admiraal (2013), who all found a positive effect of comparable methods and visualization techniques, such as concept maps, on geographical reasoning.

Content Knowledge

The quantitative analysis showed that the total score on the exam questions, measuring content knowledge, is better for the students in the experimental group although not significant. The results also show that the intervention has no or only a small effect on content knowledge in which relating is not included (Table 5). However, if curriculum objectives in which relating is required are considered, the effect of the intervention is visible. Indeed, for Item 1, students in the control group outperform students in the

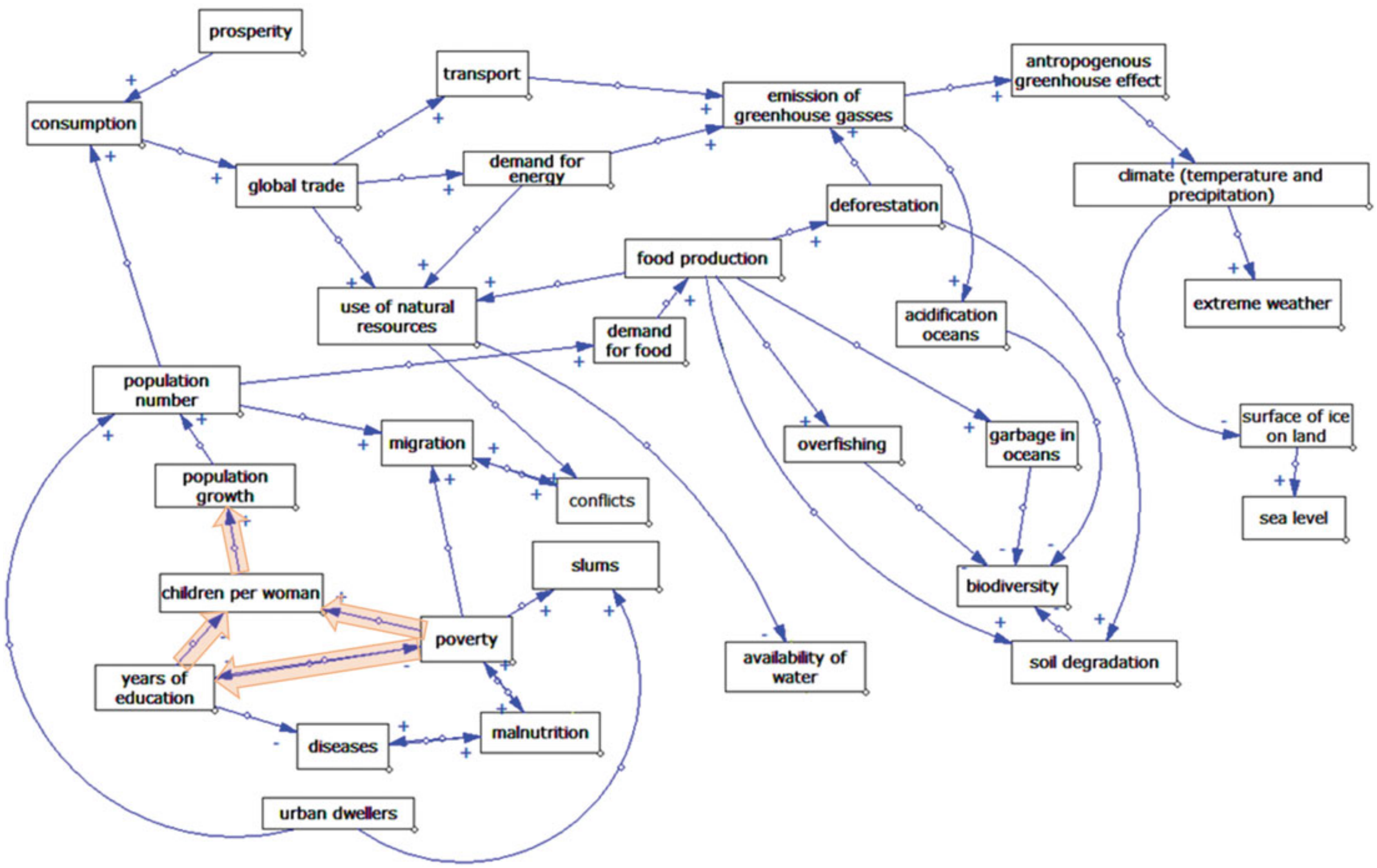

Figure 5. The diagram provided in Item 4 of the posttest. The orange arrows indicate the two paths to answer the question in the following instruction: "The first Sustainable Development Goal is to finish poverty in all its forms everywhere. If this goal is accomplished, what influence does it have on the population number on the long run? Use variables of the provided diagram in your response. Describe this in words." (Content of the diagram inspired from Hoffmann, 2013.)

experimental group, but the largest difference between the experimental and control group can be found in Item 3. This is the only item in which students have to relate several elements to each other and is therefore the most associated with systems thinking.

Concerning the importance of geographical relational thinking and systems thinking in a geography course, we advocate that more curriculum objectives contain a level of relational thinking. The outcomes in this study suggest that a focus on systems thinking in class has a leverage effect on the goals that are important to achieve in secondary geography education. It contributes to the students' geographic literacy, certainly in terms of relational thinking (Jackson 2006). Teachers of the experimental group mentioned that the students reasoned much more about the content while constructing the causal diagrams. Many enjoyed hearing their students discussing the relations and processing the geographical content.

Additional Considerations

Although students in the experimental group score better than those in the control group, these posttest scores are still rather disappointing. The relatively short period of the intervention might explain this, and we expect higher test scores if students are continuously stimulated in class to relate multiple concepts. Also, group work was not allowed during the posttest while this was stimulated during the intervention and was appreciated by teachers and students. Finally, other research has shown that systems thinking and relational thinking in general is very difficult (Cox, Elen, and Steegen 2017; Favier and van der Schee 2014; Karkdijk, van der Schee, and Admiraal 2013). The positive effect of this short intervention shows the possibility of fostering this cognitive skill.

The qualitative analysis revealed that language has an influence on student systems thinking abilities. Students take those relations and variables into account that are explicitly mentioned in the information sources, whereas

Table 4. Percentage of students of both groups that identified the variables or relations in the fifth item of the posttest.

	% of Students Who Identified the Variable or Relation					
	Variables		Relations			
Group	Population in Low Coastal Areas	Vulnerability to Floods	1	2	3	4
Experimental	27.90	84.82	77.23	4.68	14.4	7.81
Control	22.61	73.81	62.50	1.78	8.33	5.95

No. relation	Relation
1	sea level → vulnerability to floods
2	population in coastal areas → vulnerability to floods
3	poverty → vulnerability to floods
4	population number → population in coastal areas

Table 5. Overview of the mean scores and standard deviation of both groups for the three exam questions.

	Item 1		Item 2		Item 3	
Group	M	SD	M	SD	M	SD
Experimental	0.48	0.22	0.68	0.20	0.69	0.39
Control	0.52	0.21	0.65	0.22	0.57	0.39

relations that are implicitly present are much harder to identify. This was confirmed by teachers in the professional learning communities. It also corresponds to studies about causal reasoning in history courses, which found positive effects of explicit teaching strategies in intervention studies (Stoel, van Drie, and van Boxtel 2015, 2017). This implies that a clear and thoughtful use of language in the information sources of the test might be necessary. However, students should be able to understand relations and systems in different information sources that are not designed to work with in a school environment. In these information sources, such as texts, figures, or video fragments, relations are often less explicitly formulated compared to the relations in the information sources designed for this experiment. Therefore, teaching strategies should be developed gradually and researched to find out how to formulate causal relations in geographical systems on a local and global scale and how this affects students understanding.

Conclusion

This intervention study shows that the use of causal diagrams fosters students' systems thinking abilities in geography courses to improve and engage relational and geographical thinking. While constructing these diagrams, students learn to explain their reasoning and the relations between different variables on different geographical scales depicted in the diagram. The students actively process the geographical content, which increases their geographical reasoning and their insight in the complexity of the system. Therefore, it fosters students relational thinking and it contributes to the construction of geographical knowledge. Most of the teachers were positive about the use of these diagrams and the course materials used in the intervention. The intervention also showed that teachers of geography should pay attention to the formulation of relations in different kinds of information sources and their own explanations. However, more research is required to understand these effects and to detect preferable formulations. Finally, this study justifies a gradual implementation of systems thinking in primary and secondary geography education. However, research is necessary to examine which activities are feasible at different ages.

References

Arnold, R. D., and J. P. Wade. 2015. A definition of systems thinking: A systems approach. *Procedia Computer Science* 44:669–78.

Assaraf, O. B.-Z., and N. Orion. 2005. Development of system thinking skills in the context of Earth system education. *Journal of Research in Science Teaching* 42 (5):518–60. doi: 10.1002/tea.20061.

Assaraf, O. B.-Z., and N. Orion. 2010. Four case studies, six years later: Developing system thinking skills in junior high school and sustaining them over time. *Journal of Research in Science Teaching* 47 (10):1253–80.

Bala, B. K., F. M. Arshad, and K. M. Noh. 2017. *System dynamics. Modelling and simulation*. Signapore: Springer Nature.

Cox, M., A. Steegen, and J. Elen. 2018. Using causal diagrams to foster systems thinking in geography. *International Journal of Designs for Learning* 9 (1):34–48.

Cox, M., J. Elen, and A. Steegen. 2017. Systems thinking in geography: Can high school students do it? *International Research in Geographical and Environmental Education* 28 (1):37–52. doi: 10.1080/10382046.2017.1386413.

Cox, M., J. Elen, and A. Steegen. 2018. A test to measure students' systems thinking abilities in geography. *European Journal of Geography* 9 (1):105–20.

Favier, T. T. 2017. Towards a theoretical framework for geographic relational thinking. Paper presented at the CGE Lisbon Symposium, Lisbon, Portugal, October 26.

Favier, T. T., and J. A. van der Schee. 2014. The effects of geography lessons with geospatial technologies on the development of high school students' relational thinking. *Computers and Education* 76:225–36.

Fögele, J. 2017. Acquiring powerful thinking through geographical key concepts. In *The power of geographical thinking*, ed. C. Brooks, G. Butt, and M. Fargher, 59–73. London: Springer International Publishing.

Friends of the Earth Europe, Arc2020, Ceo, Concord, European Environmental Bureau, Food and Water Watch Europe, Ifoam Eu, Pesticide Action Network Europe, European Coordination Via Campesina, Act Alliance. 2012. Conference on crop rotation and legume cultivation. Accessed January 14, 2018. http://www.foeeurope.org/europe-increase-crop-rotation-legume-cultivation.

Hmelo-Silver, C. E., R. Jordan, C. Eberbach, and S. Sinha. 2017. Systems learning with a conceptual representation: A quasi-experimental study. *Instructional Science* 45 (1):53–72.

Hoffmann, T. 2013. Globale Herausforderungen-ein Mysterium? [Global challenges- a mystery?]. *Praxis Geographie*, 43 (9):36–43.

International Geographical Union. 2016. International charter on geographical education. Accessed January 14, 2018. http://www.igu-cge.org/charters-pdf/2016/igu_2016_def.pdf.

Jackson, P. A. 2006. Thinking geographically. *Geography* 91 (3):199–204.

Kali, Y., N. Orion, and B. S. Eylon. 2003. Effect of knowledge integration activities on students' perception of the earth's crust as a cyclic system. *Journal of Research in Science Teaching* 40 (6):545–65.

Karkdijk, J., J. van der Schee, and W. Admiraal. 2013. Effects of teaching with mysteries on students' geographical thinking skills. *International Research in Geographical and Environmental Education* 22 (3):183–90.

Katholiek Onderwijs Vlaanderen. 2017. Curriculum goals secondary education: Geography 11th and 12th grade. Accessed January 14, 2018. http://ond.vvkso-ict.com/leerplannen/doc/aardrijkskunde-2017-010.pdf.

Lambert, D. 2011. Reviewing the case for geography, and the "knowledge turn" in the English national curriculum. *Curriculum Journal* 22 (2):243–64.

Lambert, D. 2017. Chapter 2: Thinking geographically. In *The handbook of secondary geography*, ed. M. Jones, 20–29, 1st ed. Sheffield: Geographical Association.

Lambert, D., and J. Morgan. 2010. *Teaching geography 11-18: A conceptual approach*. Maidenhead: Mcgraw-Hill Education.

Landis, J. R., and G. G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1):159–74.

Maude, A. 2017. Applying the concept of powerful knowledge to school geography. In *The power of geographical thinking*, ed. C. Brooks, G. Butt, and M. Fargher, 27–40. London: Springer International Publishing.

Mehren, R., A. Rempfler, and E. M. Ullrich-Riedhammer. 2015. Diagnosis of system competence through concept maps. Malaria control in the Congo as an example. *Praxis Geographie* 7–8:29–33.

Mehren, R., A. Rempfler, J. Buchholz, J. Hartig, and E. M. Ulrich-Riedhammer. 2018. System competence modeling: Theoretical foundation and empirical validation of a model involving natural, social, and human-environment systems. *Journal of Research in Science Teaching* 55 (5):685–711.

Novak, J. D., and A. J. Cañas. 2008. The theory underlying concept maps and how to construct and use them. Accessed March 29, 2019. https://cmap.ihmc.us/docs/theory-of-concept-maps.

Öllinger, M., S. Hammon, M. VON Grundherr, and J. Funke. 2015. Does visualization enhance complex problem solving? The effect of causal mapping on performance in the computer-based microworld Tailorshop. *Educational Technology Research and Development* 63 (4):621–37.

Pala, Ö., and J. A. M. Vennix. 2005. Effect of system dynamics education on systems thinking inventory task performance. *System Dynamics Review* 21 (2):147–72.

Rates, C. A., B. K. Mulvey, and D. F. Feldon. 2016. Promoting conceptual change for complex systems understanding: Outcomes of an agent-based participatory simulation. *Journal of Science Education and Technology* 25 (4):610–27.

Rempfler, A., and R. Uphues. 2012. system competence in geography education development of competence models, diagnosing pupils' achievement. *European Journal of Geography* 3 (1):6–22.

Riess, W., and C. Mischo. 2010. Promoting systems thinking through biology lessons. *International Journal of Science Education* 32 (6):705–25.

Senge, P., N. H. Cambron-Mccabe, T. Lucas, B. Smith, J. Dutton, and A. R. T. Kleiner. 2000. *Schools that learn: A fifth discipline fieldbook for educators, parents and everyone who cares about education*, ed. P. Senge. New York: Doubleday.

Smetana, L. K., and R. L. Bell. 2012. Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education* 34 (9):1337–70.

Stave, K., and M. Hopper. 2007. What constitutes systems thinking? A proposed taxonomy. Paper presented at the 25th International Conference of the System Dynamics Society. Boston, July 29–August 3, 2007.

Stoel, G. L., J. P. van Drie, and C. A. M. van Boxtel. 2015. Teaching towards historical expertise. Developing a pedagogy for fostering causal reasoning in history. *Journal of Curriculum Studies* 47 (1):49–76.

Stoel, G. L., J. P. van Drie, and C. A. M. van Boxtel. 2017. The effects of explicit teaching of strategies, second-order concepts, and epistemological underpinnings on students' ability to reason causally in history. *Journal of Educational Psychology* 109 (3):321–37.

Sweeney, L. B., and J. Sterman. 2000. Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review* 16 (4):249–86.

Wehry, S., H. Monroe-Ossi, S. Cobb, and C. Fountain. 2012. Concept mapping strategies: Content, tools and assessment for human geography. *Journal of Geography* 111 (3):83–92.

Yoon, S. A., and C. Hmelo-Silver. 2017. Introduction to special issue: Models and tools for systems learning and instruction. *Instructional Science* 45 (1):1–4.