Routledge
Taylor & Francis Group

International Research in Geographical and Environmental Education

ISSN: 1038-2046 (Print) 1747-7611 (Online) Journal homepage: http://www.tandfonline.com/loi/rgee20

# Systems thinking in geography: can high school students do it?

Marjolein Cox, Jan Elen & An Steegen

INTERNATIONAL RESEARCH IN GEOGRAPHICAL AND ENVIRONMENTAL EDUCATION, 2017
https://doi.org/10.1080/10382046.2017.1386413



# Systems thinking in geography: can high school students do it?

Marjolein Cox[a], Jan Elen[b] and An Steegen[a]

[a]KU Leuven, Faculty of Science, Department of Earth and Environmental Sciences, Leuven, Belgium; [b]KU Leuven, Faculty of Psychology and Educational Sciences, Centre for Instructional Psychology and Technology, Leuven, Belgium



**ABSTRACT**

An increasing interconnectedness of people and goods enhances the complexity of many geographical problems. For students to understand geography, systems thinking is a promising approach. It helps to understand increasing complexity by looking at the entire system and at the interconnectedness between the elements in the system. In order to develop adequate systems oriented teaching and learning the current state of the art of students' systems thinking ability needs to be better understood. The authors developed a measuring tool in the form of a paper-and-pencil test in which 735 students in the last or penultimate year of secondary school (age 16–18 year) in Flanders, Belgium, took part. The main findings reveal a rather poor general level of students' systems thinking ability. Students have many difficulties recognizing relationships between variables when several elements of systems thinking come together such as in feedback loops, interactions between human and physical environment, and a combination of different information sources. Rather great differences were found according to the students' study background as well as an interaction effect between grade and gender of the students.

**KEYWORDS**

Systems thinking; geography education; secondary education; geospatial relational thinking

## Introduction

Geography as a course in secondary education aims to broaden students' worldview by helping them to better understand the complexity of our globe (International Geographical Union, 2016). This complexity often includes problems dealing with the interaction between social and natural sciences on multiple spatial and time scales, such as climate change, hunger, poverty and deforestation (Richmond, 1993). By understanding global interconnectedness, students will be able to create a geographical framework in which they can situate topical events. As globalization increases, interconnectedness will also continue to grow. Understanding complex systems is fundamental to understanding science (Hmelo-Silver, Marathe, & Liu, 2007) and part of powerful geographical knowledge (Maude, 2017).

Recently, different countries explicitly included the understanding of systems in intended learning outcomes for science courses. The prominent feature of systems in the

**CONTACT** Marjolein Cox marjolein.cox@kuleuven.be

Next Generation Science Standards for primary and secondary education in the United States is a first example (Yoon & Hmelo-Silver, 2017). In Germany, insight in complex systems is a fundamental competence in geography education (Rempfler & Uphues, 2012). In Flanders (Belgium), systems thinking does also play an important role in recent geography curricula (Katholiek Onderwijs Vlaanderen, 2017). As a first step to design adequate systems oriented teaching and learning, it is essential to understand the current state of the art of students' systems thinking ability. A good understanding of the students' prior knowledge is a crucial part of teachers' pedagogical content knowledge (Shulman, 1987). This knowledge will help teachers to develop appropriate teaching materials. The high school geography curriculum in Flanders integrates both physical and social geography in one geography course in high school. As we state further on that the interconnections between humans and their environment are essential for systems thinking in geography, the integrated organization of a course is an important start as well.

Hence, Flanders is an interesting case to study the following research questions: first, what is the current level of systems thinking in geography for students (age 16–18) in Flanders? Second, what specific aspects of systems thinking experience students most problems with? Aspects specific to systems thinking in general, to (systems thinking in) geography, and different types of tasks in the items are considered in the study. Third, to what extent is the students' level of systems thinking influenced by their degree program, their gender and the grade they are in?

## Theoretical background

### *Different approaches on systems thinking in education*

Solving global challenges requires a thorough understanding of their dynamic systems in order to predict their behavior (Arnold & Wade, 2015). Research from various fields has been done on how students handle complex systems. These studies all emphasize the importance of fostering systems thinking in education as students are the future decision makers in a rapidly changing world (Forrester, 2007; Hmelo-Silver, Jordan, Eberbach, & Sinha, 2017; Senge, 2010; Sweeney & Sterman, 2007).

First, a group of researchers active in research on systems thinking belongs to the System Dynamics Society. In system dynamics, quantitative and dynamic simulation analyses are focused upon (Forrester, 2007). Apart from studies on the improvement of the teaching on modeling (Richardson, 2014), research pertains students' understanding of the basic concepts to build a model, such as stock and flow relationships, feedback loops and time delays. Sweeney and Sterman (2000), for example, examined the understanding of stock and flow by asking students to determine how the quantity of water in a bathtub varies over time, if water is flowing in at a given flow rate and flowing out at another flow rate.

Second, several educational studies are conducted on the use of models and simulations to foster the understanding of science. These studies are similar to those in the field of system dynamics, but are often based on theoretical frameworks from education and learning such as inquiry-based learning. Several of these studies suggest that the use of simulations can be more effective than traditional instruction to achieve science content knowledge and scientific reasoning skills (Smetana & Bell, 2012). Using these simulations correctly is crucial to foster the understanding of science. Wijnen, Mulder, Alessi, and Bollen (2015), for example, studied three types of model instruction to learn about the process of glucose–insulin regulation.

Within this field of education, a subgroup adopts a framework, called Structure, Behavior and Function (SBF), as an approach to learn about complex systems. This framework guides students by structuring their thinking process and might help them to consider relevant structures and to observe the behaviors and the functional role of a complex system (Hmelo-Silver et al., 2017). This approach is used in the context of ecosystems, but also human respiration and car engines as complex systems. Structure refers to the physical structure of a system, for example, the lungs in the human respiration system. Function refers to the purpose of a system, for example supply oxygen to provide energy. The behavior of a system refers to the mechanisms that cause changes in the structural state of a system, for example, the cellular processes (Hmelo-Silver et al., 2007). Due to invisibility and time-delayed causality, the hardest aspect for students to understand is the system's behavior (Hmelo, Holton, & Kolodner, 2000). To improve the understanding of the behavior, these researchers also study the effect of computer simulations as a learning tool (Hmelo et al., 2000; Hmelo-Silver et al., 2017).

Although these different approaches all contribute to understanding complex systems, references to geography are missing. Nevertheless, geospatial relational thinking, an important goal in geography, and part of the geographic literacy that should be achieved by students (Favier & van der Schee, 2014; Jackson, 2006), connects to systems thinking. Indeed, systems thinking focuses on recognizing the interconnections between parts of the system, understanding feedback loops, and the dynamic behavior of systems, as well as the ability to predict the behavior of a system, and being able to think about modifications in order to achieve desired effects (Arnold & Wade, 2015). However, systems to foster geospatial relational thinking (Favier & van der Schee, 2014) or systems thinking in geography differ from other systems in three ways. A first, unique element in systems thinking in geography is the context as a localization of systems elements, or a possible geographical pattern, is of great importance in order to fully understand these systems. Second, in geography a connection is made between humans and natural systems (Hooghuis, van der Schee, van der Velde, Imants, & Volman, 2014; International Geographical Union, 2016). Third, scales are an important element in systems thinking in geography. Although systems are often looked at on a macro scale, e.g. international migration or global food problems, geographical systems often do include connections between different scales. The systems described in literature on Earth Sciences, e.g on the rock cycle (Kali, Orion, & Eylon, 2003), include a certain complexity, but often miss one or more from these geographical characteristics.

### Research on students' systems thinking level

Empirical studies on students' level of systems thinking in geography are scarce, but Favier and van der Schee (2014) examined, for example, the effect of geospatial technologies on the geospatial relational thinking of high school students. The students' level of geospatial relational thinking was rather disappointing before and after their intervention. These results are consistent with the findings of Karkdijk, van der Schee, and Admiraal (2013) who reported that students struggled with relating phenomena in Dutch national geography exams in 2009 and 2010. Furthermore, the field of Earth Sciences recognizes that higher order thinking takes on the characteristics of systems thinking (Batzri, Assaraf, Cohen, & Orion, 2015). Assaraf and Orion (2005) examined the development of systems thinking skills in the context of the hydro cycle. Their results show that students had an incomplete picture of the water cycle at the beginning of the learning process. Many students expressed difficulties, for instance, with respect to the basic skill of identifying the system components. These results correspond to the findings of Kali et al. (2003) who also observed a low general awareness of the rock cycle's dynamic nature before the intervention.

In the field of system dynamics, insight in some of the underlying principles has been researched. Moxnes (2000) found that participants had systematic misconceptions of stock and flows and of nonlinearities. In his experiment about managing renewable resources, the participants over-invested and over-utilized their resources. A study of Sterman and Sweeney (2002) examined the understanding of the most basic stock and flow structures governing the climate. Highly educated subjects scored poorly and consistently underestimated the delay in the response of temperature to changes in $CO_2$ concentration. In other studies with simpler tasks as filling a bathtub, participants showed a poor understanding of stock and flow relationships and time delay (Sweeney & Sterman, 2000).

Finally, Hmelo-Silver et al. (2017) found an overall pattern in the pre-test of their intervention study. Students were able to identify certain components in the ecosystem, but they failed to make connections with other components.

### Research on determinants of students' systems thinking level

Research on which elements influence students' systems thinking is rare. Sweeney and Sterman (2000) did not find a consistent pattern of influencing factors in their empirical study on the understanding of stock and flow structures. Subjects with a technical background outperformed those with a background in social sciences for only one of the tasks in their test. The degree program in which the students were enrolled had no significant effect, neither did their age or English as their first language. In other studies, there is an emphasis on the involvement of a considerable cognitive load while reasoning on complex systems. Also, the need to have sufficient domain knowledge and scientific reasoning skills such as hypothesis generation, experimentation, data collection, data analysis and communication of results are seen as important (Grotzer, Solis, Tutwiler, & Cuzzolino, 2017; Hmelo-Silver & Azevedo, 2006).

Regarding the influence of students' grade level, a correlation between progression in geospatial thinking and students' grade level was found in the study on geospatial thinking expertise of Huynh and Sharpe (2013). A study on conceptual thinking in geography revealed a similar correlation between learning gains and grade level (Trygestad, 1997).

Lastly, it is unclear whether gender affects systems thinking. Sweeney and Sterman (2000) found a marginally significant effect with males performing better than females. The studies of Favier and van der Schee (2014) and Karkdijk et al. (2013) observed no significant difference between boys and girls. Gerstner and Bogner (2009) observed females showing a deeper understanding of systems' complexity compared to their male colleagues. The study of Martin, Mintzes, and Clavijo (2000) showed the same significant difference favoring females in a biology context. Nevertheless, the effects of gender on science learning are complex and also related to attitude, motivation and self-efficacy.

## Hypotheses

Based on the meager results found in the literature described above, the hypothesis is that the level of systems thinking in geography of high school students in Flanders will be rather low. We expect to find differences across grades and the success rates in higher education of the study programs followed by students. Although the studies mentioned also found poor results for highly educated adults, we hypothesize that stronger students in upper secondary education (= students in an education program with a superior success rate in higher education) will perform better in systems thinking. Furthermore, it is difficult to predict gender differences. Based on former studies in systems dynamics and geospatial thinking, substantial differences are not expected.

## Methods

### *Description of the test used to measure the student's systems thinking ability*

The authors constructed a paper-and-pencil test based on a further operationalization of the systems thinking's definition formulated by Arnold and Wade (2015): "Systems thinking is a set of synergetic, analytical skills used to improve the capability of identifying and understanding systems, predicting their behavior, and devising modifications to them in order to produce desired effects" (Arnold & Wade, 2015, p. 675). In the operationalization process, these skills were translated into a measurable product: A person who is apt at systems thinking should have the following three skills: (1) constructing a causal diagram based on the information of a given source. This 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) describing relations between variables in words; (3) if there is interference within a system, explaining its influence on this system.

The test itself consists of four items (see attachment). The content of the test can be denominated as geography, but the topics used are usually not taught in Flemish high schools. In order to answer the first two items, the students have to read a text about the influence of climate changes on migration flows from Syria. In the first item, students construct a causal diagram with given variables. Hence, students have to recognize the relations between different variables in the text and assign the nature of the relation. In the second item, students are asked to describe in their own words a relation between two of these given variables. For the third and fourth item, a text and accompanying map are provided. This text is about the influence of air plane travel on jobs in the country of destination and the climate. The map shows which regions will be mostly affected by climate change. Based on the information in both sources, students have to identify relevant variables in the third item and have to use them in the fourth item to construct a causal diagram. A question on which the diagram in items 1 and 4 should be an answer, guides the students in their reasoning process.

Apart from the general structure, it is important to mention which elements regarding systems thinking, geography, and type of task are included in the test (Table 1). This will enable us to answer the second research question. In items 1, 2 and 4 interconnections have to be drawn or described explicitly. The fact that many variables and their interconnections are taken into account in items 1, 3 and 4 refers to the idea of seeing the broader picture rather than focusing on parts. In the systems drawn in items 1 and 4, feedback

**Table 1.** The different elements regarding systems thinking in general, systems thinking in geography, and types of task present. It shows which item in the test includes which element.

| Element | Item 1 | Item 2 | Item 3 | Item 4 |
|---|---|---|---|---|
| Systems thinking in general | | | | |
| Interconnections | x | x | – | x |
| Wholes rather than parts | x | – | x | x |
| Feedback loops | x | – | – | x |
| Dynamic behaviour | – | – | – | – |
| Systems thinking in geography | | | | |
| Identify variables on different spatial scales | – | x | x | x |
| Interconnections between different spatial scales | x | x | – | x |
| Interconnections between human and environment | x | – | – | x |
| Interconnections between different time scales | – | – | – | – |
| Type of task | | | | |
| Draw interconnections in diagram | x | – | – | x |
| Identify variables | – | – | x | – |
| Describe interconnections in words | – | x | – | – |
| Extract information from a combination of sources | – | – | – | x |

loops are included. This means that the test contains three out of the four elements which reoccurred in Arnold and Wade's (2015) comparison of eight definitions of systems thinking. The fourth element, the dynamic behavior of the systems, however, is not explicitly taken into account in this measuring tool.

Second, from the point of view of geography, dealing with systems on multiple spatial and temporal scales and including the interaction between humans and their environment is of specific importance. In all the items in the test, variables of different spatial scales (local, regional and global) are taken into account and have to be identified in items 2, 3 and 4. The drawn interconnections between these variables in items 1, 2 and 4 also reflect the different scales. In the systems drawn by the students, items 1 and 4 include interconnections between variables in social sciences and variables in natural sciences. This is shown in item 1 by the interconnections between the drought in a region and the effect on poverty. The effect of air travel on the emissions of greenhouse gasses and their effect on the earth's temperature in item 4 is also a good example. Lastly, the table and the explanation above, shows the different types of tasks students have to carry out.

In the design phase of the test instrument, several actions were undertaken regarding the validity of the test. An expert review by a professor in geography, a pilot study, at first with two and later with ten extra students, and a group discussion with teacher trainers in geography all led to several revisions before the test could be considered valid. To measure the internal consistency of the test, Cronbach's alpha was calculated after scoring all the tests, which yielded a result of 0.76. Bearing in mind the limited number of items and a not complete unidimensional definition of systems thinking, this value of Cronbach's alpha is acceptable (Tolmie, Muijs, & McAteer, 2011).

### *Sample*

The students in the sample were selected in class groups by teachers who participated in the research. These teachers were recruited by spreading a call to geography teachers who had participated at least once in professional development activities organized by the university and via the association of geography teachers in Flanders. In total, 18 teachers of 17 different schools, spread across Flanders, were willing to participate in the study with one or more of their class groups. This led to 735 students in their last ($n$ = 213) or penultimate ($n$ = 522) year of high school (age 16–18) who completed the test. Of the participants, 362 were male and 373 were female. The students were enrolled in different degree programs, which were categorized into groups of stronger and weaker programs. Strength of a program was based on the students' success rate in higher education between 2005 and 2010 for the whole of Flanders. Based on a visual analysis and a cluster analysis with the number of students that graduated in a bachelor program with professional orientation, a bachelor program with academic orientation and the participation level in higher education, four categories could be distinguished. In the first category, 68%–90% of students had successfully completed a bachelor program with academic orientation. In the second category, 69%–85% of students had graduated in a bachelor program, with an academic or a professional orientation. In the third category, 48%-67% of students had successfully completed a bachelor program with a professional orientation, but barely any student had graduated in an academically orientated bachelor program. In the fourth category, the average participation in higher education is substantially lower compared to the other categories. In this category students have a relative low success rate (<45%) in a professionally orientated bachelor program. In the sample, 257 students were in the first category, 189 in the second category, 247 in the third category and 42 in the fourth category.

### *Data gathering, scoring and analysis*

The tests and practical guidelines were transmitted to the teachers, who ensured that students took the tests within a timeframe of 30 minutes. Afterwards, the tests were scored using an ideal answer model. Students could obtain a total score on the test between 0 and 1. This interval reflects the range in systems thinking abilities from seriously lacking to fully competent. The total score on the test was the unweighted sum of the four items, thus considering the different items as equally important. The detailed scoring guide is available on request with the authors. Twenty randomly selected tests were scored by two independent raters to check the inter-rater reliability. To allow for coincidental agreement, a Cohen's kappa was calculated for each item. The Cohen's kappa was 0.95 for the first item, 1 for the second item, 0.78 for the third item and 0.74 for the last item. All these values can be considered as substantial to almost perfect according to the guidelines of Landis and Koch (1977). The remaining tests were therefore scored by one researcher.

The obtained scores were analyzed in the statistical program IBM SPSS Statistics 24. First, descriptive statistics of the overall score on the tests were explored. Second, a univariate analysis of variance was completed. Both the main and interaction effects of the independent factors gender, grade and degree program were present in the model to look at the effect on the dependent factor overall score on the test. Lastly, the results on separate items were visualized with software called Gephi version 0.9.1.

## Results and discussion

### *Total score on the test*

The descriptive statistics of the total score on the test show a frequency distribution approximating a normal distribution (Figure 1). The mean score is 0.52 with a standard

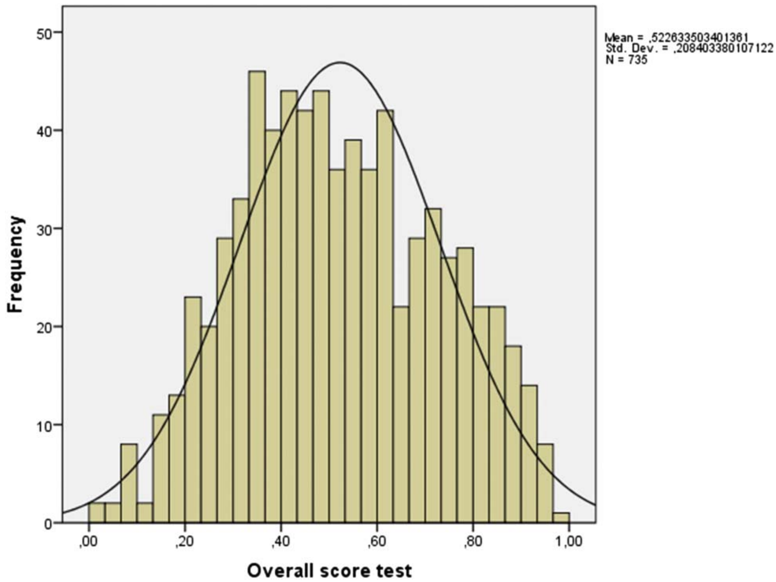


**Figure 1.** The frequency distribution of the overall scores on the test.

deviation of 0.21, indicating a wide range in students' systems thinking abilities. Thus, the test is successful in differentiating between levels in systems thinking. With a slight positive skewedness of 0.74, the distribution is slanted to the right. With a 50th percentile of 0.51, half the students did not collect half of the maximal achievable points. The 95th percentile is 0.88.

These overall scores on the test reveal great differences in the students' level of systems thinking in geography, but with a mean score of 0.52 for students at the end of secondary education, the results are not encouraging. Assuming that these results reflect the current situation, they are problematic, keeping in mind the future global challenges this student generation will be facing. Geography education aims to achieve a deeper insight in human-environmental interaction on several scales (International Geographical Union, 2016), but if students are not able to explain relations in a system with less than ten variables and with the content provided in a text, it can be questioned whether this goal is sufficiently met. This rather low level of systems thinking corresponds with earlier research on geospatial relational thinking (Favier & van der Schee, 2014), systems thinking in the Earth Sciences (Assaraf & Orion, 2005; Kali et al., 2003) and the understanding of stock and flow structures (Moxnes, 2000; Sterman & Sweeney, 2002; Sweeney & Sterman, 2000). Also the findings in the context of the structure–behavior–function model indicating causality in behavior as the hardest aspect of systems to understand (Chi, Feltovich, & Glaser, 1981; Hmelo et al., 2000), corresponds to the outcomes in our study. However, caution is required as it is impossible with the available data to conclude whether the poor results are a consequence of the current education in high schools in Flanders or whether the measuring tool used in this study is not appropriate to test students' systems thinking ability. The researchers therefore suggest to use the test instrument in a different student population outside Flanders, preferably first in countries where geography courses adopt

**Table 2.** The mean scores and standard deviation on the different items in the test.

| | Item 1 | Item 2 | Item 3 | Item 4 |
|---|---|---|---|---|
| Mean | 0.51 | 0.58 | 0.66 | 0.32 |
| Standard deviation | 0.25 | 0.34 | 0.24 | 0.24 |

an integrated approach like in Flanders, or that the same population is tested with another measuring tool.

### *Experienced difficulties in different aspects of systems thinking in geography*

In order to better understand this worrying overall result, students' responses are further discussed.

In Table 2, the average score (max = 1) and standard deviation for each item is given. Interconnections had to be drawn in items 1, 2 and 4. In item 3, variables which are used to draw relations in item 4 had to be identified. Therefore, the response to item 4 is dependent on the response given to item 3. With 0.32 on 1, students score the lowest on item 4. On item 1, where students already received the variables, the mean score is only 0.51.

The second element of systems thinking, being "the whole rather than parts", is partly reflected in the number of variables that has to be taken into account to draw the connections. In item 2, the relation between four variables had to be explained, while the relations between nine and seven variables had to be drawn in respectively items 1 and 4. Based on the mean scores on the different items, there is no clear relationship between the number of variables and the score achieved on the item. However, students score highest on the item with the least connections. This might indicate the importance of the number of variables used while fostering systems thinking in geography education, but the unclear pattern also prevents us to conclude whether the different tasks (describing in words versus drawing connections) or the different number of variables included caused a divergence in scores on the items.

The third element of systems thinking is insight in feedback loops. In the systems in both items 1 and 4, at least one feedback loop had to be drawn. In Table 3, the exact percentage for each correct relation in item 1 is given, together with the signs that students have assigned to the relation. The sign is the symbol that students add to the arrow to

**Table 3.** The percentage of students that drew correct relations in the first item, and the percentage of these students that assigned a minus, plus, no, or other sign to the drawn relations. The correct sign is shown in bold. The relations mentioned here are also indicated in Figure 2.

| | | Classification of the drawn relationship | | | |
|---|---|---|---|---|---|
| Name relation | % of students that draw this relation | Minus sign (−) | Plus sign (+) | No sign | Other signs |
| R1 | 90.2% | 5.3% | **76.0%** | 17.0% | 1.7% |
| R2 | 79.9% | 4.1% | **81.3%** | 11.6% | 3.1% |
| R3 | 44.1% | 5.2% | **75.9%** | 18.2% | 0.6% |
| R4 | 10.3% | 2.6% | **76.3%** | 17.1% | 3.9% |
| R5 | 34.6% | 2.4% | **75.6%** | 20.9% | 1.2% |
| R6 | 66.7% | 3.5% | **77.1%** | 17.1% | 2.2% |
| R7 | 74.8% | 4.2% | **66.5%** | 27.1% | 2.2% |
| R8 | 70.6% | 4.0% | **74.2%** | 20.4% | 1.3% |
| R9 | 78.5% | 1.4% | **76.6%** | 18.7% | 3.3% |
| R10 | 53.9% | 8.6% | **75.3%** | 15.2% | 1.0% |

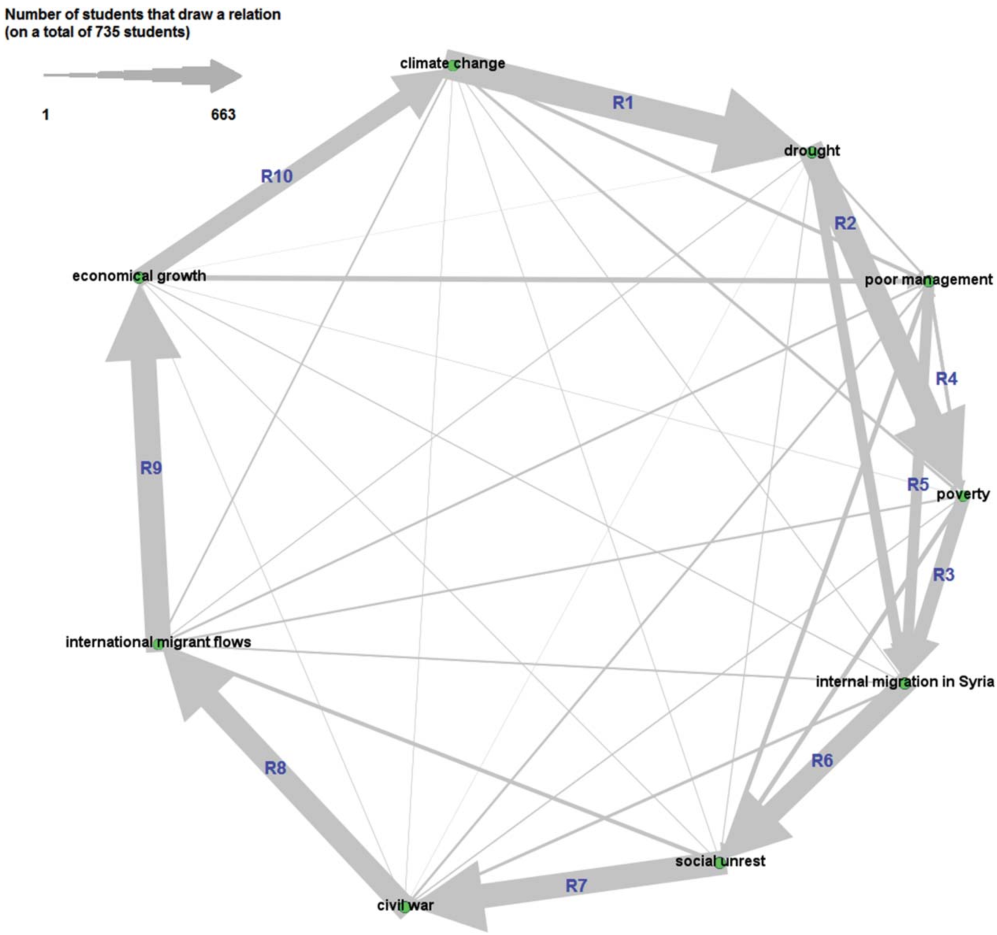


**Figure 2.** Relations drawn by students as response to the first item in the test. The arrows' thickness represents the number of students that effectively draw each relation.



show whether the variables have an increasing (+) or decreasing (−) effect on each other. The students sometimes added words instead of symbols or used a plus and minus sign together or used an equation sign (=), etc. All these are included here in the category "other signs". Table 3 indicates that relations 1, 2, 7, 8 and 9 are all drawn by more than 70% of students, but that relations 4 and 5 were hard to identify for most students. Respectively, only 10.3% and 34.6% of students drew these relations. Those two relations are lying outside a bigger feedback loop. The structure of the system in this case may have been confusing for students (Figure 2). Students who have drawn the relation are mostly able to add the correct sign to it.

In the fourth item, students drew fewer relations overall compared to the first item. As item 4 is dependent on the variables found in item 3, Figure 3 shows the percentage of students that drew each relation correctly in item 4 of those that had also found all the variables and thus had a maximum score on item 3 ($n$ = 105). It can be observed that relations 7 and 8, compared to the other relations, are drawn much less by students. Remarkably, connections that are drawn much less in item 4, also by students who had found all the

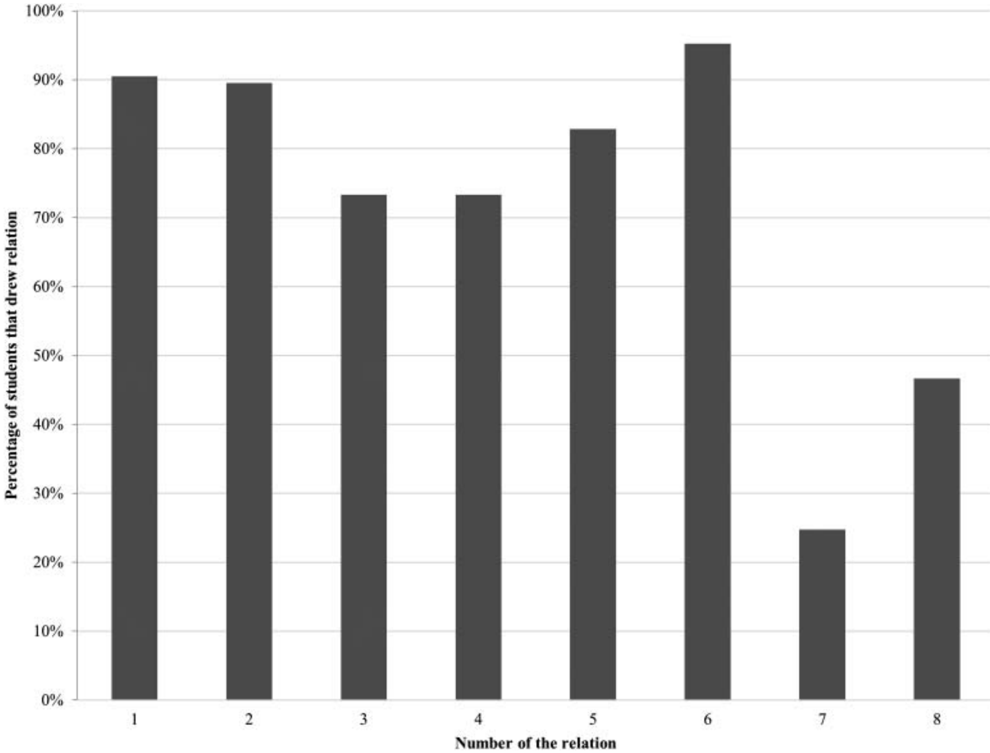


**Figure 3.** Percentage of students that drew a correct relation between variables. Only those students who found all variables are taken into account (105 out of 735).



necessary variables in item 3, have two specific characteristics (Figure 4). First, these are the only two feedback loops in the system. Second, these are the only two connections of which the information had to be extracted from the accompanying map instead of from the text. Interpreting and synthesizing information from different types of sources can be considered as part of scientific reasoning skills. From these results, it is impossible to deduct why students do not take into account the information presented on the map. But if systems thinking in geography is considered, it is important to mention that students should be able to take into account information presented in maps. In addition, the fact that these two connections were feedback loops creates an extra difficulty since students are more prone to linear thinking (Hmelo-Silver et al., 2017).

Concerning the first important element in geography, namely the variables on different spatial scales, it is interesting to look at item 3, in which students had to identify the variables themselves. From the seven variables, four are global, namely "air plane travel", "$CO_2$ emissions", "earth temperature" and "globalization", while the other three are local, namely "welfare country of origin", "jobs in country of destination" and "welfare country of destination". Figure 5 shows the percentage of students that identified the variables. It is remarkable that the variables "welfare country of origin" and "welfare country of destination" are only identified by respectively 50% and 43% of students. The distinction between country of origin and destination was necessary to create a correct causal diagram. Therefore, the word "welfare" on itself was not coded as correct. Although students

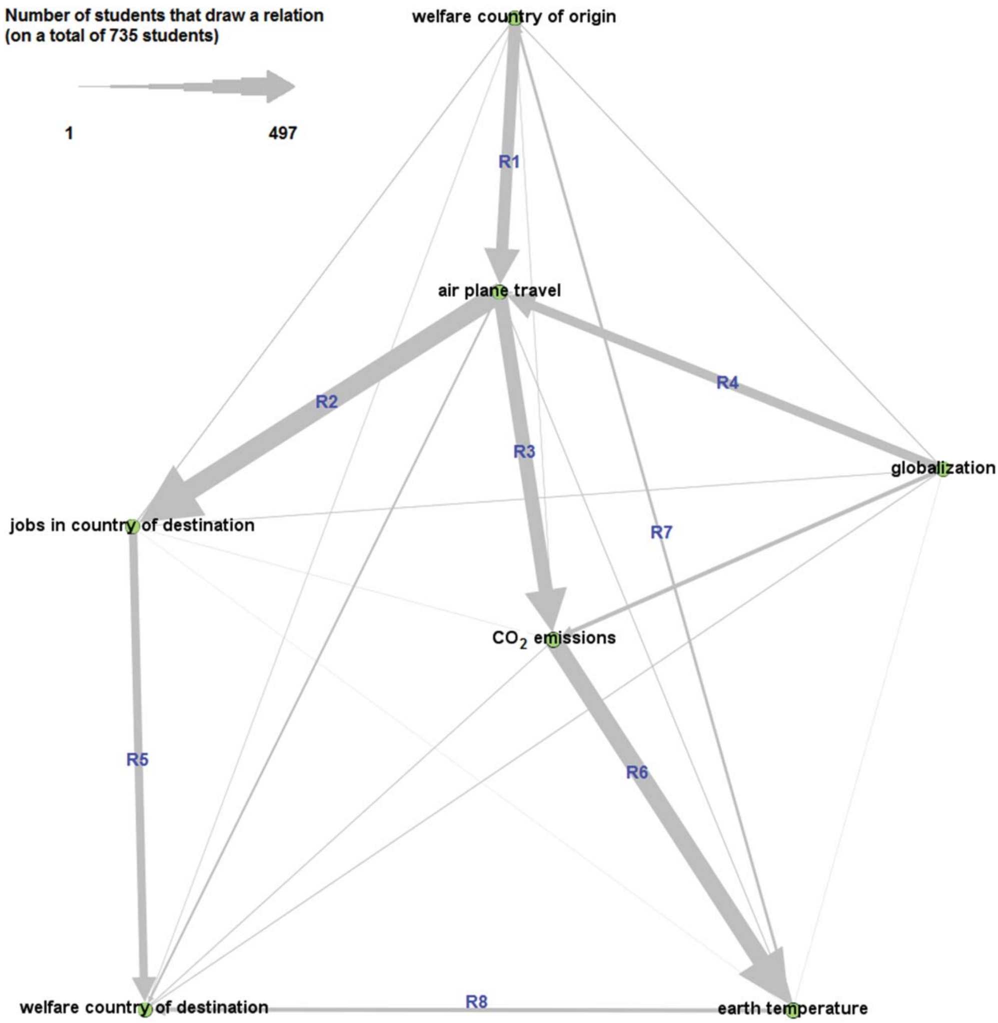


**Figure 4.** Relations of students drawn as a response on item 4. The arrows' thickness represents the number of students that effectively draw each relation.



were asked to look at the effect of both country of origin and destination, about half of them did not make a distinction in their causal diagram. On the contrary, more than 80% of students identified the other local variable "jobs in country of destination". A second element important in geography is the relation between variables on different scales. In the first item, these are reflected in relations 1 and 10 (Figure 2). In the fourth item, these are reflected in relations 1, 2, 7 and 8 (Figure 4). However, no clear pattern is observed regarding students drawing these relations more or less compared to the other relations. The third element in geography is the relation between variables from social sciences and variables from natural sciences. This is reflected in relations 2 and 10 in the first item (Figure 2) and in relations 6, 7 and 8 in the fourth item (Figure 4). Here also, no clear patterns are observable.

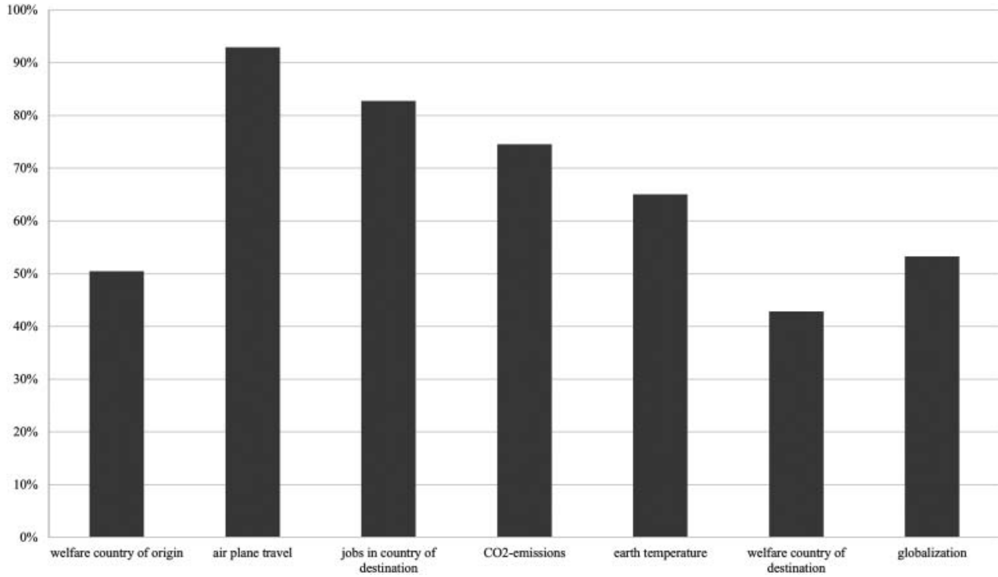


**Figure 5.** Percentage of students who identified the variables in item 3.

### *Influencing factors on student's systems thinking level*

The results of the analysis of variance in Table 4 show a significant main effect for success rate in higher education ($F(3, 722) = 105.515$, $p < 0.05$). The group of students in stronger study programs has a higher mean score on the systems thinking test in secondary school. Students with a high chance to achieve an academic oriented bachelor degree have the highest mean score of 0.66, followed by a mean score of 0.52 for students who will have a rather great chance on graduation in either an academic or professional oriented bachelor. The mean score for students with a rather great chance on achieving a professional oriented bachelor degree and students with a lower chance on a professional oriented bachelor degree does not significantly differ with, respectively, 0.40 and 0.39. Students' grade is a second main effect ($F(1, 722) = 6.597$, $p < 0.05$) as students from 12th grade performed better than those from 11th grade. No main effect for gender was found in the analysis of variance.

The analysis of variance also shows a significant interaction effect of gender and grade ($F(1, 722) = 8.831$, $p < 0.05$). In Grade 11, female students scored better on the test compared to their male colleagues. But while female students from 12th grade did not perform better than female students from 11th grade, male students from 12th grade did score

**Table 4.** The influence of success rate in higher education, gender, and grade level on the overall test score. The significant variables at $p < 0.05$ are indicated in bold.

| Variables | DF | F | P |
|---|---|---|---|
| Success rate in HE | 3 | 105.515 | **0.001** |
| Gender | 1 | 3.297 | 0.070 |
| Grade | 1 | 6.597 | **0.010** |
| Success rate in HE* Gender | 3 | 0.279 | 0.840 |
| Success rate in HE* Grade | 3 | 2.219 | 0.085 |
| Gender* Grade | 1 | 8.831 | **0.003** |

Model DF = 12; Error DF = 722; Total DF = 735; Model F = 30.550; $R^2 = 0.337$

better than male students from 11th grade and also slightly better than their female colleagues. However, with these factors taken into account, only 34% of the variance can be explained.

In contrast to the limited effect of study program found by Sweeney and Sterman (2000), the results in our study show a significant difference regarding stronger and weaker programs. The authors are not surprised that students in stronger programs and hence with a high chance to achieve an academic oriented bachelor degree have the highest mean score of the four groups, but the mean score of this highest scoring group is only 0.66 on a maximum of 1 which is concerning. It is unclear whether these students further develop their systems thinking abilities in higher education or whether they graduate in higher education without being able to see the bigger picture and understanding connections in complex systems. But the results do show a better mean score for students from 12th grade compared to students from 11th grade. Taken into account these differences between students in different study programs and grade level in interaction with gender, a differentiated approach in classrooms might be appropriate to foster systems thinking in geography.

## Conclusion and further research

The explorative analysis of responses shows that students have many difficulties recognizing relationships if several elements of systems thinking come together such as feedback loops, interactions between human and physical environment, and a combination of different information sources. However, rather great differences were found when looking at students' study background.

These findings contribute to geography education in at least two ways. First, it can raise awareness among teachers about the difficulties students experience to understand relations between different elements in geographical issues. It cannot be taken for granted that students themselves can establish the relationships between these elements. Second, it shows the need to develop appropriate teaching strategies. These teaching and learning strategies can differ according to the students' study background, but the authors suggest treating relations in an explicit way in the classroom as a first step. Furthermore, the authors point to the importance to visualize the relations in a system. Different tools might be appropriate such as concept maps, computer simulations or causal loop diagrams. Apart from a visualization of the complexity which is helpful for the students to understand the bigger picture, the activity to construct these maps, simulations or diagrams in group can create discussion among students which helps them to reason. Further research is necessary to examine these suggested teaching strategies in geography education. More in particular the effectiveness of visualization tools to help students understanding the bigger picture and reasoning about it deserves research attention. A first attempt to design lessons with these tools is done in Cox, Steegen, and Elen (in press). Moreover, research can help to understand how many variables are appropriate in a system to teach students about at certain ages.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This research was supported by the KU Leuven.

## ORCID

*Jan Elen* http://orcid.org/0000-0003-1611-5075

## References

Arnold, R. D., & Wade, J. P. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science, 44*, 669–678. http://doi.org/10.1016/j.procs.2015.03.050

Assaraf, O. B.-Z., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching, 42*(5), 518–560. http://doi.org/10.1002/tea.20061

Batzri, O., Assaraf, O. B.-Z., Cohen, C., & Orion, N. (2015). Understanding the earth systems: Expressions of dynamic and cyclic thinking among university students. *Journal of Science Education and Technology, 24*(6), 761–775. http://doi.org/10.1007/s10956-015-9562-8

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121–152. http://doi.org/10.1207/s15516709cog0502_2

Cox, M., Steegen, A., & Elen, J. (in press). Using causal diagrams to foster systems thinking in geography education. *International Journal of Designs for Learning*.

Favier, T. T., & van der Schee, J. A. (2014). The effects of geography lessons with geospatial technologies on the development of high school students' relational thinking. *Computers and Education, 76*, 225–236. http://doi.org/10.1016/j.compedu.2014.04.004

Forrester, J. W. (2007). System dynamics – a personal view of the first fifty years. *System Dynamics Review, 23*(2/3), 345–358. http://doi.org/10.1002/sdr.382

Gerstner, S., & Bogner, F. X. (2009). Concept map structure, gender and teaching methods: An investigation of students' science learning. *Educational Research, 51*(4), 425–438. http://doi.org/10.1080/00131880903354758

Grotzer, T. A., Solis, S. L., Tutwiler, M. S., & Cuzzolino, M. P. (2017). A study of students' reasoning about probabilistic causality: Implications for understanding complex systems and for instructional design. *Instructional Science, 45*(1), 25–52. http://doi.org/10.1007/s11251-016-9389-6

Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *Journal of the Learning Sciences, 9*(3), 247–298. http://doi.org/10.1207/S15327809JLS0903

Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *The Journal of the Learning Sciences, 15*(1), 53–61. http://doi.org/10.1207/s15327809jls1501_7

Hmelo-Silver, C. E., Jordan, R., Eberbach, C., & Sinha, S. (2017). Systems learning with a conceptual representation: A quasi-experimental study. *Instructional Science, 45*(1), 53–72. http://doi.org/10.1007/s11251-016-9392-y

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences, 16*(3), 307–331. http://doi.org/10.1080/10508400701413401

Hooghuis, F., van der Schee, J., van der Velde, M., Imants, J., & Volman, M. (2014). The adoption of thinking through geography strategies and their impact on teaching geographical reasoning in Dutch secondary schools. *International Research in Geographical and Environmental Education, 23*(3), 242–258. http://doi.org/10.1080/10382046.2014.927168

Huynh, N. T., & Sharpe, B. (2013). An assessment instrument to measure geospatial thinking expertise. *Journal of Geography, 112*, 3–17. http://doi.org/10.1080/00221341.2012.682227

International Geographical Union. (2016). International charter on geographical education. Retrieved from http://www.igu-cge.org/Charters-pdf/2016/IGU_2016_def.pdf

Jackson, P. A. (2006). Thinking geographically. *Geography, 91*(3), 199–204. http://doi.org/10.1016/j.polgeo.2004.04.005

Kali, Y., Orion, N., & Eylon, B. S. (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), 545–565. http://doi.org/10.1002/tea.10096

Karkdijk, J., van der Schee, J., & Admiraal, W. (2013). Effects of teaching with mysteries on students' geographical thinking skills. *International Research in Geographical and Environmental Education, 22*(3), 183–190. http://doi.org/10.1080/10382046.2013.817664

Katholiek Onderwijs Vlaanderen. (2017). *Leerplan secundair onderwijs: Aardrijkskunde derde graad tso/kso*. Brussels: Author

Landis, J. R., & Koch, G. G. (1977). The Measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. http://doi.org/10.2307/2529310

Martin, B. L., Mintzes, J. J., & Clavijo, I. E. (2000). Restructuring knowledge in Biology: Cognitive processes and metacognitive reflections. *International Journal of Science Education, 22*(3), 303–323. http://doi.org/10.1080/095006900289895

Maude, A. (2017). Geography and powerful knowledge: A contribution to the debate. *International Research in Geographical and Environmental Education, 2046*, 1–12. http://doi.org/10.1080/10382046.2017.1320899

Moxnes, E. (2000). Not only the tragedy of the commons: Misperceptions of feedback and policies for sustainable development. *System Dynamics Review, 16*(4), 325–348. http://doi.org/10.1002/sdr.201

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

Richardson, G. P. (2014). "Model" teaching. *System Dynamics Review, 30*(1–2), 81–88. http://doi.org/10.1002/sdr.1512

Richmond, B. (1993). Systems thinking: Critical thinking skills for the 1990s and beyond. *System Dynamics Review, 9*(2), 113–133. http://doi.org/10.1002/sdr.4260090203

Senge, P. M. (2010). Education for an interdependent world: Developing systems citizens. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (pp. 131–151). Dordrecht: Springer Netherlands.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23. http://doi.org/10.1007/SpringerReference_17273

Smetana, L. K., & Bell, R. L. (2012). Computer simulations to support science instruction and learning : A critical review of the literature. *International Journal of Science Education, 34*(9), 1337–1370. http://doi.org/10.1080/09500693.2011.605182

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

Sweeney, L. B., & Sterman, J. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review, 16*(4), 249–286. http://doi.org/10.1002/sdr.198

Sweeney, L. B., & Sterman, J. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. *System Dynamics Review, 23*(2/3), 285–312. http://doi.org/10.1002/sdr.366

Tolmie, A., Muijs, D., & McAteer, E. (2011). *Quantitative methods in educational and social research using SPSS*. Berkshire: Open University Press McGraw-Hill Education.

Trygestad, J. (1997). Students' conceptual thinking in geography. Retrieved from http://eric.ed.gov/?id=ED422231

Wijnen, F. M., Mulder, Y. G., Alessi, S. M., & Bollen, L. (2015). The potential of learning from erroneous models: Comparing three types of model instruction. *System Dynamics Review, 31*(4), 250–270. http://doi.org/10.1002/sdr.1546

Yoon, S. A., & Hmelo-Silver, C. (2017). Introduction to special issue: Models and tools for systems learning and instruction. *Instructional Science*, 1–4. http://doi.org/10.1007/s11251-017-9404-6