Mysteries to support geographical relational thinking in secondary education

Karkdijk, Jan

Mysteries to support geographical relational thinking in secondary education

Jan Karkdijk

Mysteries to support geographical relational thinking in secondary education

Jan Karkdijk

This research project was supported by the Dutch Research Council (NWO) under grant number 023.001.046

Printed by: Quaeris Media BV, Goes/Breda, The Netherlands

Cover design: DMARC

ISBN: 978-94-643-7235-9

VRIJE UNIVERSITEIT

Mysteries to support geographical relational thinking in secondary education

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. C.M. van Praag,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op woensdag 1 december 2021 om 11.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door
Jan Karkdijk
geboren te Enschede

promotoren:	prof.dr. W.F. Admiraal
	prof.dr. M. Meeter
copromotor:	prof.dr. J.A. van der Schee
promotiecommissie:	prof.dr. M.E.J. Raijmakers
	prof.dr. T. Béneker
	prof.dr. F.J.J.M. Janssen
	dr. R.F. Grol
	dr. B. Andeweg

Table of contents

Chapter 1

Introduction

1.1 Rationale

Teaching geography as a secondary school subject is about teaching students how to make sense of our wonderful, diverse and dynamic world. It is teaching how people at specific locations, in different circumstances, try to make a living on this planet and how their interrelationships with people in other places and their interaction with the environment form places and landscapes (International Geographical Union IGU-CGE, 2016). This means that in order to gain more understanding of our world, students have to learn to think relationally in geography: looking for interrelationships between human activities in different places and for the interaction between man and nature (from the local to the global). “Geography, the world subject, tries to keep things whole, geographical thinking includes relating the local and the global, the near and far, the physical and the human, people and environments, the economic and the social, time and distance ... and so on” (Geographical Association GA, 2012, p. 1, 2). Relational thinking constitutes the heart of thinking geographically (Lambert, 2017). Being able to give integrated, holistic explanations is mentioned as a way of thinking that makes geographical knowledge powerful for young people (Maude, 2017). Geographical relational thinking contributes to making students responsible citizens who understand how society and individuals interact with nature and with others in near and far regions in our tightly interconnected world. This relational consciousness, this awareness of interdependence, will inspire young people to live sustainably and humbly, because so much in our own life is made possible by the sacrifices of others (Murre, 2019).

Although relational thinking is important in the school subject, the rather scarce evidence available reveals that it is difficult for students. This is not to be expected, because in geography lessons secondary and even primary school students are used to questions such as Why? and Why there?, the latter being a relational question. And relational thinking is also an important skill in the national geography exams in The Netherlands (Centrale Examencommissie Vaststelling Opgaven CEVO, 2007), so it might be expected that students would master this skill. However, analysis of the results of all students in Dutch higher general secondary education (havo) on the 2009 and 2010 national geography exams revealed that students had most difficulties with relational questions (Cito, 2009, 2010; Karkdijk & Van der Schee, 2012). A research study on the use of geospatial technologies on the relational thinking skills of students aged 14-15 years also found that these thinking skills were rather low: students recognised less than half of all relationships in the representations and used only some of the relevant factors. Students had also difficulties expressing relationships verbally and visually (Favier & Van der Schee, 2014a, 2014b). Cox, Elen, and Steegen found that the overall level of systems thinking in Belgian secondary geography education was "not encouraging". Students reached only half of the maximal achievable points on their systems thinking test (2017, p. 8).

In order to promote thinking in geography lessons, the Thinking Through Geography (TTG) program was launched in the late 1990s. The "probably (…) most powerful strategy" (Leat, 2001, p. 51) of the program is the Mystery, a thinking strategy focusing on causality. Students have to establish a "variety of causal relationships (…) underpinned by geographical concepts", in order to "explore possible explanations for an event" (Leat & Nichols, 2003, p. 9). Mysteries can be used (1) to foster reasoning and relating skills (Vankan & Van der Schee, 2004) and (2) as a diagnostic and formative instrument to assess and foster students' thinking in geography (Leat & Nichols, 2003). Although mysteries are "a powerful research tool" (Leat & Nichols, 2003, p. 9) and widely used in geography lessons, very limited research has been conducted to obtain more evidence on student's relational thinking in geography by using this strategy.

The importance in secondary geography of learning to think relationally, the difficulties students have with relational thinking, the scarce evidence available on students' relational thinking skills together with the almost complete absence of research in the strategy of the Mystery to assess and foster relational thinking in geography lessons, provided the rationale for this dissertation. The aim was to obtain more evidence on students geographical relational thinking, in order to gain more insight into possible ways that teachers could help their students to master this skill. Therefore, mysteries, "excellent for promoting peer discussions" (Weeden & Lambert, 2006, p. 12), were used to elicit and analyse students' relational thinking when working on a mystery in small groups, to analyse differences between small student groups in their relational thinking and to evaluate the effects of using mysteries on students' geographical relational thinking.

1.2 Relational thinking in secondary geography education

The prime goal of secondary education is knowledge building (Young, 2008). Hattie (2009) uses Bereiter's application of Popper's three worlds to build his model of learning. These three worlds of achievement are: surface knowledge of facts and ideas (the first world); deep learning, qualified by elaborative and relational thinking processes (the second world); and the world of "constructed or conceptual understanding" (p. 29). Learning is a "journey, from ideas to understanding to constructing and onwards. (…). When students can move from idea to ideas and then relate and elaborate on them we have learning" (p. 29). These three worlds correspond to Lambert's description of the three related parts of powerful geographical knowledge. Powerful knowledge refers to knowledge that is discipline-based, has to be acquired in schools, is different from everyday knowledge, offers better explanations and new ways of thinking, and gives intellectual power and a language to those that possess it to participate in public debate (Young, 2008, p.14). According to Lambert, powerful geographical knowledge consists of: "(1) the acquisition and development of deep descriptive and explanatory `world knowledge'; (2) the development of the relational thinking that underpins geographical thought; and (3) a propensity to apply analysis of alternative social, economic and environmental futures to particular place contexts" (2014, p. 9). Both Hattie and Lambert mention relational thinking processes as essential in secondary education, in order to analyse and interrelate facts and lower order concepts into coherent wholes, necessary for making sense of the world.

Relational thinking connects to holistic thinking and to systems thinking. All three constructs refer to thinking in interconnected wholes. Whereas holistic thinking focuses on the functioning of systems as a whole, relational thinking focuses on thinking in interwoven (multi-causal) relationships. Systems thinking encompasses holistic and relational thinking (Lezak & Thibodeaux, 2016). For example, the German Geographical Society DGfG takes the systems approach to study the relationships within and between the physical and human subsystems of the Earth system at different scales as can be seen in Figure 1.1 (Rempfler & Uphues, 2012).

In order to formulate a definition that describes how geographical relational thinking is understood in this dissertation, the Geographical Analysis Model of Van der Schee (2000, 2009) has been used (Figure 1.2). Van der Schee used the complicated Geographical Structure Model of Hoekveld (1971) to devise a workable model for education, in order to help teachers to structure the content of their geography lessons. In the model, the geographical key concept "place" is recognisable in the region (at different scales) as a unit of analysis. Relationships

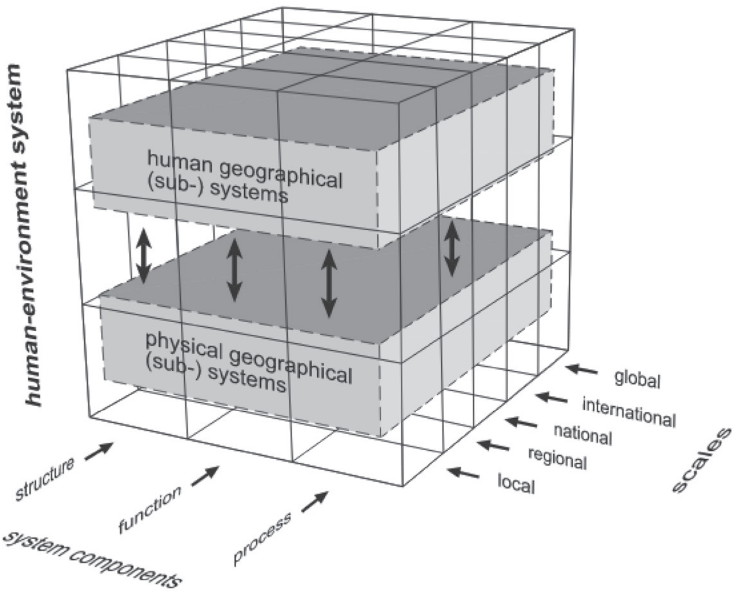


Figure 1.1 The cube model of the systems approach in German educational standards for geography. Source: DGfG, 2012, p. 11).

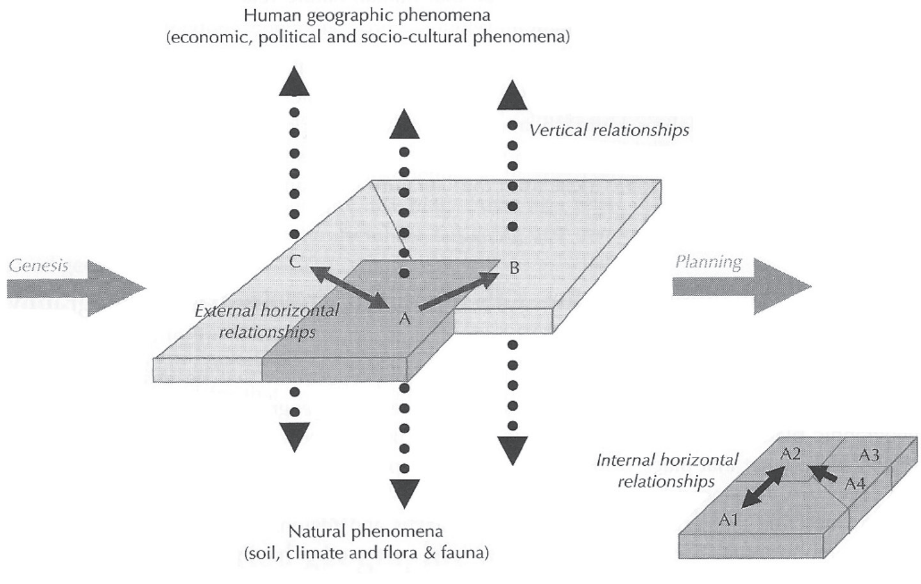


Figure 1.2 The Geographical Analysis Model (Source: Favier, 2011, p. 30).

within a region are called vertical relationships. These are interrelationships between man and environment within a region, between human phenomena and/or physical phenomena within a region. Geography's key concept "space" underpins the horizontal relationships in the model: relationships between regions. These can be analysed at different scale levels: between countries, but also between regions within a country and so on (the geographical key concept of "scale" is used here). As a result of the interaction between vertical and horizontal relationships, regions change continually. For example, the construction of the Belo Monte dam in the Amazon region in Brazil caused an influx of tens of thousands of migrants into the region of Altamira and the displacement of thousands of inhabitants of the region to Altamira city with rapid urbanisation of Altamira and accelerated deforestation as a result (horizontal relationships on two scale levels). Deforestation causes regional climate change (less precipitation), which decreases river discharge. These are examples of vertical relationships that change the region (Faiola, Lopes, & Mooney, 2019; Stickler et al., 2013).

Using Van der Schee's model, geographical relational thinking in this dissertation is defined as *analysing, explaining, evaluating and/or predicting the horizontal and vertical relationships and the interactions between them, on different scales, that cause regional change.* This implies that the research focus was on the ability to think in interrelated, multi-causal relationships. There are alternative interpretations of relational thinking in geography without this focus on causality, for example, from a more normative perspective as engagement with interdependencies or inequalities (Varró & Van Gorp, 2020, p. 3). However, in Dutch secondary education the model of Van der Schee is used to specify relational thinking (CEVO, 2007), so we used this interpretation. This geographical relational thinking is a kind of higher order thinking, for most of the mentioned thinking processes in the definition belong to the higher order cognitive process dimensions ("analysing", "evaluating") of the revised Bloom Taxonomy (Anderson & Krathwohl, 2001).

The central concept of this dissertation is *geographical* relational thinking, which is strongly related to *geospatial* relational thinking (Favier & Van der Schee, 2014a). Spatial thinking has been defined as "a set of abilities to visualise and interpret location, position, distance, direction, relationships, movement, and change through space" (Baker et al., 2015, p. 120) and consists of the abilities: spatial visualisation, spatial orientation and spatial relations (Huynh & Sharpe, 2013; Lee & Bednarz, 2009). Geospatial thinking is a specialised form of spatial thinking, bounded by a location on planet Earth, because in geo sciences location matters (Baker et al., 2015; Favier & Van der Schee 2014a). Favier and Van der Schee (2014a) argue that the subcomponent geospatial relational thinking should be distinguished, to stress the geospatial relational reasoning processes that are demanded here. Whereas geospatial relational thinking makes use of a lot of digital technologies like satellite images, digital maps, computer games or GIS, in this dissertation the emphasis is on geographical thinking, for the focus is on meaning-making with the help of geographical concepts and questions (Van der Schee, Trimp, Béneker, & Favier, 2015; Uhlenwinkel, 2013, 2017). This is why we used the concept of *geographical* relational thinking.

1.3 Mysteries and relational thinking

The mystery was devised as one of the strategies of the Thinking Through Geography (TTG) program. The TTG program was developed to infuse generic thinking skills into geography lessons like categorising, hypothesising, analysing, relating, and reasoning. (Leat, 2001; Moseley et al., 2005; Morgan, 2017). The TTG program is intended to teach thinking, in order to make geography lessons more stimulating and challenging for students, to promote understanding of geographical key concepts and to foster the intellectual development of students (Leat, 1997, p. 145). The program was inspired by cognitive acceleration projects, like the Cognitive Acceleration in Science Education (CASE). Adey and Shayer (1994) provide evidence for the significant gains in achievement the CASE project brought about, not only in science, but also in mathematics and English, even two or three years after the project had finished. Motivated by the CASE project and its design principles, TTG strategies have been developed within the framework of constructivism. The strategies make use of cognitive conflict to challenge students' existing knowledge, are organised as small group work where talk is crucial for student learning and are completed with a whole-class debriefing session. The debriefing is necessary to help students to understand what they have learned, not only of the specific geographical problem at hand and the geographical concepts involved, but also of their own thinking and reasoning (metacognition) in order to facilitate bridging of the learned thinking strategies to other contexts (Adey & Shayer, 1994; Leat, 1997, 2001).

The mystery as one of the TTG strategies is a card-sorting activity. In order to 'solve' (quotation marks by Leat and Nichols, 2003, p. 7) a problem formulated as a triggering, open question to create a cognitive conflict, students receive 16-30 strips containing information on the problem, but without causal relationships. In small group work students read the strips, move and categorise them to make groups and subgroups, establish linear or web-like causal relationships and write their "detailed and thoughtful explanations" (Leat, 2001, p. 51) as answers to the mystery question. The quotation marks around the verb solve and the answer to a mystery typified as "detailed and thoughtful explanations" indicate that the strategy is not a typical problem-solving activity. As quoted earlier, for tackling a mystery, Leat and Nichols (2003) use the expression "students explore possible explanations for an event" (p. 9). For this reason, it is difficult to position the strategy within one of the 11 types of problem solving in educational settings as discerned by Jonassen (2000), because mysteries are mostly used to explain the complexity of an event or phenomenon instead of finding a solution for a problem. Therefore, expressions like "solving a mystery" or "the solution of a mystery" (also in this dissertation) should be interpreted as explaining the mystery or the explanation of a mystery.

After the group work, some groups present their explanation in the debriefing session to the whole class, after which a teacher-guided whole class discussion follows. The design principles for a mystery are: (1) the number of information strips should be between the 16 and 30; (2) some strips should provide a storyline, a narrative thread of a person or people, in order to personalise the events and to capture the interest of students; (3) an open, central question is necessary in order to trigger students to think; (4) some strips have to provide abstract information, background and peripheral factors; (5) some strips should act as red herrings by giving misleading or irrelevant information; and (6) no causal relationships should be expressed on the strips (Leat, 2001; Leat & Nichols, 2003). Box 1.1 presents an example of a mystery describing the Rio mystery, one of the mysteries used in the research project. The individual information strips of this mystery can be found in Appendix B.

The mystery as a TTG thinking strategy is a powerful thinking strategy according to Leat (2001), because it includes a wide range of thinking activities such as analysing, classification, hypothesising and reasoning (Leat & Nichols, 2003, p. 9). The main thinking activity within the mystery concerns cause and effect (Leat, 2001, p. 51). Students have to establish a "variety of causal relationships (…) underpinned by geographical concepts" (Leat & Nichols, 2003, p. 9). The list of "big concepts" in TTG strategies includes: cause and effect, planning, decision making, location, classification, inequality, development, and systems (Leat, 2001, p. 161, 162). One could question whether these big (or key) concepts are really geographical. Except for the concept of location (place), the concepts are quite generic (cause and effect, development, systems, inequality) or actually skills (planning, decision making and classification). Not

Box 1.1 The Rio Mystery

The introduction of this mystery to the whole class starts with the damaging landslides in Rio de Janeiro in April 2010. In particular, favelas on steep slopes near and in the central city were severely hit. The government of Rio offered the favela residents new, safe dwellings on the outskirts of Rio. One of these residents was Fabio Pereira who, however, refused to move. After this whole-class introduction, the mystery question to be explained in small groups is presented: Why does Fabio Pereira want to stay in his threatened house in the favela? All groups receive the 28 information strips. Personal information is provided on Fabio Pereira (a father of four children, his house built by his father in a neighbourhood with residents who, like his parents, are migrants from North East Brazil). Other strips contain background information on the landslides and on the geography of Rio (distances in Rio, the proximity of Fabio's present house to the centre of Rio and high real estate prices in Rio). One third of the strips contain information on government policy towards the favelas (like projects to build luxury apartments in parts of the city where the government had also removed favelas and the start of the rehousing very shortly after the landslides without consulting the slum inhabitants). Four strips present information on upcoming big events in Rio (Olympics and World Cup), Rio as a tourist city and North East Brazil as a poor region with outmigration. A few strips contain irrelevant information (like closed schools in favelas hit by the landslides and numbers of inhabitants of favelas). The core inference students have to make from these information strips is Fabio's distrust of government intentions with regard to the rehousing plan, because his neighbourhood could easily be transformed into a more profitable part of the city. Other factors that explain his decision not to move are the great distance from the outskirts in Rio to the city centre (work), the community bonds in his present neighbourhood and the emotional fact that his house was built by his father.

surprisingly, mysteries are also used in history and economics. The relational thinking on cause and effect mentioned by Leat (2001) as the main thinking activity of the mystery, is therefore a more generic concept than geographical relational thinking in our definition. Leat and Nichols (2003) report that the outcomes of mysteries are not always strongly related to the subject of geography (p. 33). Therefore, the geography teacher has to design a geographical mystery by focusing on the region as the unit of analysis and the interaction of horizontal and vertical relationships that cause regional change. And it is also the teacher who has to focus in the debriefing session on geographical key concepts such as place, space and scale. In this way, mysteries are very useful to foster and diagnose the geographical relational thinking of students.

Research on the use of mysteries in geography lessons is limited and has covered different themes. Leat and Nichols (2000b, 2003) conducted research on the different, physically observable stages that reveal the thinking activity of groups in unravelling a mystery. Van der Schee, Leat, and Vankan (2006) investigated the combined effect of the use of Dutch adaptations of three TTG strategies (Five Ws, Reading Photographs and Mystery) on classification and relating skills of secondary school students and found a significant positive effect. Applis (2014) and Szymanski (2008) found that mysteries are suitable for dealing with complex issues within the framework of Global Learning and Value-Oriented Geography in Germany. Recently, Benninghaus, Mühling, Kremer, and Sprenger (2019) constructed a reference diagram in order to assess students' "influence diagrams" (concept maps) that resulted from the group work on a mystery. There is a lack of evidence on the effects of the use of mysteries on students' relating skills. It also remains unclear what the use of mysteries reveals about the quality of students' geographical relational thinking: what relationships students establish in their group work and what differences can be observed between small groups in established relationships. Evidence on these topics would not only evaluate the recommendations made for the mystery as a powerful strategy to diagnose and foster relational thinking, but also give teachers possible insights into strategies on how to foster relational thinking in students.

1.4 This dissertation

The research conducted for this dissertation can be qualified as teacher research. The researcher is a teacher who investigated the relational thinking skills of students within his own subject: geography. Besides many other students and teachers in secondary schools in the Netherlands, his own students also participated in the research project, guaranteeing a close connection with educational practice. Most teacher research consists of action research, lesson studies, self-studies and design-based research and is mostly small-scale, qualitative research and focused on the improvement of their own teaching practice. Teacher research rarely aims at extending the knowledge base on teaching and learning, precisely because it is mostly focused only on teachers' own teaching (Admiraal, Buijs, Claessens, Honing, & Karkdijk, 2017; Zwart, Smit, & Admiraal, 2015). The teacher research conducted for this dissertation aimed to extend the knowledge base on students' relational thinking framed within the group work on a mystery, in order to improve teaching and learning of geographical relational thinking in secondary education. This dissertation may contribute to bridging the gap between scientific educational research and teaching practice in secondary schools as described by Vanderlinde and Van Braak

(2010). They found that teachers often do not see the practical relevance of educational research for their daily teaching in class.

The main research question was:

How does the use of mysteries in secondary geography education support geographical relational thinking in students?

The use of "support" in this question implied that we used mysteries to foster geographical relational thinking skills (Vankan & Van der Schee, 2004) as well as to diagnose their geographical relational thinking (Leat & Nichols, 2003). The research question was approached by four studies, each with its own research questions. The focus was on the ability to think in interrelated, multi-causal relationships in order to give coherent explanations for a phenomenon or an event framed as a mystery question.

1.4.1 Study 1

Mysteries were used in the research project to provoke relational thinking, individually or within small groups. As the main concept addressed by this thinking strategy concerns cause and effect (Leat, 2001, p. 51), the first question to be answered before the next larger research project was launched, was whether mysteries *do* foster causal relational thinking. If no effect was found, it would still be possible to use mysteries as a framework for a follow-up study into students' geographical relational thinking, but a positive effect would provide an underpinning for the use of mysteries for this goal and also for recommending mysteries as a tool to foster geographical relational thinking. The first study aimed to find evidence for the effects of the use of mysteries on geographical relational thinking in terms of establishing causal relationships between human and environmental phenomena in developing countries. The guiding research question was: *(1) What is the effect of the use of mysteries on students' geographical relational thinking?*

This research was designed as a quasi-experimental study, with a pre-test-post-test control group design. Seven secondary schools in the Netherlands participated. At each school the teacher(s) assigned one class to the experimental condition (where teaching with three or four mysteries was added to the regular geography curriculum) and another class of the same educational level and year to the control condition (the same number of lessons, the regular geography curriculum, but without mysteries). A total of 221 students, aged 15-17, in upper higher general secondary education (havo) and pre-university education (vwo), participated: 105 in the experimental classes and 116 in the control classes.

The pre-test and post-test that were devised for this study focused on topics linked to the Millennium Goals, like poverty, demographics, illness, sustainability, environment and debts (see Appendix A). These topics are related to geography (especially of developing countries), but are not strictly geographical. For example, debts, the position of women or child mortality are also topics in other non-geographical subjects. On the other hand, these interrelated topics determine the living conditions of people in developing countries and can be understood within the Geographic Analysis Model (Van der Schee, 2000) as mainly vertical relationships and therefore as geographical relationships.

The pre-test and post-test each consisted of three assignments. In the instructions on the test, students were given an introduction to the Millennium Goals and the variables involved. Each assignment was to explain the relationship between two Millennium Goals variables by making use of other variables that were mentioned in the introduction. Students had to construct a relational scheme with cause-and-effect arrows indicating the relationships between the variables. Each established relationship had to be explained verbally. All correct relationships within an assignment were counted and a total score as the sum of all correct relationships of the three assignments was obtained for the pre-test and the post-test of each student.

1.4.2 Studies 2,3 and 4

In the next three studies the geographical relational thinking qualities of small groups working on a mystery were considered. The same database was used for these studies. Two geographical mysteries, designed in line with the design principles of Leat and Nichols (2003), acted as the learning activity to elicit relational thinking and reasoning. One mystery concerned the devastating landslides in Rio de Janeiro in 2010 and the resulting favela rehousing plans of the government and the other concerned the recurring floods in Jakarta. Information was gathered from newspapers, blogs and radio programs. Information on the Jakarta floods was also obtained from the dissertation of Van Voorst (2014). The central question, task and the information strips of both mysteries can be found in Appendix B.

The research was carried out in six schools in the Netherlands. In total 205 students aged 15-18 years in upper higher general secondary education (havo) and pre-university education (vwo), participated. The students collaborated in 69 small groups to understand and explain one of the two mysteries. Thirty-five groups worked on the Rio mystery and 34 on the Jakarta one. Students had to represent their explanations of the mystery as a concept map with labelled cause-and-effect arrows. Concept maps are a suitable format for organising and representing relational thinking and for provoking elaborative talk (Hwang, Yang, & Wang, 2013; Novak & Cañas, 2008; Srinivasan et al., 2008; Van Boxtel, Van der Linden, Roelofs, & Erkens, 2002). Research has found evidence for the construct-a-map technique, an open task where students have to construct the concept map with their own variables and structure, as the gold standard for revealing students' knowledge (Ruiz-Primo, Shavelson, Li, & Schultz, 2001; Ruiz-Primo, Schultz, Li, & Shavelson, 2001; Yin, Vanides, Ruiz-Primo, Ayala, & Shavelson, 2005). Concept maps and group discussions were used for analysis of established relationships and the coherence of these relationships. The coherence of the relationships, or the structure of their explanation, was assessed using the SOLO (Structure of Observed Learning Outcomes) taxonomy of Biggs and Collis (1982).

1.4.2.1 Study 2

This descriptive research was conducted to gain more insights into the quality of the geographical relational thinking of students working in small groups to explain a mystery. Evidence was sought for the relationships small groups established in their group work on a mystery and the coherence of their geographical relational thinking. The first research questions of this study was: (*2) Which geographical relationships do students in small groups establish to solve a mystery?*

Group discussions and concept maps were analysed for all correct and relevant relationships. For each mystery, a total concept map was constructed, representing all correct and relevant relationships groups established.

The second research question of this study was: (*3) How coherent are the solutions to the mystery posed?*

All correct and relevant relationships a group established were put into the standard format of the criterion map to make comparison between groups possible. This standard concept map was analysed in conjunction with the group's discussion to determine the appropriate level of the SOLO taxonomy of Biggs and Collis (1982), indicating the structural quality of the explanation.

In this second study, we used the verb "to solve" and "solution" in the meaning of Leat and Nichols (2003, p. 7, 9): "to explain" and "explanation".

1.4.2.2 Study 3

Where the second study summarised all correct and relevant relationships groups had established in a total concept map for each mystery, the third study focused on explaining differences between individual groups in their geographical relational thinking to explain a mystery. The first research question of this study was *(4) How does geographical relational thinking in terms of the SOLO taxonomy differ between groups?*

Differences between groups in geographical relational thinking were illustrated in the study by presenting the explanation of three groups working on the Rio mystery. The structure of the explanation of each group was indicated with the appropriate level of the SOLO taxonomy and all correct and relevant relationships a group had established were put into the format of the standard concept map of the mystery.

The second research question was: *(5) How does geographical relational thinking in groups differ between the two mysteries?*

The two mysteries were compared on three indicators of relational thinking: the total proposition score, the number of established cross-links and the number of factors used.

The third research question of this study was: *(6) How can differences between groups in geographical relational thinking be explained by characteristics and collaborative behaviour of the groups?*

Four multiple regression analyses were performed with the total proposition score, the number of cross-links, the number of factors and the SOLO levels of the groups used as dependent variables and five independent variables as predictors in the following sequence: educational level, school year, standardised mean geography grades of the groups, proportion of off-task words of all words uttered prior to the construction of the concept map and the number of on-task words uttered prior to the construction of the concept map.

1.4.2.3 Study 4

This study was a follow up to the third study on differences in geographical relational thinking between groups and comprised an in-depth qualitative analysis of the strategies groups employed to explain a mystery. This study was guided by the research question: *(7) What are the differences between low-performing and high-performing groups in their strategies to understand a geographical mystery?*

Twelve groups that worked on the Rio mystery were selected for analysis: six groups with the highest total proposition scores and six groups at the lowest end of the total proposition scores. The strategies groups employed were analysed using the stages groups progress through as observed by Leat and Nichols (2000b, 2003).

Table 1.1 summarises the main characteristics of the four studies. More detailed descriptions of the methods employed are presented in the Chapters 2 to 5, which comprise the four studies. Chapter 6 presents the main findings and the discussion of this dissertation.

Table 1.1 Overview of the main characteristics of the four studies.

	Study 1	Study 2	Study 3	Study 4
Research subject	Effect of mysteries on geographical relational thinking	Quality of geographical relational thinking	Differences between small groups in geographical relational thinking	Strategies of geographical relational thinking
Research question	1	2, 3	4, 5 & 6	7
Design	Quantitative, quasi-experimental	Quantitative, descriptive	Quantitative, correlational	Qualitative, explorative
Participants	7 secondary schools, 7 teachers in experimental condition, 14 classes with 221 students	6 secondary schools, 12 teachers, 69 small groups with 205 students		12 small groups with 36 students*
Unit of analysis	Individual students	Small student groups		

* These groups and students belonged to the participants of studies 2 and 3.

Chapter 2

Effects of teaching with mysteries on students' geographical thinking skills*

Abstract

Thinking Through Geography strategies are popular in secondary education. Geography teachers see these strategies to be powerful to stimulate thinking geographically. However, empirical evidence is scarce. Based on a quasi-experimental design, effects of mysteries, one of the more famous Thinking Through Geography strategies, were examined on students' geographical thinking in terms of their skills to relate phenomena. A multilevel regression analysis showed that students who used mysteries in geography lessons reported significantly more correct geographical relationships than students who attended the regular curriculum. We conclude that a mystery can be an effective TTG strategy for the development of students' geographical thinking skills. Improvements of the pedagogy of using mysteries in geography lessons are discussed.

*This chapter is based on: Karkdijk, J., Van der Schee, J. A., & Admiraal, W. F. (2013). Effects of teaching with mysteries on students' geographical thinking skills. *International Research in Geographical and Environmental Education*, *22*(3), 183-190.

2.1 Introduction

Geographical thinking is one of the most challenging aspects of both teaching and learning geography. Students should be enabled to learn these higher-order thinking skills and teachers need support on how to teach these skills. In the geography curriculum of Dutch secondary schools, students' thinking skills are fully addressed: students learn to relate phenomena and processes within and between regions. However, our analysis of the outcomes of the 2009 and 2010 Dutch national geography exams suggests that students had difficulties with linking phenomena and processes accurately. Yet almost 60% of the assignments of the exams asked for this skill. So, a more appropriate training of this important geographical skill seems to be needed.

In the current study, an intervention was examined that focused on triggering students' thinking skills in terms of relating phenomena and processes. The intervention included the use of mysteries, a challenging problem that students should solve with a card-sorting activity. The use of mysteries is one of the intriguing teaching approaches of Leat's well-known Thinking Through Geography program (Leat, 1997, 2001), although there is no empirical evidence for its effects.

2.2 Geographical thinking and mysteries

The reproduction of facts and concepts seems to be dominant in today's education in many countries. This is not a bad thing in itself but a necessary element in the learning process to come to deeper learning. "There needs to be a major shift from an over-reliance on surface information (the first world) and a misplaced assumption that the goal of education is deep understanding or development of thinking skills (the second world), towards a balance of surface and deep learning leading to students more successfully constructing defensible theories of knowing and reality (the third world)" (Hattie, 2009, p. 28). Creating cognitive maps with links between facts and concepts, theories and real-life, and different contexts seem to a promising way to enable meaningful learning. An influential initiative to promote this kind of meaningful learning and that has ensured continuing curriculum renewal in geography in many countries is Leat's Thinking Through Geography (Morgan & Lambert, 2005).

The Thinking Through Geography project uses information about meaningful learning collected in the thinking skills program of Adey and Shayer (1994) that showed that the

Cognitive Acceleration in Science Education project (CASE) improved results in science but also in other subjects like mathematics and English two or three years later. Leat's work focuses on strategies to motivate students to do geography and to help them to develop thinking skills for a better understanding of the world they live in. Hopwood (2009) states that pupils' conceptions of geography may be somewhat messy rather than fitting straightforwardly into neat and coherent categories. TTG is a teaching approach that can make learning more meaningful for students, improves understanding, helps to reveal misconceptions and helps teachers to reinterpret their subject matter. The TTG project has set three broad aims: (1) to devise adaptable strategies and curriculum materials that make geography lessons more stimulating and challenging; (2) to help students understand some fundamental concepts in geography in an explicit way so that these can be transferred to new contexts; and (3) to aid the intellectual development of students so that they can handle more complex information and achieve greater academic success (Leat & Nichols, 2000a; Nichols & Kinninment, 2001). TTG strategies are assumed to support the intellectual development of students by stimulating higher-order thinking skills, but empirical evidence is still scarce.

The use of mysteries is one of the teaching approaches in the TTG. Mysteries are perceived as excellent means to trigger a student to "understand and (…) explain of causes, processes and consequences" (Leat & Nichols, 2000b, p. 118). A mystery starts with a challenging question that triggers students to investigate the issue and solve the problem. This challenging question includes information that intuitively seems to be inconsistent with what the students already know. For example, one of the mysteries used in our research is about slum dwellers in Rio de Janeiro who built their houses on steep slopes around the Central Business District despite the great danger of frequently occurring landslides. Some dwellers refuse to move even when they are offered alternative housing elsewhere in Rio. Why do they prefer to live in danger instead of moving to a safer area? To solve the mystery, students get 16-30 pieces of information to help them answer the question. Students work in groups of two to four, investigate the issue and solve the mystery by producing an essay, picture, map or poster (see Figure 2.1). At the end of the mystery task a debriefing follows. In the debriefing teachers discuss various student products as well as their strategies to answer the question. The key features of mysteries are that:

(1) they are built around a natural or human phenomenon with a great impact, such as a tsunami, a landslide, a factory shutdown or an outbreak of a disease;

(2) the challenging question has to do with the real-life situation of someone to enhance students' imagination of the issue at stake;

(3) there are different correct ways to answer the question;

(4) students are provided with an overload of diverse and unconnected information. Information can be specific and detailed or general and abstract. So-called "red herrings" might also be included: information that seems to be useful, but is not necessary or even misleading;

(5) students rearrange collaboratively the 16-30 pieces of information, evoking shared and productive reasoning. Students have to clarify why they create sets of data, causal chains or other links. Students set up hypotheses about possible relationships between pieces of information with an argumentation to convince each other;

(6) group discussions are setup (Figure 2.1) to create a safe environment in which students with different abilities can feel confident enough to contribute.

Figure 2.1 Secondary school students at work with a mystery (Jan Karkdijk).

One of the key characteristics of the TTG is the phase of debriefing in order to enable students' transfer of thinking across subjects and situations (Van der Schee, Vankan, & Leat, 2006). Moseley et al. (2005) describe debriefing as the reflection of teacher and students on the processes and outcomes of the mystery task. According to Lambert and Balderstone (2010) this kind of debriefing in geography should be underpinned by two important principles: metacognition and bridging. They describe metacognition as an understanding of their own thinking that helps students interpreting different patterns of reasoning so that they can be applied to appropriate problems and situations, while bridging is the transfer of these new insights to other contexts in geography in order to generalise and consolidate learning.

So, mysteries can be used to engage students in learning and to support them in understanding complex issues through a combination of declarative knowledge (facts, concepts, generalisations), procedural knowledge (skills) and meta-cognitive knowledge (reflection). Although Leat's TTG strategies have been successfully introduced in the Netherlands and warmly welcomed by geography teachers, the more complex strategies like mysteries are scarcely used in geography lessons. Teachers prefer to use the simple and short strategies like "Odd one out" in their classes (Leat, Van der Schee, & Vankan, 2005). However, geography is essentially about relationships between human society and nature. So, to think geographically, students have to be able to formulate correct geographical relationships explaining the world around them (IGU-CGE, 1992). Being able to use relationships in explanations is also seen as an essential part of higher-order thinking skills by many authors (e.g., Biggs & Collis, 1982; Lipman, 1991; Stimpson, 1992). In this study, we want to provide an empirical basis for the effects of the use of mysteries on geographical thinking in terms of relating human and environmental phenomena. The provided empirical basis could foster teaching with mysteries in classrooms.

2.3 Method

Data were collected in seven secondary schools in the Netherlands using a pre-test-post-test control group design. Schools volunteered to be part of this "2011 mystery-project". All classes were higher general secondary education classes with students in the age of 15-17 years. At each school, the teacher(s) selected two classes to participate, one in the experimental condition (i.e. teaching with mysteries added to the regular geography curriculum) and one in the control condition (i.e. the regular geography curriculum). Five of the seven teachers instructed students in both the experimental group and the control group. In total, 221 students participated, 105 in the experimental group and 116 in the control group. The teachers involved had 5-30 years teaching experience. Although all teachers knew the work of David Leat, only two of them had used mysteries before.

In the experimental condition, the intervention consisted of teaching with three or four mysteries and each mystery took one lesson (of 45-60 minutes). The teachers could choose from five mysteries that in general were useful for the geography lessons in the fourth grade of higher general secondary education. Apart from the well-known Kobe mystery in the TTG program (about the earthquake in Kobe in Japan in 1995), four other mysteries were constructed using the design principles of TTG. These four mysteries were about:

- blood minerals and mobile phones in the Democratic Republic of Congo;
- the 2011 tsunami in Japan;
- the revolution in Benghazi, Libya and the Great Man Made River Project, and
- the Rio mystery described above.

All mysteries were tested in classes that were not part of this study.

Teachers received the series of mysteries and were instructed how to use them. For each of the mysteries the teachers received the information slips, a short introduction for the students and an instruction on how to organise the lesson. In each instruction, we stressed the need for enough debriefing time. We also mentioned in the instruction that the students should make a presentation of their solution of the mystery by creating a map with information (words, sentences) linked by arrows. Each mystery was prepared to be done in one lesson of 45-60 minutes, although we realised that at least 80 minutes should be used for each mystery in order to become most effective. We would have lost the readiness of the teachers to participate in the program if we had claimed two lessons for each mystery. The teachers did not receive a hands-on training as the schools were scattered over the Netherlands.

In the control condition, students attended the regular curriculum and teachers used the same geography textbooks as in the experimental condition. For example, when the curriculum topic was plate tectonics, teachers in the control condition followed the text book that also includes assignments. Teachers in the experimental condition also followed these textbooks and assignments, but for the mystery lesson they could choose to teach with the Kobe mystery or the tsunami mystery.

The pre-test was piloted with a small group of students of the same age and school level and evaluated by the researcher. After that, the test-instructions for the students were improved. The pre-test and post-test were taken individually in the classroom, with the teacher present (but not helping). Together with the tests, the teachers received a detailed instruction on how to organise the data collection (see Appendix A). After the post-test, each teacher reported circumstances that could have possibly influenced the outcomes of the tests (see questionnaire Appendix A). The data of one school were deleted as there were some disturbances during the intervention. This means that data of seven schools were included in the analyses.

The pre-test and post-test had the same structure (see Appendix A). Both tests included assignments to create a map with accurate relationships between variables that were connected to the UN Millennium Goals. Working on the Millennium Goals was new for all students. In the test-instruction for the students information was provided on the Millennium Goals and the variables involved. The pre-test and the post-test consisted of three assignments to explain the relationship between two variables of the Millennium Goals by making use of other variables belonging to all Millennium Goals. These Millennium Goal variables were mentioned explicitly in the student instruction. Students were asked to draw a map (a relational arrow diagram) of these variables and their relationships, and to explain each of the relationships. Both pre-test and post-test consisted of three of these assignments. Reading the instruction carefully, constructing three relational maps and explaining each relationship explicitly was time consuming. In order to give every student the opportunity to complete the test, no other assignments were added. The scores of both the pre-test and post-test were the total number of accurate relationships in the three maps. Doubtful or incorrect relationships were not counted. In Appendix A an example of a student's pre-test and the correction model of the tests can be found. The inter-rater reliability, based on the codes scores of two skilled and experienced geography teachers of 14% of the data, was satisfactory ($r = 0.96$). As schools varied in students' scores, multilevel regression analysis was used with the condition (experimental vs. control) as independent variable, the total number of accurate relations as dependent variable and the pre-test score as covariate.

2.4 Results

In Table 2.1, we present the descriptive statistics for each school. The multilevel regression analysis showed that students in the experimental condition scored significantly higher at the

Table 2.1 Average pre-test and post-test scores in experimental and control condition for each school and in total.

School		Group	N	Mean	Std. Deviation
All schools	Pre-test	Control group	116	5.89	2.87
		Experiment group	105	5.89	2.89
	Post-test	Control group	116	7.92	3.40
		Experiment group	105	8.82	3.42
School 1	Pre-test	Control group	24	7.17	2.697
		Experiment group	19	7.89	3.398
	Post- test	Control group	24	10.83	3.074
		Experiment group	19	11.63	3,004
School 2	Pre-test	Control group	16	4.06	2.265
		Experiment group	15	4.40	1.724
	Post-test	Control group	16	7.13	2.729
		Experiment group	15	9.27	3.654
School 3	Pre-test	Control group	13	4.69	2.287
		Experiment group	14	4.29	2.463
	Post-test	Control group	13	7.69	2.869
		Experiment group	14	8.14	3.110
School 4	Pre-test	Control group	7	3.29	2.563
		Experiment group	12	5.83	2.823
	Post-test	Control group	7	5.00	2.000
		Experiment group	12	7.25	2.301
School 5	Pre-test	Control group	22	6.09	2.706
		Experiment group	18	5.72	3.495
	Post-test	Control group	22	5.82	2.152
		Experiment group	18	7.67	3.678
School 6	Pre-test	Control group	18	5.22	2.365
		Experiment group	17	6.06	1.886
	Post-test	Control group	18	7,44	2.975
		Experiment group	17	8.35	2.644
School 7	Pre-test	Control group	16	8.38	2.391
		Experiment group	10	6.60	2.011
	Post-test	Control group	16	9.25	3.856
		Experiment group	10	8.50	3.689

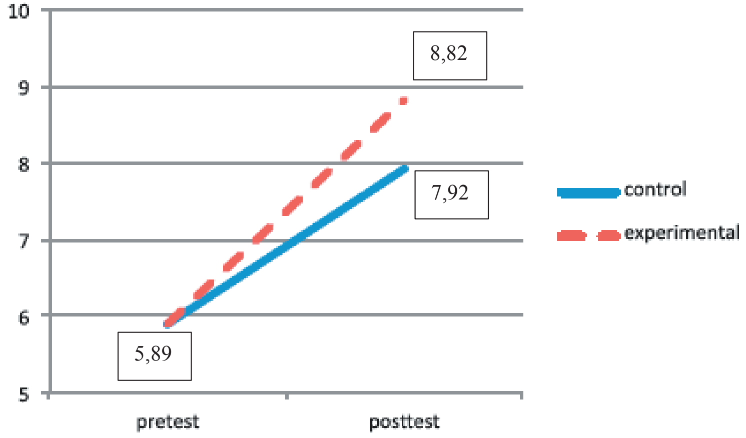


Figure 2.2 Mean number of correct relationships established by the experimental and control group in the pre-test and the post-test.

post-test than the students in the control condition did, after controlling for the pre-test scores (β = 1.01; *s.e.* = 0.40, *n* = 221). We did not find any significant differences between boys and girls. At the school level, about 15% of the variance in student scores was explained, leaving 85% of the variance on student level. The intervention explained 8% of the variance on student level, a moderate effect according to Cohen (1988). In Figure 2.2, we present this main result graphically. In the pre-test, the mean of correctly established relationships for students in the control condition was the same as for students in the experimental condition (5,89). In the pre-test, the total of correctly established relationships had a range from 0 to 16, in the post-test this range was from 1 to 19.

2.5 Conclusion and discussion

We found a significant effect of teaching with mysteries on students' skills to establish geographical relationships. So, there seems to be some empirical evidence for the effects of teaching with mysteries on fostering students' geographical thinking. We examined the use of mysteries in regular classroom conditions. The conditions of testing the effect could have been improved by a teacher training and more time. Teachers reported that the mysteries need to fit more with the content of the lessons that were instructed before using the mysteries. Additionally, teachers mentioned that they would like to have more than one lesson for working with a mystery, because students need time to read, to order, to relate and to rework the information and teachers need time for the debriefing session. A more precise fit with the curriculum, enough time for students to do their task and a good debriefing might result in mystery lessons being more effective for understanding geography and enhancing students geographical skills.

In addition to these improvements, we also might improve the mystery itself. In their influential paper on the effects of teaching with conflicts, Johnson and Johnson (2009) indicated that the use of conflicts in instruction can have positive effects on student's learning and their well-being. The authors mention four types of conflicts: (a) controversy, (b) conceptual conflict, (c) conflict of interests and (d) developmental conflicts. Conceptual conflicts are the most useful for mysteries, because this type of conflict is what mysteries are about. The authors define a conceptual conflict as a conflict "which occurs when incompatible ideas exist simultaneously in a person's mind or when information being received does not seem consistent with what one already knows" (p. 38). Formulating the challenging question of each mystery in such a way that it triggers a conceptual conflict might support the effectiveness of teaching with mysteries for students' geographical thinking. In the debriefing session first opposing solutions or opinions of students could be presented and then class discussion led by the teacher could solve conflicting views.

Adaptive teaching and supporting higher-order thinking skills of students are major issues in initial teacher training and in-service training. Many beginning teachers and student teachers are capable of managing classrooms, instructing their subject, and using a variety of teaching approaches. But many of them have difficulties with the more complex teaching strategies. This also counts for TTG, and teaching with mysteries in particular. Teachers should learn to teach students how to relate information that is presented in mysteries. Students cannot be expected to connect the first, second and third world themselves unless there is a step-by-step support of the learning process in the use of geographical thinking skills and the exploration of geographical knowledge. Using mysteries is quite different from recitation or making assignments with just one right answer. Mysteries involve critically analysing information and challenging others' positions, seeing issues from a variety of perspectives and seeking reasoned judgements. Additionally in in-service training or teacher training courses, more attention should be paid to debriefing, for example by discussing video tapes of debriefing practices, to further improve the effectiveness of the approach.

2.5.1 Limitations

We also should consider some limitations of our study. Roberts (2010) argues that qualitative studies carried out in authentic teaching and learning situations which recognise the complexities of classrooms are important to inform practice. First, a more in-depth analysis of the geographical relations that were included in the tests and in the mysteries of the experimental group might provide additional explanations for the effect we found. Second, to get a more profound view of the impact of mysteries on students' skills of relating geographical phenomena within and between regions it is necessary to collect data about the process of geographical thinking by, for example, interviews with students and teachers, class observations or thinking-aloud protocols. Such an analysis can help to better understand how students learn this kind of complex skills.

2.5.2 Recommendation

Many teachers report that mysteries are challenging and motivating for students (Van der Schee, Vankan, & Leat, 2003). That is a good start, but not good enough. Actively engaged students are a sign that teaching with mysteries is going well, but the question is what is going well exactly, the classroom management or the learning process or both (Nuthall, 2005)? The educational challenge is to make mysteries an instrument to come to higher levels of thinking. Based on the results of the current study, we suggest that mysteries can be a way to help teachers and students to think about real-life issues within and between domains of geography as well as about learning goals and learning processes.

Chapter 3

Students’ geographical relational thinking when solving mysteries*

Abstract

Geographical relational thinking is an important part of geographical thinking. This descriptive research was conducted to seek evidence on students’ ability to establish geographical relationships which could help teachers to foster their geographical relational thinking. Sixty-nine small student groups from six secondary schools in the Netherlands were observed when solving a mystery. All relationships students established were analysed and the SOLO taxonomy of Biggs and Collis (1982) was used to analyse how coherent their solutions were. The results revealed that students had difficulties with complex, abstract and physical geographical relationships. A large proportion of the groups also had difficulties understanding the interdependence of the relationships. These findings underpin the usefulness of activities like mysteries which offer opportunities to practise, assess and teach geographical relational thinking in geography lessons.

* This chapter is based on: Karkdijk, J., Van der Schee, J. A., & Admiraal, W. F. (2019). Students' geographical relational thinking when solving mysteries. *International Research in Geographical and Environmental Education, 28*(1), 5-21.

3.1 Introduction

Relational thinking is at the core of the secondary school subject of geography. Thinking about interconnections from global to local scales characterises the geographical perspective (GA, 2012; Jackson, 2006). Lambert (2004, p. 1) argues that interdependence is "perhaps the key geographical concept". Recent research has revealed, however, that identifying, analysing, explaining and evaluating geographical relationships is difficult for many students in secondary education (Favier & Van der Schee, 2014a; Karkdijk, Van der Schee, & Admiraal*,* 2013). For this reason deliberate attention should be paid to geographical relational thinking in education and research. We conducted this descriptive research to gain more insight into which geographical relationships students were able to establish and which caused difficulties. We used the mystery, a problem-solving strategy devised to provoke geographical reasoning and relational thinking by students (Leat & Nichols, 2003). We also wanted to know how coherent students' solutions to the mystery would be. Our research aim was to provide more evidence on students' ability to establish geographical relationships, which can help teachers to find ways to advance their geographical relational thinking.

3.2 Relational thinking in school geography

How can we describe geographical relational thinking? The International Charter on Geographical Education declares "the study of human activities and their interrelationships and interactions with environments from global to local scales" as the content of geography (IGU-CGE, 2016, p. 4). Understanding human-environmental relationships has always been important within the discipline (Golledge, 2002). The units of analysis to study these relationships are concrete regions (Favier & Van der Schee, 2014a), which change continually in our globalised and interconnected world. In order to describe geography as a school subject, Van der Schee (2000) designed the Geographical Analysis Model. This model distinguishes two kinds of interactions that cause regional change. First, vertical relationships are distinguished: interactions within and between human and natural systems, within regions. These interactions are also possible between factors on different scales but they cause changes in particular regions. For example, global climate change will affect farmers' practices in the inner Niger Delta. Second, the model distinguishes horizontal relationships: interactions between (sub)regions. Changes in one region cause changes in another connected region. For example, a flow of migrants from one region to another causes changes in both regions. Horizontal and vertical relationships together are geographical relationships: they make and change regions where people live.

We define geographical relational thinking as a core element of geographical thinking, containing the analysis, explanation and/or evaluation of the vertical and horizontal relationships (the geographical relationships) that cause change in regions on different interconnected scales. In this reasoning process, students have to apply their geographical conceptual knowledge to specific regional contexts. It therefore demands higher order thinking skills (Favier & Van der Schee, 2014b).

3.3 Mysteries in geography education

A mystery is a complex strategy of the Thinking Through Geography (TTG) program. Developed in the 1990s in the UK, the TTG program focuses on the "infusion" of higher order thinking skills in geography lessons (Moseley et al., 2005, p. 28) such as creative thinking, reasoning and establishing relationships (Leat, 2001; Leat & Nichols, 2003; Vankan & Van der Schee, 2004). The mystery is a complex strategy and therefore not widely used by geography teachers (Hooghuis, Van der Schee, Van der Velde, Imants, & Volman, 2014; Leat, Van der Schee, & Vankan, 2005), but it is "probably the most powerful strategy" of the program (Leat, 2001, p. 51). The problem of the mystery is always an open question formulated as a cognitive conflict that triggers students to think. A mystery consists of three parts: first, the introduction of the problem and the required instructions; second, small group collaboration where students have to use 16-30 information strips to solve the mystery; and third, a whole-class debriefing. Mysteries offer teachers opportunities for diagnostic and formative assessment. Teachers can listen to students' discussions and observe their manipulation of information strips to signal misconceptions and assess their level of understanding (Leat & Nichols, 2000b; Leat & Nichols, 2003).

To date little research has been done to analyse student learning by solving and evaluating mysteries. Leat and Nichols (2000b) conducted a descriptive study into student activities in small groups to solve a mystery in secondary education. They focused on the process of manipulating the information strips and observed five stages. In the *display stage*, the strips were read to comprehend the information they contain. In the *setting stage*, groups analysed and classified the information on the strips. Most groups proceeded to the next stage, the *sequencing and webbing stage*, where relationships between information strips were established. In the *reworking stage*, new and more coherent relationships were established. A few groups moved to the *abstract stage*, where discussions were more abstract, extending beyond the given data. According to the authors, student activities and discussions in these stages coincided with the thinking processes described by the levels of the SOLO taxonomy, indicating a progress in complexity of thinking as a group moved from one observed stage to the next (Leat & Nichols, 2000b; 2003). The SOLO taxonomy, developed by Biggs and Collis (1982), describes the quality of responses on the basis of the structure of the learning outcomes (SOLO) and has five levels:

(1) the *pre-structural level*: no relevant datum is given to the question;
(2) the *uni-structural level*: one datum is correctly related to the question;
(3) the *multi-structural level:* two or more data are correctly related to the question but without interrelationships between the data;
(4) the *relational level*: two or more data are correctly related to the question and interrelationships between these data are given to make a coherent explanation;
(5) the *extended abstract level* where abstract principles are used to hypothesise beyond the given data (Biggs & Collis, 1982).

Two effect studies showed the positive effect of the use of mysteries on students' relating skills. Van der Schee, Vankan, and Leat (2006) measured a significant positive effect of using three TTG strategies (five W's, reading photographs and mystery) on the number of relationships established by students in lower secondary education. A larger scale research study revealed that the repeated use of mysteries helped secondary school students to establish relationships (Karkdijk et al., 2013).

Although some research has been carried out on the learning processes and the effects of the use of mysteries, there is a lack of evidence on the nature of students' geographical relational thinking while solving a mystery.

3.4 Research aim and research questions

The aim of this descriptive study was to seek more evidence on students' ability to establish geographical relationships, which could help teachers to find ways to advance their geographical relational thinking. We used mysteries to elicit students' reasoning. Our research questions were:

(1) *Which geographical relationships do students in small groups establish to solve a mystery?*
(2) *How coherent are the solutions to the mystery posed?*

We expected our study to provide insights into possible deficiencies in students' geographical relational thinking. Deficiencies could consist of a lack of certain important relationships within their reasoning or the formulation of incomplete relationships (first question). Because of the difficulties students have with relational thinking, we also expected that a large proportion of the students would come up with a solution below the relational level of the SOLO taxonomy, showing a lack of insight into the interdependence of relationships (second question).

3.5 Method

3.5.1 Materials

We designed two mysteries using the design principles of Leat and Nichols (2003). Both were reviewed by an educational geographer familiar with the TTG program and mysteries and tested by two geography teachers in classrooms. The mysteries were regional case studies concerning relationships between man and environment and between local actors within a specific region, and hence we considered them to be geographical mysteries (see Appendix B). The Rio mystery questioned the decision of a slum dweller (Fabio) not to move out of his favela which was threatened by landslides into another house in the suburbs of Rio. The focus was on understanding the geography, society and economy of Rio to explain his decision. The Jakarta mystery asked students to evaluate the complaint of a Jakarta official that slum dwellers along rivers in Jakarta were causing the floods in Jakarta. The focus was on understanding the hydrological system of the river basin and delta where Jakarta is situated in order to evaluate the accusation. We present the content of each mystery in more detail in the results section of this study.

3.5.2 Participants

Twelve qualified and experienced geography teachers and 205 students from six secondary schools in the Netherlands participated in our research project. The project was carried out between January and June 2015 in the senior years of higher general secondary education (havo, 4th and 5th years) and pre-university education (vwo, 4th, 5th and 6th years) with students in the age range 15-18 years. The teachers formed three groups of three students each from their class, using their geography grades: one group of students belonged to the highest 30% of the class, one group to the lowest 30% and one was a mixed group (1 student from the highest group and 2 from the lowest). Because of absenteeism among the selected students, the teachers had no option other than to select another student in some cases. In 11 cases, therefore, the group composition did not fit into one of the three categories and those groups were assigned an intermediate position (Table 3.1). The teachers decided which class would work on the mystery, because it had to match the content of their curriculum. We observed 69 groups: 35 groups solved the Rio mystery and 34 groups solved the Jakarta mystery. Table 3.1 shows the distribution of the groups regarding geographical ability, gender, educational level and year.

Table 3.1 Distribution of groups regarding geographical ability, gender, educational level and year.

	Rio	Jakarta
Total number of groups	35	34
Geography grades		
Groups in highest 30%	11	11
Groups in lowest 30%	9	12
Mixed groups	8	7
Intermediate groups	7	4
Gender		
Girls only groups	5	11
Boys only groups	7	5
Mixed groups	23	18
Educational level and year		
havo-4	17	3
havo-5	0	19
vwo-4	9	3
vwo-5	6	6
vwo-6	3	3

3.5.3 Intervention

After the teacher had introduced the mystery and given the instructions to all the students, each of the three selected groups had to work in separate rooms. They had to represent their solution as a free style concept map (structured by the group as they wanted to represent their solution). There is evidence that concept maps are a representation of learners' knowledge structures and give insights into students' understanding of relationships between concepts (Srinivasan, McElvany, Shay, Shavelson, & West, 2008). Because the teachers lacked the time, their students did not practise with constructing a concept map, but were given an example and a short instruction on how to construct it. They had to formulate the concepts, draw the cause-and-effect arrows and verbalise the linking phrases by themselves. This task was clearly at the low end of teacher-directedness, to be as "content rich" and "process open" as possible (Ruiz-Primo, Shavelson, Li, & Schultz, 2001, p. 102). Evidence suggests that this construct-a-map-technique is the best, the "gold standard", for revealing students' knowledge structures (Yin et al., 2005, p. 166). Although more difficult to analyse, "concept mapping tasks that do not constrain the responder have the highest validity for measuring student knowledge" (Wehry, Monroe-Ossi, Cobb, & Fountain, 2012, p. 86). Each group was given the time they needed to complete their work and received only limited scaffolding from the researcher (how to construct the concept map). A camera captured the group discussions.

3.5.4 Analysis

3.5.4.1 Relationships

The observed group discussions were transcribed in full and analysed in conjunction with the concept maps for formulated correct and relevant relationships. All relevant connections that the groups established between pieces of information provided within the mystery or added from outside to solve the problem, were considered as correct geographical relationships. Some connections were on the conditional side of the problem, others on the direct cause and effect side or on the spatial side, but we did not draw a distinction between these different kinds of connections, since we were only interested in the students' ability to establish the necessary relationships in order to understand and explain the regional problem of the mystery. We also decided not to distinguish between vertical and horizontal relationships, because our focus in this contribution was on the establishment of geographical relationships to solve the mysteries, no matter whether they were horizontal or vertical.

3.5.4.2 Concept maps

The scoring system for the established relationships on the concept maps, and within the transcriptions of the group work, was constructed with two raters who were qualified and experienced geography teachers. As many relationships as possible were discussed for each mystery and then formulated and summarised in an "ideal" concept map. These relationships consisted of concepts provided within the mystery and of concepts that students added to their solutions. After the scoring system was completed, each rater analysed four groups for established relationships. The inter-rater reliability between the researcher and the two raters for concept maps and transcriptions together was satisfactory (κ= 0.823). The intra-rater reliability test of four groups (two from each mystery) for concept maps and transcriptions together was also satisfactory (κ= 0.725).

Groups represented the structure of their solutions in their concept maps, visualising the degree of coherence of the relationships. However, great representational differences between maps hindered analysis. For this reason, the relationships a group established were put into the same standard format of the raters' "ideal" concept map. This standard concept map consisted of relationships a group established, grouped together into one of the main causes (factors) for the problem of the mystery. These factors were determined by the designer of the two mysteries (the first author) and discussed with both raters. The factors were used to design the ideal concept map. For example, Figure 3.1 gives the factors "house" and "government actions" and some of their relationships (Rio mystery). The factor "house" in the standard concept map contains all relationships concerning the quality of Fabio's house and

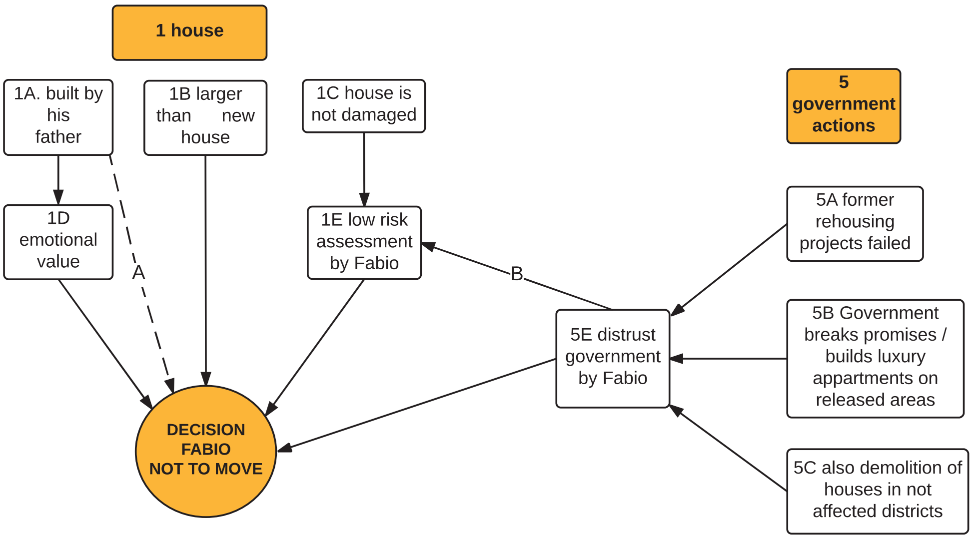


Fig. 3.1 The factors "house" and "government actions" and some of their relationships.

the factor “government actions” contains all relationships concerning past and present government actions on housing, demolition of favelas, construction of new apartments et cetera.

As groups did not include all their established relationships in their concept maps, we added all other correct and relevant relationships established in the group discussion to construct a “complete” standard concept map for each group. Finally, we constructed a “total” concept map of relationships established by all groups together for each mystery and analysed these for established relationships.

3.5.4.3 SOLO levels

The complete standard concept map of each single group was used in conjunction with the transcriptions of their group discussions to determine the SOLO level of the solution of each group. We carefully followed the general description of the SOLO taxonomy of Biggs and Collis (1982), their approach to operationalising the criteria for each SOLO level for their geography study and Stimpsons’ approach (1992), which is in line with Biggs and Collis’ approach. In order to analyse students’ geographical explanations or to design SOLO based ordered outcome questions for geography, three criteria were used by Biggs and Collis and by Stimpson to distinguish between SOLO levels. They were: (1) the number of explanatory factors / pieces of information needed; (2) the interrelationships between these factors; and (3) the use of deductive, abstract arguments, alternative hypotheses and/or generalisations. Our criteria for distinguishing between SOLO levels, in line with this approach, were:

(1) the number of factors correctly connected with the question of the mystery;
(2) the use of branches and cross-links as expressions of interrelationships within and between factors; and
(3) abstract reasoning and the transfer to other regional contexts.

An incompletely connected factor was characterised by one or more incomplete relationships. Incomplete relationships lacked one necessary concept, and thus remained partly unclear. For example, within the relationship: “house built by his father → Fabio’s decision not to move” (dashed arrow A in Figure 3.1), the necessary concept “emotional value” is lacking. If this was the only relationship a group established in the factor “house”, this factor would be incompletely connected with Fabio’s decision. The use of branches, several relationships coming together in or departing from one concept (5E in Figure 3.1), illustrated the ability to establish interrelationships within a factor. Cross-links indicated the ability to formulate interrelationships between factors, and gave the mystery solution coherence (arrow B in Figure 3.1). A cross-link with an incompletely connected factor was considered as an incomplete cross-link. Group discussions were analysed for the use of abstract reasoning (using alternative hypotheses and/or reasoning with generalisations) and transfer to other regional contexts.

An outcome on the pre-structural SOLO level indicated that no factor was correctly connected with the question of the mystery. An outcome on the uni-structural level indicated that the group was able to connect only one factor correctly to the question of the mystery. One or more incompletely connected factors besides the correct one resulted in the transitional level to multi-structural. A multi-structural response indicated that several factors were correctly connected with the question of the mystery, but without any cross-link between factors. When a group established only incomplete cross-links between factors, the output was on the transitional level to relational. An output on the relational level indicated the use of most factors, the establishment of branches within factors and the establishment of one or more cross-links between factors. The transcriptions of groups with an output on the relational level were analysed for the use of abstract reasoning and other regional contexts. If a group used both, their work was classified as extended abstract. If a group used only one of these, their work was classified as transitional to extended abstract. After detailed instructions and several try-outs and discussions, an experienced geography teacher analysed the work of 12 groups as a second rater and found the same SOLO levels as the researcher had established.

3.6 Results

The results for each mystery are presented separately. After a concise description of the explanatory factors necessary for the solution of the mystery, we present a set of two concept maps of relationships that were established by all groups: one of relationships without cross-links and one of only cross-links (Figures 3.2-3.5). Cross-links are separately presented for clarity reasons. The concept maps were designed in a non-hierarchical way to illustrate the essence of each mystery. The weight of an arrow between the concepts is an expression of the number of groups that established that relationship; that number is shown in each arrow. Concepts which were provided within the mystery have a dashed frame. Finally, we present the structure of the solutions, using the SOLO taxonomy.

3.6.1 Study 1: Jakarta mystery

Six factors were necessary in order to clarify the causes of the annual floods, the contribution of slum dwellers to these floods and to evaluate the complaint of the government official. These were: 1) the deforestation and urbanisation in the region that cause peak flows in the Ciliwung, the main river in Jakarta; 2) the on-going construction of slums in the river beds in Jakarta that causes obstruction, narrowing and hardening of these river beds; 3) the lack of municipal services in Jakarta especially for slum dwellers, resulting in blockages with waste and garbage of the rivers and badly maintained drainage channels; 4) the geomorphology of the region, where rivers come down from the mountains into the delta and lose velocity; 5) the relative sea level rise caused by the absolute rise in sea level and the subsidence of downtown Jakarta as a result of groundwater withdrawal; and 6) the torrential rains of the monsoon that cause peak flows in the river. These factor numbers correspond with the numbers in Table 3.2 and in the concept maps (Figures 3.2 and 3.3). Figures 3.2 and 3.3 show, respectively, the number of relationships without cross-links and the number of cross-links established by all groups on their concept maps and in their group discussions. The share of total relationships each factor had is presented in Table 3.2.

Table 3.2 Relationships per factor within the Jakarta mystery.

Factor	1	2	3	4	5	6	Total
1. Total number of relationships	86	160	131	19	97	55	548
2. % of total relationships	15.7	29.2	23.9	3.5	17.7	10.0	100.0
3. Index total relationships (548/6 = 91.33 = 100)	94.2	175.2	143.4	20.8	106.2	60.2	600.0
4. Concepts per factor	7	9	5	5	7	1	34

Table 3.2 gives the percentages of the total relationships per factor and the indexes which show whether a factor had an above or below average share of total relationships (mean = 100). Table 3.2 shows that factors 2 (slums along the riverside) and 3 (municipal services in Jakarta) had an above average share of total relationships, while factors 1 (deforestation), 5 (relative sea level change) and 6 (monsoon) took a position in the middle ground. Factor 4 (geomorphology) had a far below average share, indicating its possible difficulty for students. Factor 6 (monsoon) also had a below average share, but it only had one concept. This pattern was fairly similar for each year.

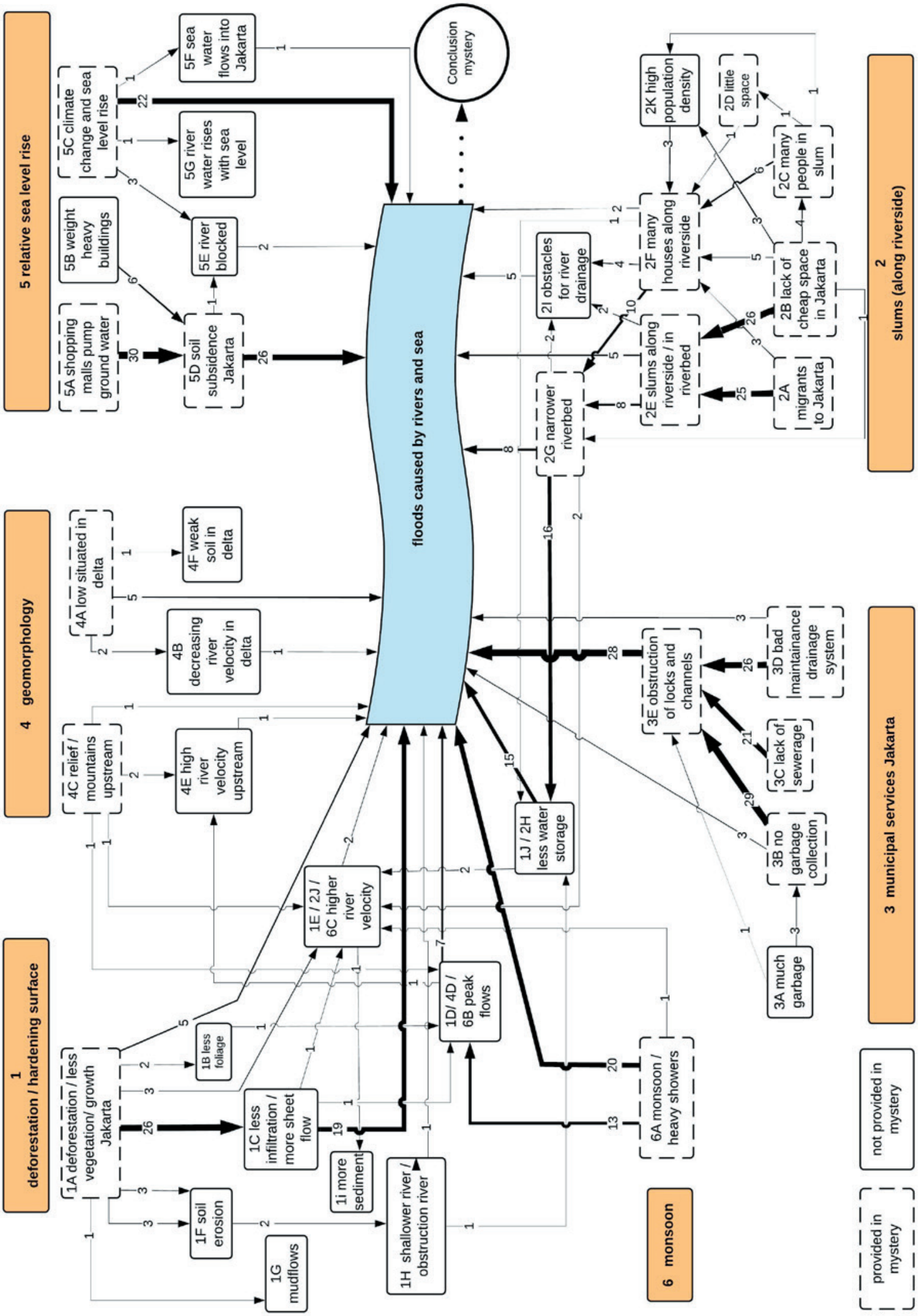


Figure 3.2 Total map of established relationships in the Jakarta mystery, without cross-links.

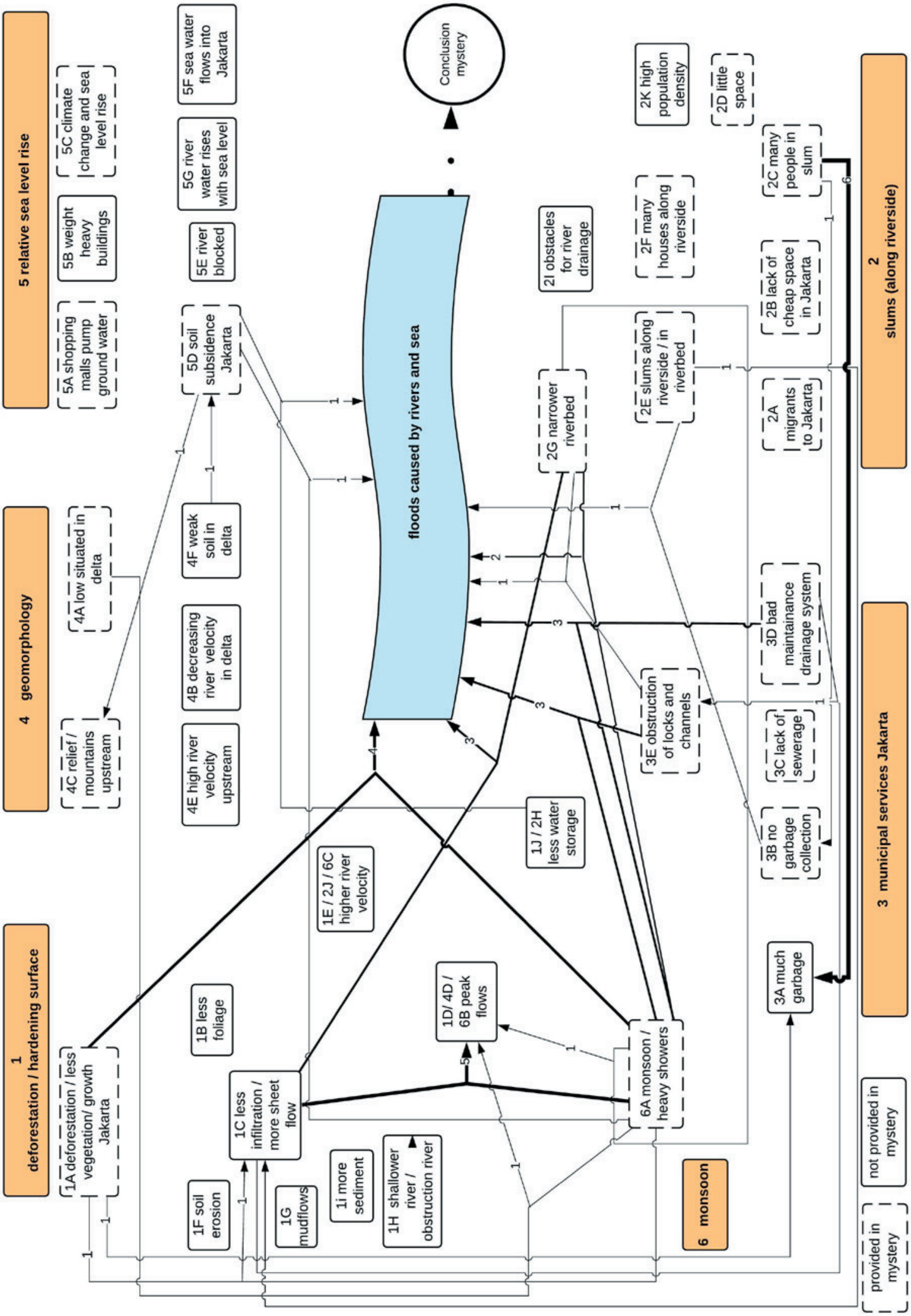


Figure 3.3 Total map of established cross-links in the Jakarta mystery.

A more detailed analysis of the relationships within each factor (Figures 3.2 and 3.3) revealed differences between the frequencies of the use of concepts (this was also fairly similar for all years). Within three factors (1, 2 and 5), relationships between only a small number of concepts dominated. Within factor 1 (deforestation), the relationships "deforestation causes less infiltration and more sheet flows, which causes floods", were dominant. The other relationships within this more physical geographical factor were mostly neglected. Most cross-links were also established using the concepts of deforestation and less infiltration in combination with peak flows, the monsoon (factor 6) and narrower riverbeds (factor 2). Within factor 2 (slums in riverbeds), the difference in use of concepts was somewhat less pronounced. Most groups formulated the relationships "migrants come to Jakarta and lack of cheap space in Jakarta causes migrants to build slums in riverbeds". The more complex relationship between the building of slums in the riverbeds and the narrowing of these riverbeds to explain floods was established by only approximately half of the groups. Most groups that established this relationship connected this with the not provided concept "less water storage" to explain floods. Within factor 5 (relative sea level rise), three straightforward relationships dominated: "shopping malls pump ground water up which causes soil subsidence which causes floods" and "absolute sea level rise causes floods". The more complex relationship using the not provided concept "river blocked" to explain why absolute sea level rise causes floods, was only rarely formulated. Almost all of the provided, fairly simple and concrete concepts belonging to factor 3 were used by most groups. The physical geographical relationships belonging to factor 4 were complex and abstract: students needed to relate the decreasing slope of the riverbed to the decreasing water velocity in the delta that causes Jakarta's vulnerability to flooding. Factor 4 was also hardly ever used to make cross-links. Factor 6 (heavy rains of the monsoon cause floods) was used by most groups, but the necessary, not provided physical geographical concept "peak flows" to explain this relationship, was far less used. The monsoon, as source of extra water transportation, was used most frequently of all the factors to formulate cross-links.

3.6.1.1 Structure of solutions

Table 3.3 shows the distribution of the groups on SOLO levels. Fifteen groups had an outcome on relational level or higher: they had a coherent solution with the use of most factors and characterised by interconnections between factors. Four of these groups also used abstract reasoning to explain the floods. Seven groups had an outcome on the transitional

Table 3.3 Distribution of groups on SOLO levels and means of correctly connected factors and relationships and loose relationships per level in the Jakarta mystery.

SOLO-level	Number of groups	Correctly connected factors	Correctly connected relationships	Loose relationships	% Loose relationships
1. Uni-structural U	0	0	0	0	0
2. Transitional U/M	1	2	3	6	66.7
3. Multi-structural M	11	4.5	11.5	2.2	15.9
4. Transitional M/R	7	4.1	11.6	2.1	15.6
5. Relational R	11	5.3	16.2	0.8	4.8
6. Transitional R/EA	3	5.7	20.3	1.3	6.2
7. Extended Abstract EA	1	6	22	0	0
Total	34	4.7	13.9	1.7	10.9

level towards the relational level and showed only rudimentary coherence in their solutions. Twelve groups operated (almost) on a multi-structural level. Their solution was limited to some isolated factors and was not coherent. There were no groups operating on the pre-structural or uni-structural level. The mean SOLO level of groups working on the Jakarta mystery was 4.21: on the transitional level towards the relational level. We found no significant relationships between SOLO level and educational level (havo or vwo) or SOLO level and year (4, 5 or 6). Our research on the structure of the solutions also gave insight into "loose" relationships, relationships without any connection to floods in Jakarta. For example, one group related the heavy monsoon rains to peak flows in the rivers, but did not relate peak flows to floods. The Jakarta mystery had 58 loose relationships out of a total of 530 relationships (10.9%).

3.6.2 Study 2: Rio mystery

Six factors were necessary to explain Fabio's decision to stay in his neighbourhood which was threatened by landslides. These were: (1) the emotional value of his house that was built by his father; (2) the neighbourhood with migrants also coming from north-east Brazil that binds to the community; (3) the great distance from his new home to his work in the centre of Rio that would cause loss of time and money or even put him at risk of losing his job; (4) high land values in the centre of Rio that give opportunities to developers to force favela dwellers out to build luxury apartments for high profits; (5) government actions that reveal that government concerns for the safety of favela dwellers was not the only incentive for the rehousing project; and (6) the organisation of the football World Cup in 2014 and the Olympics in 2016 in combination with the fact that Rio is a major tourist destination. As a consequence of this last factor, Rio attracts many people, which creates a lot of work but could also trigger the government to move the slums away from the centre in order to create a better image of the city. Distrust of government intentions with their rehousing project (an underlying concept not explicitly provided within the mystery) was therefore one of Fabio’s main reasons for deciding not to move. These factor numbers correspond with the numbers in Table 3.4 and in the concept maps (Figures 3.4 and 3.5). Figures 3.4 and 3.5 show the number of relationships without cross-links and the number of cross-links established by all groups on their concept maps and in their group discussions. The share of total relationships each factor had is presented in Table 3.4.

Table 3.4 Relationships per factor within the Rio mystery.

Factor	1	2	3	4	5	6	total
1. Total number of relationships	110	11	93	56	85	65	420
2. % of total relationships	26.2	2.6	22.1	13.3	20.2	15.5	100
3. index total relationships (420/6 =70 =100)	157.1	15.7	132.9	80.0	121.4	92.9	600
4. Concepts per factor	5	2	6	5	5	7	30

Table 3.4 gives the percentages of total relationships per factor and the indexes which show whether a factor had an above or below average share of total relationships (mean = 100). The pattern which Table 3.4 shows was fairly similar for each grade. Factors 1 (house) and 3 (location and work) had an above average share of total relationships, as had factor 5 (government actions). Factor 2 (neighbourhood), with a far below average share, was more complex, for students first had to recognise that migrants from the same region in Brazil often constitute close communities in cities by chain migration (not provided in the information). Factor 4 (location and land values), the most abstract one, also had a below average share of total relationships. Factor 6 (World Cup, Olympics and tourists) had a slightly below average share. This factor was a more complex one, for students had to add intermediate concepts to understand Fabio’s decision, like the image of Rio. A more detailed analysis of the relationships within each factor revealed the differences in use of concepts within each factor (Figures 3.4 and 3.5). The more concrete and fairly simple relationships within factor 1 between the concepts “house built by his father” and “old house larger than the new house”

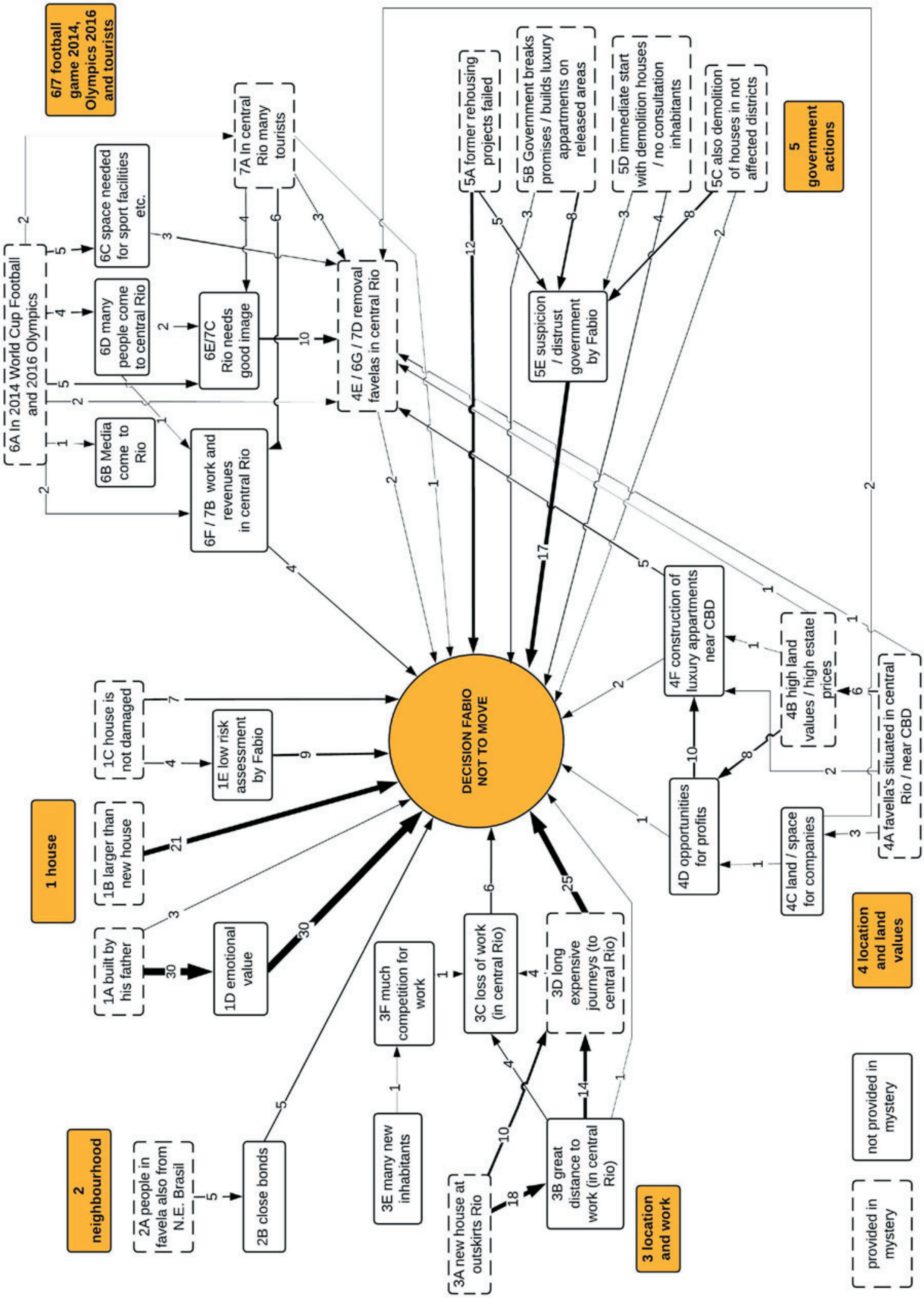


Figure 3.4 Total map of established relationships in the Rio mystery, without cross-links.

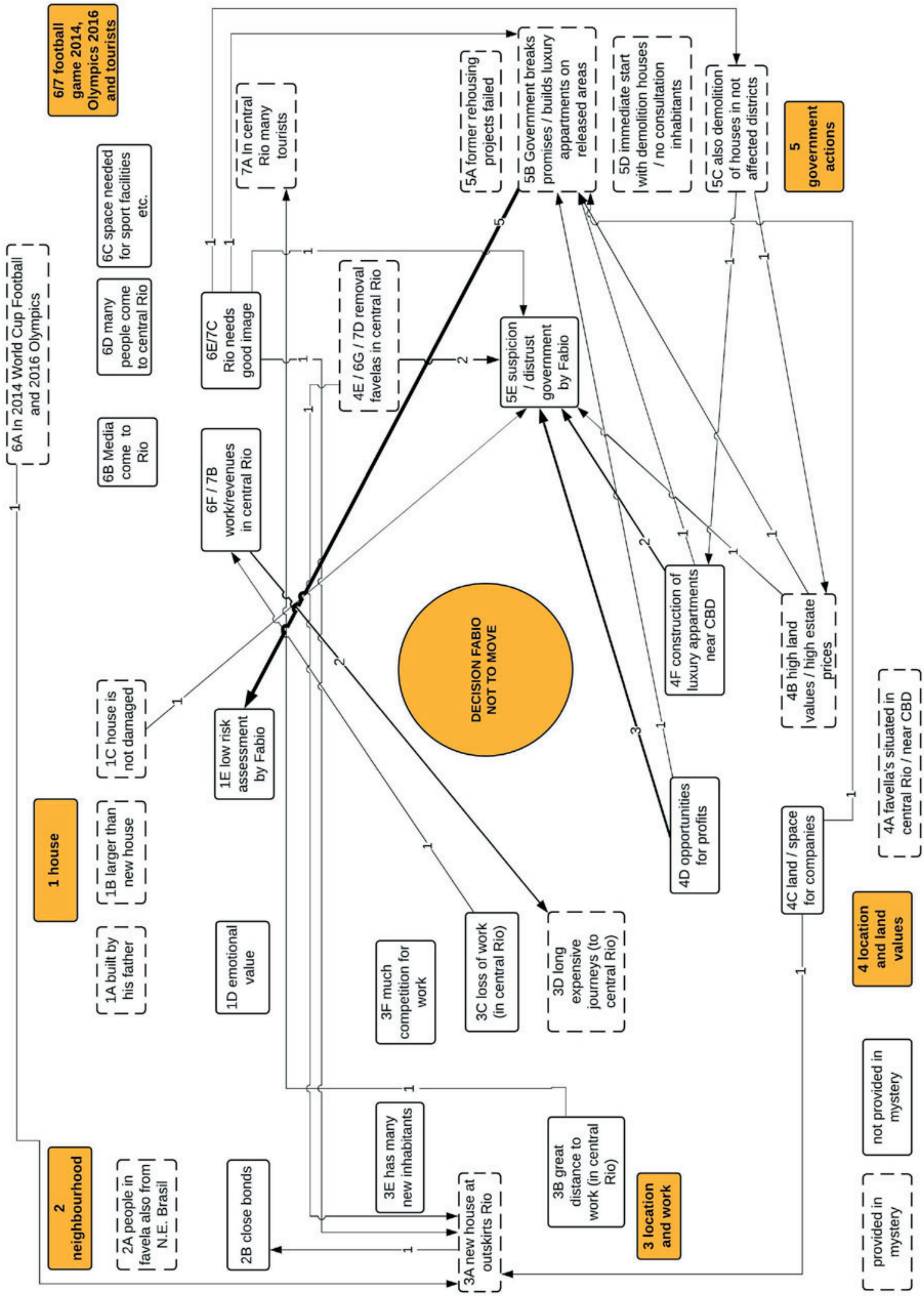


Figure 3.5 Total map of established cross-links in the Rio mystery.

and Fabio's decision were the most used. Within factor 3, the relationships "new house on outskirts of Rio" means "great distance to work" and "long and expensive journeys to central Rio" so Fabio will not move, dominated. Information on the actual location of Fabio's work was not provided; students had to make this inference by themselves. Approximately half of the groups concluded that Fabio's move to the outskirts of Rio would mean a great distance to or loss of his work in central Rio, so they included the concept "great distance to work" in their reasoning. Ten other groups made a shortcut by simply relating a house on the outskirts of Rio to longer and more expensive journeys to central Rio, without any consideration as to why Fabio has to travel to central Rio. Factor 4 was a difficult one, because its concepts were abstract and not all provided. Information was provided on land values in Rio and on the location of Fabio's favela close to the central business district (CBD). Students had to make the inference from this information that this central location is very expensive and therefore gives good opportunities for investors (also the government) to make high profits in real estate. Cross-links were mainly established with factor 5 (distrust of government intentions). The relationship "suspicion of the government" with Fabio's decision not to move was a dominant one within factor 5. This relationship was important within the mystery to understand Fabio's decision and students had to make this inference by themselves. Almost half of the groups explained Fabio's decision by using this relationship. Most cross-links were made with factor 5, with the profits to be made by replacing favelas by luxury apartments (factor 4) and with the good image that Rio needed as a tourist city and as host of the World Cup and the Olympics (factor 6). Within factor 6 no relationship clearly dominated.

3.6.2.1 Structure of solutions

Table 3.5 Distribution of groups on SOLO levels and means of correctly connected factors, relationships and loose relationships per level, Rio mystery.

SOLO level	Number of groups	Correctly connected Factors	Correctly connected relationships	Loose relationships	% loose relationships
1. Uni-structural U	1	1	2	5	71.4
2. Transitional U/M	3	2.3	5.7	0.7	10.5
3. Multi-structural M	11	2.5	6.3	0.4	5.5
4. Transitional M/R	6	3.2	9.3	1.8	16.4
5. Relational R	13	4.3	14	1.2	7.6
6. Transitional R/EA	1	6	24	2	7.7
7. Extended Abstract EA	0	0	0	0	0
Total	35	3.3	10	1.1	10

Table 3.5 depicts the distribution between the groups on SOLO-levels. Fourteen groups working on the Rio mystery operated on relational level or higher; only one of these also explicitly used general rules to explain Fabio's decision. Fourteen groups working on the Rio mystery (almost) operated on a multi-structural level and six groups took a transitional position between the multi-structural and the relational level. There were no groups operating on the pre-structural level. The mean SOLO-level of the output of groups working on the Rio mystery was 3.86, between the multi-structural and the transitional (M/R) level. We found no significant relationships between SOLO-level and educational level (HAVO or VWO) or SOLO-level and year (4, 5 or 6). In the Rio mystery, groups made 58 loose relationships (total 389 = 10%)

3.7 Discussion

This descriptive study was conducted to gain insight into which geographical relationships students establish to solve a mystery (first research question) and how coherent their solutions to the mystery posed are (second research question). Both types of information can be understood as indicators of students' geographical relational thinking skills. Concerning the first question, we found that more complex, abstract relationships were formulated less than other relationships and also led to incomplete reasoning.

Concerning the second question, we found that although a large minority of the groups had an outcome on the relational SOLO level or higher, a large number still had an outcome on the multi-structural level or lower, neglecting the interdependence of the factors that caused regional change. Only five groups had an outcome above relational level. Although the mysteries did not ask the students to compare the regional situation in Rio or Jakarta with other regions, groups could have applied theoretical reasoning, using generalisations.

3.7.2 Implications

These findings of deficiencies in students' geographical relational thinking when solving a mystery raise the question of how teachers could help students to advance their geographical relational thinking skills. We make three suggestions.

First, the regular use in geography lessons of teaching strategies like mysteries, focused on geographical reasoning and involving the construction of a coherent solution or explanation, is indispensable. The repeated use of mysteries fosters students' relational skills (Karkdijk et al., 2013). Mysteries also give teachers the opportunity to teach with an explicit focus on interdependence. Renshaw and Wood (2011) used interdependence explicitly as a structuring tool for lessons about interconnected physical geographical themes and conclude that it acted as a threshold concept, giving students new insights into geography as an interconnected whole. Strategies like mysteries also offer teachers opportunities to gain insight into students' shortcomings and difficulties in geographical relational thinking. These insights can be detected from the group discussions and representations, but also during whole-class discussions at the debriefing sessions. These whole-class discussions should be focussed on relationships that are difficult for students, like more complex, abstract and physical geographical relationships.

Second, whether it be when teaching with mysteries or with other geographical assignments or exercises, it would be very useful to ask students to speak aloud in class mentioning all relational steps. Incomplete reasoning and misconceptions would then be detected and could be corrected. In order to be able to help students in their reasoning, teachers should have the necessary subject knowledge of the regional problem at hand.

Third, although our research was descriptive, we suggest the relevance of geographical subject knowledge as an important prerequisite for understanding the problem of the mystery. Knowledge of the spatial pattern of land values in mega cities like Rio will help students understand the economic forces that push favela dwellers away from central locations. Physical geographical knowledge about relationships between a location in a delta and diminishing river velocities on the one hand, and between rising sea levels and water levels in rivers on the other, can help students understand the significance of living on the edge of a delta. This conceptual subject knowledge also facilitates its application to other regional contexts.

3.7.2 Limitations and strengths

A limitation of our research was our decision not to distinguish between different kinds of connections that students made. More detailed research into difficulties that may emerge when students have to differentiate between regional conditions and direct cause-effect relationships or between vertical and horizontal relationships could reveal other obstacles to sound geographical relational thinking. Another limitation was the reliance on students' discussions and concept maps, without being able to talk to the students afterwards and evaluate their results with them. This would have elicited why certain concepts were employed less than others. One of the merits of our research was that it covered many groups from different schools and that students had enough time to work on the mysteries.

3.8 Conclusion

We conclude that the use of mysteries, as a tool to engage secondary school students in the geography of changing regions, provides teachers with great opportunities to foster their geographical relational thinking. Attention to students' shortcomings in establishing geographical relationships and a clear focus on interdependence should help students to understand regional change within our dynamic and interconnected world. Being able to reason with geographical relationships should also be an area for teacher trainers in higher geography education to focus on, because teachers' shortcomings will inevitably have repercussions in secondary geography classrooms. Continuing research on the ability of students and teachers to reason with geographical relationships in order to understand, explain and evaluate regional change is strongly recommended.

Chapter 4

Small-group work and relational thinking in geographical mysteries: a study in Dutch secondary education*

Abstract

Relational thinking is a necessary skill for building students' individual capabilities and a core concept in geography education. Geographical relational thinking refers to being able to give interrelated, causal explanations for geographical phenomena such as regional change. The aim of this study was to gain more insight into differences in relational thinking between small groups of students working together on an assignment to explain a regional event which was framed as a geographical mystery. This insight could help teachers to advance students' geographical relational thinking skills. Two geographical mysteries were examined with data from 69 small groups in Dutch upper secondary education. The two mysteries resulted in differences in the level of relational thinking, which were partly explained by small-groups' on-task behaviour. Many student groups showed a low level of geographical relational thinking. Findings point to the need to incorporate exercises into geography lessons which require the use of thinking and reasoning with interrelated causal relationships.

* This chapter is based on: Karkdijk, J., Admiraal, W F., & Van der Schee, J. A. (2019). Small-group work and relational thinking in geographical mysteries: A study in Dutch secondary education. *Review of International Geographical Education Online (RIGEO)*, *9*(2), 402-425

4.1 Introduction

Teaching geography in secondary education introduces students to the beauty and fascinating variety of our planet. Geography lessons might also contribute to the development of students' individual capabilities outside the subject. For example, students learn how to live responsible lives with regard to issues such as sustainability, inequality and poverty. In their concluding remarks on the contribution of geography to the development of students' individual capabilities as human beings, Lambert, Solem, and Tani (2015) propose that teaching geography should contribute to students': (1) deep descriptive world knowledge, to know how the world works; (2) conceptual knowledge providing a relational understanding of our existence on the planet; and (3) thinking through alternative futures in specific places and locational contexts (p. 732). In this study, we focus on the second of these contributions of geography lessons. We studied differences in relational thinking between small student groups in upper secondary education in the Netherlands working together on a geographical relational assignment framed as a mystery. More insight into students' geographical relational thinking could help teachers to find ways to support students in gaining a relational understanding of their existence on the planet.

4.2 Geographical relational thinking

Relational thinking belongs to the core of geographical thinking (Jackson, 2006) and offers students a powerful way of thinking. Using descriptions of powerful knowledge borrowed from Young, Maude interprets knowledge as powerful when, among other things, it "enables students to discover new ways of thinking; better explain and understand the natural and social worlds; think about alternative futures and what they can do to influence them" (2017, p. 30). Our students live their lives in a globalised and interconnected world with many interrelated societal and ecological problems. To address these problems, students need to think in terms of interwoven causal relationships, as our contemporary problems cannot be explained by simple causalities (Arnold & Wade, 2015; Brown, 2018; Chee, 2010; DeVane, Durga, & Squire, 2010; Fögele, 2017; Lezak & Thibodeaux, 2016). Being able to understand relationships between geographical phenomena and to give integrated, holistic explanations is an example of powerful thinking (Maude, 2017). Besides this, relational thinking in geography can also be powerful in Maude's interpretation when used as analogical reasoning. This means that objects or events in the world are seen as systems of relationships that can be compared with each other to find differences, similarities and deeper structures (Richland &

Simms, 2015). For example, students can compare deforestation in Brazil for soy plantations with deforestation in Indonesia for palm oil plantations, or, as a next thinking level, between deforestation for palm oil plantations in Indonesia and gentrification in New York. Comparing these systems of relationships will reveal the similarities in the deeper economic forces between those processes and the power relationships in them that result in the forcing out of weaker parties like indigenous people or poor inhabitants. Differences will show the importance of the specific regional context.

One of the most important concepts in geography is interdependence (Lambert, 2004; Massey, 2014). “Interdependence is a potentially powerful way of recognising the scale, but also subtlety, of interactions among and between global and local processes of economic, social and environmental change” (Smith, 2015). With this focus on relationships and interaction, it is possible to understand, explain or predict changes in particular places or regions in the world. The study of the interaction between human society and its environment, in particular, has a long tradition in geography and is still the core of geography education (IGU-CGE, 2016). But places and regions are also the product of their connections with other places in the world; therefore, students need an understanding of how places or regions are connected to the rest of the world (Massey, 2014). Van der Schee’s (2000) Geographical Analysis Model for structuring geography lessons distinguishes between these two different kinds of relationships that cause regional change. Vertical relationships are relationships within a region. These are the interactions within and between natural and human geographical systems that cause regional change. Changes in one region that cause change in another region are horizontal relationships. For example, the large-scale irrigation agriculture in the south of Spain causes serious water shortages that trigger regional desertification (vertical relationships). Transportation of water from the river Ebro in the north of Spain to the south will probably solve these water shortages, but will also harm rice production in the Ebro delta (horizontal relationship).

Using Van der Schee’s model, geographical relational thinking can be described as analysing, explaining and/or predicting the relationships that cause regional change on different scales and the interactions between them. The concept of “interaction” in this description underlines the interconnectedness of the relationships. The concept of “geographical” refers to relational thinking as a part of geographical thinking, for the focus is on meaning-making with respect to regional change with the help of geographical concepts and questions (Van der Schee, Trimp, Béneker, & Favier, 2015; Uhlenwinkel, 2013, 2017). Our description of geographical relational thinking connects to systems thinking. Systems thinking, relational thinking and holistic thinking are strongly related constructs and are often used interchangeably for thinking in dynamic, interconnected wholes as opposed to reductionist thinking. Actually, holistic thinking and relational thinking refer to different elements of systems thinking. Whereas holistic thinking refers to the functioning of the system as a whole and considers the consequences of events or decisions for the system, relational thinking refers to the causal relationships within the system (Lezak & Thibodeaux, 2016).

4.2.1 Relational thinking in secondary geography education

Only a limited number of studies have included empirical work on the ability of secondary school students to establish causal relationships to explain geographical phenomena or events.

Two studies on systems thinking from the earth sciences, a related discipline to geography, provide evidence of students' difficulties with cyclic causal relationships. Kali, Orion, and Eylon (2003) conducted a study on students' systems understanding of the rock cycle. Understanding the rock cycle implies that students are able to think and reason in cyclic chains of cause and effect. After a learning program, students still had great difficulties understanding and representing these cyclic chains of causality. Assaraf and Orion (2005) analysed the progress of junior high school students in various components of systems thinking concerning the hydro cycle. At the start of the learning program, students had an incomplete picture of the hydro cycle and also showed many misconceptions, but they showed a significant increase in understanding after learning about the hydro cycle. Two main factors explained differences in this progress in students' system thinking: students' individual cognitive abilities and their involvement in the learning activities.

In a quasi-experimental study on the effects of the use of geospatial technologies on geospatial relational thinking of students in Dutch secondary geography education, Favier and Van der Schee (2014a) developed a geospatial relations test which was used as a pre-test and as a post-test. In their experiment a three-lesson series with geospatial technologies on water management in Dutch polders was compared with a conventional geography lesson series which had the same content. The students were 14 and 15 years old and in their 3rd year of havo (higher general secondary education) and 3rd year of vwo (pre-university education). The authors concluded that students identified only a proportion of the possible relationships, well below the maximum possible. Students in pre-university classes did better on both the pre-test and the post-test compared to students in higher general secondary education. Students in the experimental groups showed more progress in relational thinking than students in the control group. In the experimental groups, the effect size of their experiment was higher for lower complexity assignments than for assignments involving more complex geospatial relational thinking.

Cox, Elen, and Steegen (2017) investigated the current state of the art in systems thinking in geography of students aged 16-18 in their final and penultimate years in secondary education in Belgium. They developed a systems-thinking test in which students had to identify relevant variables and establish and describe relationships between them in order to construct a causal diagram. They found large differences between students in geographical systems thinking. Students with the highest chance of achieving an academic-oriented bachelor degree performed better than the other students and students in their final year did better than students in their penultimate year. The overall level of systems thinking in geography was, however, "problematic, keeping in mind the future global challenges this student generation will be facing" (p. 8).

This low level of relational thinking corresponds with an analysis of Dutch geography exams in higher general secondary education in 2009 and 2010, which revealed that students had particular difficulties with relational questions (Karkdijk & Van der Schee, 2012). Furthermore, research on established relationships to explain mysteries (Karkdijk, Van der Schee, & Admiraal, 2019) revealed that many small student groups neglected both the more difficult causal relationships and the interrelationships (the cross-links) necessary to give a coherent explanation for the mysteries.

The evidence of the low level of relational thinking in secondary geography education corresponds with findings of studies on relational thinking with ecosystems, which show that students have more difficulties with complex causality than with simple causality and linear flow (Hmelo-Silver, Jordan, Eberbach, & Sinha, 2017; Perkins & Grotzer, 2001).

4.2.2 Mysteries and relational thinking

In the current study, we used two mystery assignments to elicit students' geographical relational thinking. A mystery is an educational strategy originally designed within the Thinking Through Geography program to advance students' relational thinking and reasoning skills by investigating complex situations (Leat, 2001; Leat & Nichols, 2003). A mystery is centred around a problem formulated as an open question with more than one answer or solution. After a short introduction to the whole class, students have to explain or solve the mystery in small groups. They use 16-30 information strips that contain some information about a person's life, facts, background information and some red herrings. These strips contain no relationships: students are challenged to think about and discuss relationships between them. A student has to explain to the other group members the reasons why certain strips should be seen as categories or as related to each other. In this way, a mystery triggers shared reasoning amongst the group members (Leat & Nicols, 2003).

Leat and Nichols (2000b, 2003) observed the explanation of a mystery in small groups and found five stages that characterise this. After the reading (1) and categorisation of the data (2), most groups created one or several separated or more integrated cause and effect chains (3). In the next stage, the higher ability groups reworked their explanation by formulating new sets of relationships, which were increasingly abstract, and incorporated more data in order to give a more coherent explanation (4). The highest ability groups moved on to the last stage: they hypothesised and generalised beyond the given data (5). The authors suggest that these stages coincide with the progression in the thinking processes of the levels of the SOLO taxonomy (cf. Biggs & Collis, 1982).

4.2.3 The SOLO taxonomy and relational thinking

Biggs and Collis (1982) constructed the SOLO taxonomy in an attempt to measure the quality of learning as represented by the outcome of a learning process. They identified five levels that indicate a progression in the structural complexity of an outcome: each higher level compared to the former is more abstract, uses more organising dimensions and shows more internal consistency and coherence due to established interrelationships. The highest level of the taxonomy is characterised by using self-generated principles. In a *pre-structural* response no relevant datum is used to answer a question; in a *uni-structural* response a student uses correctly one relevant datum to answer the question; in a *multi-structural* response several correct but unrelated data are used to answer the question; in a *relational* response a student is able to use several interrelated data to give a coherent explanation; while in an *extended abstract* answer, abstract principles beyond the given data are formulated.

Stimpson (1992) conducted a study on the relationship between the levels of the SOLO taxonomy and geographical thinking of secondary students in Hong Kong and found evidence that the progression of these levels reflects an increase in the level of geographical thinking as in the what-where-why-sequence in geographical questions. He also found that 12-13 year-old students in the first year of secondary education operated mostly on an uni-structural or a multi-structural level, while many fourth-year students had started to operate on a relational level, demonstrating relational thinking.

4.3 Research questions

The current study aimed to gain more insights into secondary school students' geographical relational thinking. To this end, we applied small-group work to a geographical mystery. We designed two mysteries to elicit collaborative relational thinking and reasoning. One was about floods in Jakarta and another about landslides and slum dwellers in Rio de Janeiro. The two mysteries are presented in the method section and are included as Appendix B. Our first and second research questions were:

(1) How does geographical relational thinking in terms of the SOLO taxonomy differ between groups?

(2) How does geographical relational thinking in groups differ between the two mysteries?

Leat and Nichols (2003) suggested the influence of students' age, group ability (mixed or more homogeneous) and gender (male or female groups) on the outcomes of mysteries as an area of educational research. Explaining complex problems such as a mystery in small groups also requires effort and focused attention. We therefore expected that the time students spent on-task would also be a factor explaining differences between student groups. Our third research question was therefore:

(3) How can differences between groups in geographical relational thinking be explained by characteristics and collaborative behaviour of the groups?

4.4 Methods

4.4.1 Research design

A quantitative research project was carried out on a relational task framed as a mystery. We used the concept maps constructed by small student groups and the transcriptions of their group discussions. The project was carried out between January and June 2015 in six secondary schools in the Netherlands.

4.4.2 Sample

Twelve qualified and experienced geography teachers from six schools and 205 students in higher general secondary education (havo, 4th and 5th years) and pre-university education (vwo, 4th, 5th and 6th years) were part of the project. These teachers and their schools were selected, because they responded positively on a call to participate in our research project. The schools are located in different parts of the Netherlands: two schools are located in small cities in a rural, less densely populated region in the south-western part of the Netherlands (Goes and Middelharnis), three schools are located in cities in the most urbanised and densely populated region in the western part of the Netherlands (Rotterdam, Gouda and Hilversum) and another school is located in a small city in the less densely populated north-western part of the Netherlands (Hoorn). Students were between 15 and 18 years old. Teachers were instructed to arrange three groups of three students based on the students' geography grades: (1) one group of students belonging to the highest 30% of the class; (2) one group of students belonging to the lowest 30% of the class; and (3) one mixed group (one student from the highest 30% and two from the lowest). In some cases, because of the absence of selected student(s), an intermediate ability group was created (with two or more students between the highest thirty per cent and lowest 30% of the class). Each class worked on one of the two mysteries. Table 4.1 gives the composition of the 35 groups who worked on the Rio mystery and the 34 groups who worked on the Jakarta mystery.

Table 4.1 Description of student groups.

	Rio	Jakarta
Total number of groups	35	34
Geography grades		
Groups in highest 30%	11	11
Intermediate groups	7	4
Groups in lowest 30%	9	12
Mixed groups	8	7
Gender		
3 girls	5	11
Mixed groups	23	18
3 boys	7	5
Educational level and year		
havo-4	17	3
havo-5	0	19
vwo-4	9	3
vwo-5	6	6
vwo-6	3	3

4.4.3 Data collection

The two mysteries that were part of the assignment were designed for students in upper secondary education as a regional event that challenged them to use complex causal relational thinking and reasoning, although each mystery had its own regional context and problem. The regional context of both mysteries is a mega city in two developing countries (Rio in Brazil and Jakarta in Indonesia), but the Rio mystery was designed with a human geographical focus and the Jakarta mystery with more emphasis on physical geography, to capture the breadth of geography as a school subject. The geographical content of both mysteries is part of the curriculum of the previous years in higher general secondary education and pre-university education in the Netherlands. They were designed in line with the design principles of Leat & Nichols (2003), reviewed by a professional educational geographer familiar with mysteries and tested in upper secondary geography classes. Groups had to represent their explanation of the mystery as a concept map consisting of causal relationships. Each group worked in a separate room, supervised by the researcher (the first author). Their discussions were recorded by a video camera. The rest of their class discussed the same mystery in small groups in the classroom, guided by the teacher. The group discussion was used together with their draft concept map to construct a final concept map that included all correct and relevant relationships the group had formulated.

Both mysteries were based on real data, gathered from a wide variety of sources. One was about Fabio Pereira, a favela dweller in Rio who refused to move to a new dwelling on the outskirts of Rio. The government offered him this new place because his favela was threatened by landslides. Students had to think about the geography and society of Rio, in order to understand and explain why Fabio wanted to stay in his old threatened house. Factors that explained Fabio's decision were: his present house was built by his father; the community bonds in his present favela; the current proximity to his work in the centre of Rio; the high real estate prices in central Rio; and former government actions to remove favelas which had fuelled distrust in government intentions with regard to the rehousing of slum dwellers; and the football World Cup in 2014 and the Olympics in 2016 in Rio, which is also a famous tourist city. A core element of the mystery was Fabio's distrust in the intentions of the government, because his favela, relatively close to the central business district of Rio, could become gentrified and transformed into a more profitable district. Students had to make this inference by themselves based on the information provided.

The second mystery was about slum dwellers on the riverbeds in Jakarta, like Nani, who were accused by a high government official of being the main cause of regular floods in Jakarta. Students had to understand and explain the hydrological system of Jakarta in order to evaluate this accusation. Factors that explained Jakarta's vulnerability to floods were: heavy monsoon rains; deforestation of the surrounding region; the construction of slums in the beds of the main rivers in Jakarta by migrants and the resulting narrower riverbeds; blockage of rivers by waste and badly managed water channels; Jakarta's location in a delta near the sea; and the rising sea level. For a complete and coherent explanation of the mysteries, students had to establish interrelationships between the different factors.

Both mysteries were geographical mysteries, for the focus was on regional change, human-environment interaction and on the significance of a particular location within a specific regional context.

4.4.4 Measures

4.4.4.1 Relational thinking

Groups had to create a concept map that represented their explanation. A concept map consists of relationships between concepts (propositions) and provides a representation of the knowledge structure of the student group: it provides insight into how students understand relationships between concepts (c.f. Novak & Gowin, 1984; Ruiz-Primo, Shavelson, Li, & Schultz, 2001; Srinivasan, McElvany, Shay, Shavelson, & West, 2008). Creating a concept map as a collaborative activity also elicits and stimulates reasoning within a group (Van Boxtel et al., 2002; Cox, Steegen, & Elen, 2018).

Each group was given a piece of paper, pen or marker and short instructions on how to construct the concept map. No concepts were provided and all propositions had to be formulated by the student groups in their own words. They could use the information provided and were also allowed (and encouraged) to add extra concepts necessary for their explanation. We intended the task to be as student-led as possible. Compared to fill-in-the-map techniques, this construct-a-map technique better reveals differences between the knowledge structures of students, for the low-directedness of the task offers students more opportunities to show what they know about a specific topic and to express misconceptions and partial understandings (Ruiz-Primo, Shavelson et al., 2001). The arrows students were required to draw between concepts had to express causal relationships and had to be provided with linking phrases that expressed the relationship. This meant that their concept maps could be characterised as causal schemes and did not necessarily have a hierarchical structure.

We wanted to use all the correct and relevant relationships a group formulated in our analysis. However, most students were unfamiliar with the construction of concept maps, so there was a risk that not all of the correct and relevant relationships a group identified would be represented in the concept map. Assaraf and Orion (2005) observed that some students who presented relationships in other tools preceding the construction of a concept map were not able to represent these relationships in the concept map itself. For this reason, we also gathered the data from the group discussions to be analysed in combination with the concept maps. First, the group's concept map was analysed and scored for relevance and accuracy. Then the same was done for the data from the group discussion, which was video-taped and transcribed verbatim. Relevant and correct relationships that were rejected later in the group discussion were deleted from the data set. The data from the discussion and the concept map were merged into a total data set that was used to construct the final concept map of a group. We used a criterion map as a benchmark in the coding process. The criterion map for each mystery was dynamically constructed, because of the task requiring the students to formulate relationships in their own words and to use extra concepts besides those provided. To construct the mystery, the designer (first author) used different factors or dimensions of the problem, which were also used as a frame for the criterion map. Both criterion maps were discussed with two other raters (experienced geography teachers who were familiar with the content of the mysteries) and reworked. Analysis of the outcomes of the group work revealed that groups often established correct and relevant extra concepts to explain the problem. We added the new propositions to the final criterion maps. These criterion maps are included as Appendix C.

Four aspects of the total data set for each group were used as indicators to analyse relational thinking in a group. Three were based on research on analysing concept maps (McClure, Sonak, & Suen,1999; Rye & Rubba, 2002; Turns, Adam, & Atman, 2000): (1) the accuracy of the propositions; (2) the number of cross-links; and (3) the number of factors used (in this study factors are the different dimensions to explain the mystery). The number of accurate propositions reflects the depth of understanding, while cross-links indicate the connectedness of understanding, necessary for a coherent explanation. The number of factors used to explain the mystery refers to the breadth of understanding.

Table 4.2 Scoring system of the group concept maps.

Indicators	score
Total proposition score	(a + b + c)
a. each correct and relevant proposition	2
b. each correct and relevant proposition, but intermediate concept(s) missing	1
c. each correct and relevant proposition, but unclear formulation	1
Number of cross-links	total number
Number of factors used	total number

In addition to the data on the three indicators shown in Table 4.2, the SOLO level of the outcome of each group was determined as a fourth indicator of students' relational thinking. The SOLO taxonomy has five levels as described in section 4.5 and between these levels there are transitional levels. To determine the SOLO level of the outcome of each group, the number of factors which were correctly connected with the outcome of the mystery, the use of branches within factors, the number of cross-links and the use of abstract reasoning or transfer to other regional contexts in the group discussion were used as criteria. Our operationalisation accurately followed the way Biggs and Collis (1982) characterise responses at different SOLO levels for geography assignments, and this has been described elsewhere in detail (Karkdijk, Van der Schee, & Admiraal, 2019).

The first author coded all the data. Two other raters coded the concept map and discussions of four groups each. One other rater analysed the data of the four groups working on the Rio mystery and another rater coded the four groups working on the Jakarta mystery. The inter-rater reliability between the scores of the first author and the two raters was satisfactory (Cohen's k = 0.823). After instruction and several try-outs and discussions on the determination of the SOLO level of the outcome of each group, an experienced and qualified geography teacher analysed the work of 12 groups as a second rater and found the same SOLO levels as determined by the researcher.

4.4.4.2 Group characteristics and group effort

The educational level of the groups was higher general secondary education (havo) or pre-university education (vwo) and students were in their 4th, 5th or 6th year. Teachers provided students' mean geography grades based on their grades starting from the 4th year. We standardised these grades to Z-scores for each class, from which we calculated mean Z-scores for the groups. The gender composition of the groups included girls-only groups, boys-only groups and mixed groups (see Table 4.1).

To analyse group effort to explain the mystery, we coded the group discussions on on-task, procedural and off-task words. The on-task category contained that part of the discussion aimed at understanding the content of the mystery or designing a strategy to solve it. The procedural category comprised those parts of the discussion aimed at organising the task without discussing the content of the strategy ("We need paper"; "Who is going to write?"; "How much time do we have?" or "What are we supposed to do?"). Off-task words were all the utterances irrelevant for the task. As indicators of group effort, we used:

1) the proportion of off-task words of all words uttered from the start of the group discussion prior to the construction of the concept map; and
2) the number of on-task words uttered from the start of the group discussion prior to the construction of the concept map.

4.4.5 Data Analysis

We used the SOLO taxonomy of Biggs and Collis (1982) to analyse differences in relational thinking between groups (first research question). A series of independent t-tests on the indicators of relational thinking (total proposition score, number of cross-links and number of factors used) were carried out to test differences in relational thinking between both mysteries (second research question). To answer the third research question, four hierarchical multiple regression analyses (MRAs) were performed with the total proposition score, the number of cross-links, the number of factors and the SOLO levels of the groups used as dependent variables. The four MRAs were performed for both mysteries together. We standardised the variables total proposition score and factors used to Z-scores for each mystery. Standardisation of the number of cross-links was not necessary, because means for both mysteries did not differ significantly. The MRAs were performed with five independent variables as predictors in the following sequence: educational level, school year, standardised mean geography grades of the groups ("mean geography grades"), proportion of off-task words of all words uttered prior to the construction of the concept map ("proportion off-task words") and finally the number of on-task words uttered prior to the construction of the concept map ("number on-task words"). The categorical variables educational level and school year were transformed into dummies. Because a majority of the groups were mixed boys' and girls' groups, gender was not included in the analyses.

As the group data were nested within teacher and school, multilevel variance components analyses were carried out but they did not indicate significant variance in three of the four dependent variables at either school or teacher level. For this reason, all subsequent regression analyses were performed at the group level only.

4.5 Findings

First the differences between the groups are presented in section 4.8.1. Thereafter the differences between the two mysteries in groups' geographical relational thinking are shown in section 4.8.2 and finally the results of the multiple regression analyses are given in section 4.8.3.

4.5.1 Differences in relational thinking between groups

We used the SOLO taxonomy to analyse differences in relational thinking between the groups. Table 4.3 shows the distribution of the groups on SOLO levels. We found no groups with an outcome on the prestructural level: all groups were able to find at least one relevant and correct relationship to explain the mystery. One group connected only one factor to the problem of the mystery accurately (unistructural level). Twenty-two groups with an outcome on the multistructural level used two or more factors to explain the mystery, but without establishing any cross-link between these factors. Twenty-four groups showed an outcome on the relational level: these groups accurately connected four or more factors to the problem of the mystery and established at least one cross-link between these factors and one branche within a factor. This means that these groups provided a more or less integrated, coherent explanation of the problem of the mystery. Only one group was able to give a coherent explanation for the mystery as well as to use abstract reasoning in their discussion: this group showed an outcome on the extended abstract level. Twenty-one groups had a transitional level.

To illustrate these differences in relational thinking between groups, three examples will be described: (1) a group (22) with an outcome on the multi-structural level; (2) one (64) with an outcome on the transitional level (between multi-structural and relational); and (3) a group (33) with an outcome on the relational level. All three groups worked on the Rio-mystery. To

Table 4.3 Distribution of groups on SOLO-levels.

SOLO-level	Total	Jakarta	Rio
Pre-structural P	0	0	0
Uni-structural U	1	0	1
Transitional U/M	4	1	3
Multi-structural M	22	11	11
Transitional M/R	13	7	6
Relational R	24	11	13
Transitional R/EA	4	3	1
Extended Abstract EA	1	1	0
Total number of groups	69	34	35

facilitate comparison of the student groups, we revised all their propositions (from both their concept maps and group discussion) into the format of our criterion map. The additional propositions from the group discussions are represented with dotted lines.

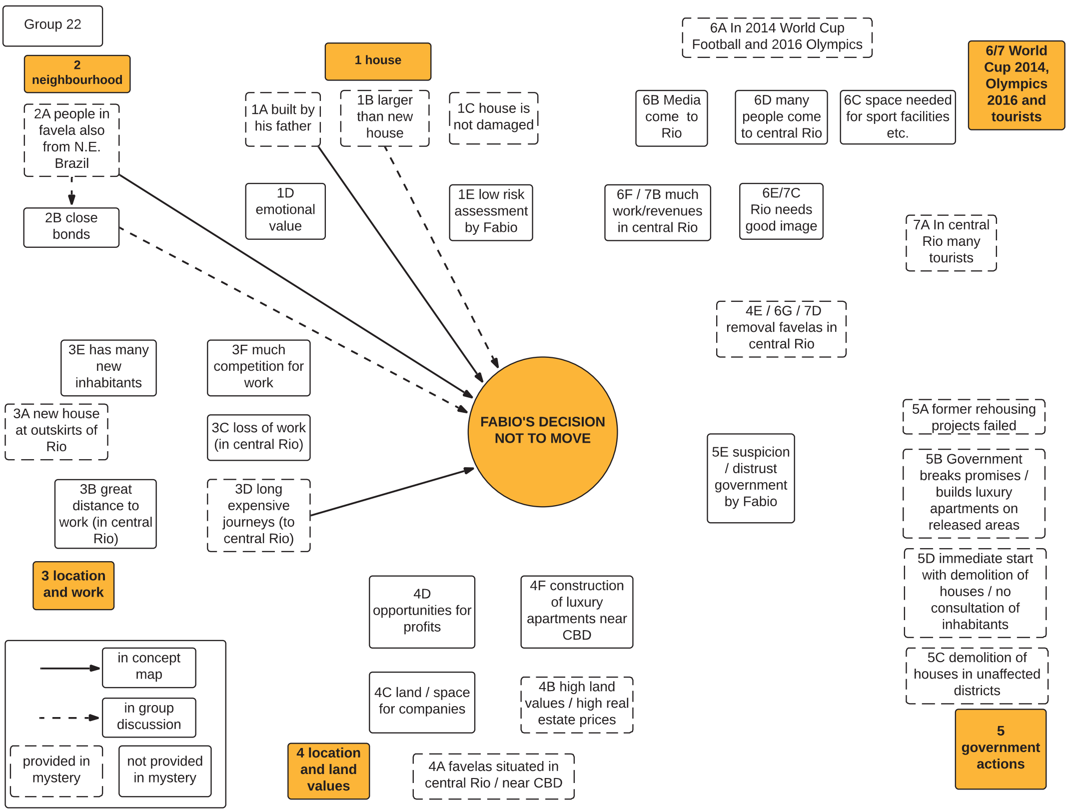


Figure 4.1 Propositions from concept map and group discussion of group 22.

The concept map of group 22 is shown in Figure 4.1. This group consisted of two girls and one boy aged 15-16 from a fourth year havo class. It was an intermediate ability group: two students had an intermediate position in their class and one student belonged to the highest 30% as regards their mean geography grades. This group had an on-task discussion, albeit relatively short: they started to work out their solution in the concept map after 14 minutes of discussion. As can be seen in Figure 4.1, the outcome was on the multi-structural level: the students in this group used only three, not interrelated, factors (1, 2 and 3) to explain the mystery. Their explanation of the mystery was very limited. Their total proposition score is 8 (three correct propositions, one incomplete proposition and one unclearly formulated proposition). The group used three factors to explain the mystery (the score for factors is therefore 3) and made no cross-links between these factors (the score for cross-links is therefore 0).

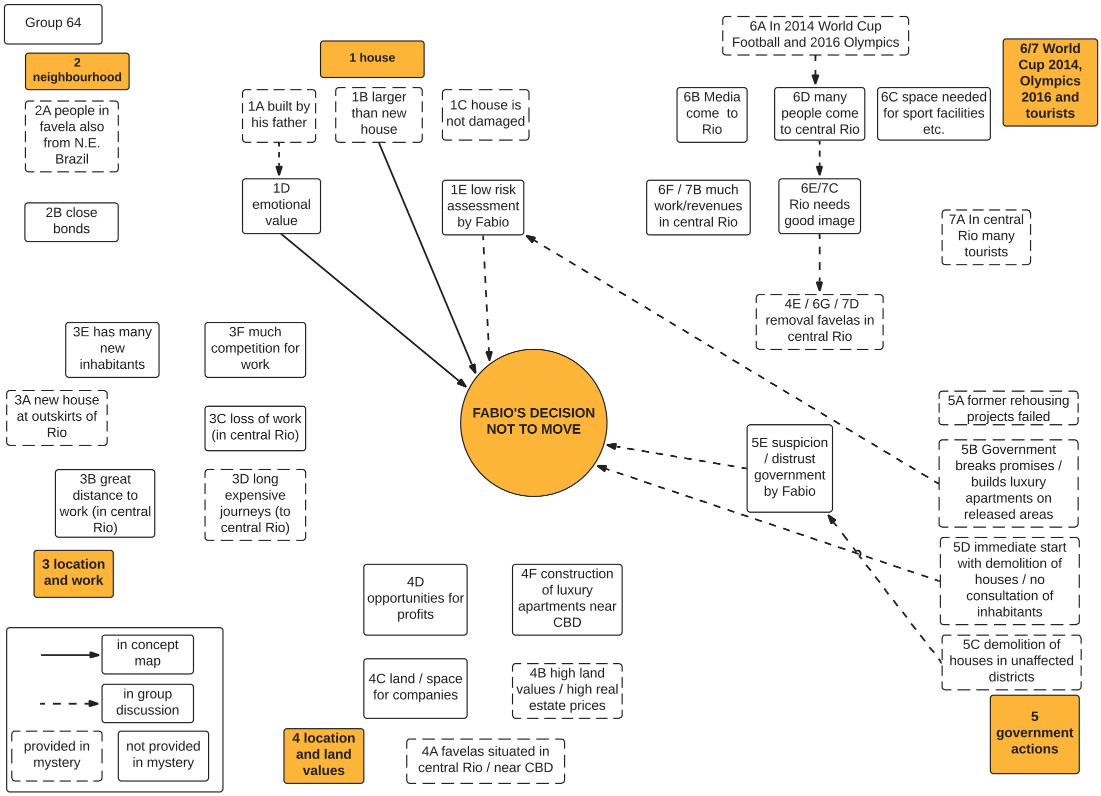


Figure 4.2 Propositions from concept map and group discussion of group 64.

The second example is group 64 (see Figure 4.2). This was a low-ability group from a fifth-year vwo class. The group consisted of two girls that belonged to the lowest 30% of their class as regards their mean geography grades and one boy from the highest 30%. The students were 16-17 years old. This group had problems constructing a representative concept map, so most propositions were added from the group discussion. They also had an on-task discussion but this took much longer than the first group: they started their concept map after 27 minutes of discussion. Figure 4.2 shows that this group was able to establish a cross-link between two factors: governments' building activities in former released areas and Fabio's risk assessment for his favela. The group used three factors to explain Fabio's decision. However, one factor, "World Cup, Olympics and tourists", with some correctly formulated propositions, was not connected to Fabio's decision and therefore not integrated into their explanation. With only two factors correctly connected to Fabio's decision and one cross-link, the outcome is on the transitional level between multi-structural and relational. The group gave a more integrated, albeit narrow, explanation for Fabio's decision than the former group. Their total score was 18.

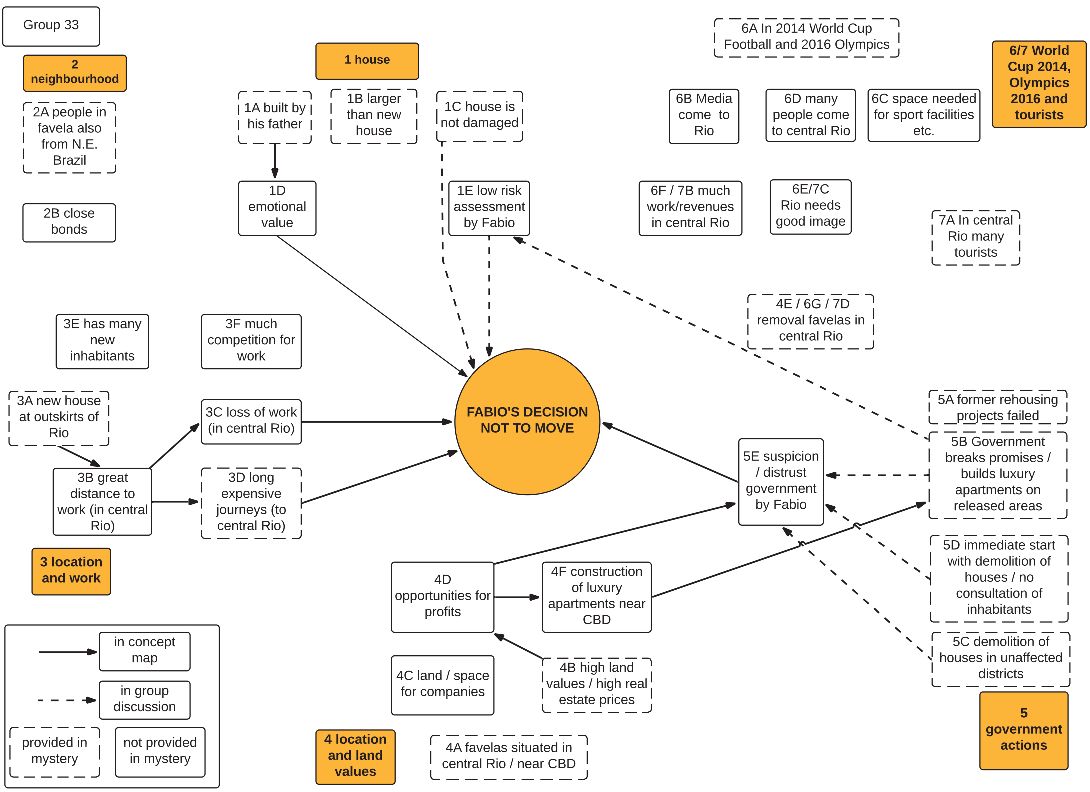


Figure 4.3 Propositions from concept map and group discussion of group 33.

The third example is group 33 (see Figure 4.3). This was a low ability group from a fourth-year vwo class: the students, three girls aged 15-16 , belonged to the lowest 30% of their class as regards their mean geography grades. They had an extensive on-task discussion and started with their concept map after 29 minutes. As can be seen in Figure 4.3, their explanation is more extensive and coherent than the other groups: they established more propositions, used more factors and established three cross-links, illustrating their understanding of the interrelationships between factors. They established the interrelationship between the opportunities to gain high profits by building luxury apartments on the future released area and Fabio's distrust of the intentions of the government. Another interrelationship they identified was between the opportunities to gain profits and the construction of luxury apartments on released areas near the CBD. The third cross-link was between governments' building activities in former released areas and Fabio's risk assessment for his favela. Their discussion was limited to the problem of the mystery and no transfer to other contexts or generalisations were made. The group outcome is therefore on the relational level. The group concluded that Fabio's distrust of the government was crucial to understanding his decision not to move. Their total proposition score was 36.

Comparison of Figures 4.1-4.3 shows the differences between the groups in their geographical relational thinking skills. Whereas group 22 was able to formulate only simple, linear relationships to explain Fabio's decision, group 33 showed an ability to reason with more complex relationships to explain his decision. They clearly provided a more integrated, coherent explanation.

4.5.2. Differences in relational thinking between the two mysteries

Table 4.3 shows that the distribution of the student groups on the SOLO levels was similar for the Jakarta and Rio mysteries. Table 4.4 shows the differences between the mysteries on the other three indicators of relational thinking. The number of propositions students formulated was significantly lower in the Rio mystery than in the Jakarta mystery ($t(67) = 2.72$; $p = .008$; Cohen's $d = .66$). With respect to the number of factors used, the difference between the two means was also significant ($t(51.29) = 6.67$; $p < .001$; Cohen's $d = 1,61$). No difference between the two mysteries was detected with respect to the number of cross-links ($t(67) = 1.01$; $p = .32$; Cohen's $d = .25$).

Table 4.5 shows the correlations between these three indicators of relational thinking for the two mysteries. In the Rio mystery (the shaded cells below the diagonal), the three indicators were positively correlated. Thus, for the Rio mystery, groups that established one or more cross-links,

Table 4.4 Means and standard deviations of three indicators of relational thinking for both mysteries.

	Jakarta	Rio
N	34	35
Total proposition score		
Mean*	26.59	20.83
St. deviation	7.62	9.82
Min	15	8
Max	49	49
Number of cross-links		
Mean	1.18	0.91
St. deviation	0.99	1.15
Min	0	0
Max	4	4
Number of factors used		
Mean*	5.24	3.74
St. deviation	0,61	1.17
Min	4	2
Max	6	6

*Differences are significant (2-tailed with α=0.05)

also formulated more correct propositions obtained from more factors, while groups who did not establish a single cross-link, usually formulated fewer propositions. This could be an indication that higher scores for the Rio mystery were obtained by a more integrated understanding of the mystery problem. For the Jakarta mystery the scores obtained for the three indicators of relational thinking were less correlated.

Table 4.5 Correlations between the three indicators of relational thinking for the Jakarta mystery (above the diagonal) and the Rio mystery (below the diagonal).

		Jakarta (N=34)		
	Indicator	Proposition score	Cross-links	Factors used
Rio (N=35)	Proposition score		0,261	0.481*
	Cross-links	0.625**		0,079
	Factors used	0.804**	0.464*	

*p ≤ .05 **p ≤ .01

4.5.3 Group characteristics, group behaviour and relational thinking

The results of the regression analyses for each of the four indicators of geographical relational thinking are summarised in Table 4.6. The only significant predictor of the standardised total proposition score in the final regression model was the on-task behaviour of the students (β = .43; sr^2 = .16): the number of on-task words they used before they started to construct their concept map. This model explained significantly 13% of the variance in the total proposition score (F = 2.63; p < .05). For the number of cross-links, again the only significant predictor in the final regression model was students' on-task behaviour (β = .26; sr^2 = .06). The proportion of variance in the total number of cross-links explained by the final regression model (5%) remained statistically non-significant. In the final regression model, there were no predictors significantly explaining differences in the standardised total number of factors groups used. Finally, we performed a hierarchical multiple regression analysis for the variance in the SOLO levels of the groups with the same predictors and found no different results: again only the number of on-task words uttered prior to the construction of the concept map was a significant predictor in the final model (β = .39, sr^2 = *.13*). The final regression model did not significantly explain the variance in the SOLO levels of the groups.

Table 4.6 Results of the final regression models of four Hierarchical Multiple Regression Analyses predicting the standardised total proposition score, total number of cross-links and the standardised number of factors used to explain a mystery, by group characteristics and group effort (N=68).

Dependent variables	*Total proposition score*		*Total number of cross-links*		*Total number of factors*		*SOLO level*	
Predictors in final model	*β*	sr^2	*β*	sr^2	*β*	sr^2	*β*	sr^2
Educational level	-.001	.00	-.075	.00	.070	.00	-.138	.01
School year, dummy, 5,4=0	-.062	.00	.111	.01	.019	.00	.036	.00
School year, dummy, 6,4=0	.092	.01	.177	.02	-.153	.02	.116	.01
Mean geography grades	.063	.00	-.073	.01	.187	.03	.099	.01
Proportion of off-task words	-.034	.00	-.219	.04	.138	.02	-.079	.01
Number of on-task words	.431	.16***	.259	.06*	.152	.02	.391	.13**
Adjusted R^2	.127		.047		-.006		.090	

*$p \leq .05$ **$p \leq .01$ ***$p \leq .001$

4.6 Discussion

With respect to the first research question, we found substantial differences between groups with respect to their relational thinking skills. Many groups had serious difficulties with relational thinking: nearly 40% of the groups operated on the multi-structural level or lower, so they were not able to give an integrated, coherent explanation. Yet almost half of the groups operated on the relational level or higher: they were able to establish one or more cross-links and gave an integrated, coherent solution. The difficulty that many groups had with relational thinking is in line with the findings of the studies on geographical relational thinking we described in the introduction section.

With respect to the differences in relational thinking between the two mysteries (research question 2), we found that the Rio mystery was more difficult for students than the Jakarta mystery, based on lower mean scores overall. The problem of the Jakarta mystery was probably more familiar to Dutch students, because floods and hydrological systems are common topics in geography lessons in lower and higher secondary education in the Netherlands. Moreover, our analysis revealed that the design of the Rio mystery was more complex for students, because it required them to think more in terms of interconnections between factors. Most groups that made connections between different factors to explain the Rio mystery also established more relevant and correct propositions and used more factors. The information strips on Jakarta contained more pieces of information that were relatively easy to connect linearly to floods: heavy monsoon rains, a lot of waste in the rivers, smaller riverbeds, slums in riverbeds and a rising sea level. It was therefore easier to obtain a relatively high total proposition score without establishing the interconnections between factors. These findings suggest that the design of a mystery affects the level of relational thinking.

Concerning the third research question, we found that the only significant predictor that explained a small proportion of the variance of three indicators of relational thinking was students' on-task behaviour during the group discussion: the amount of on-task discussion they had on the problem of the mystery before they started to construct the concept map. More extensive discussion on the different aspects of the problem before integrating these relationships into a concept map seemed to be more fruitful than quickly starting with the construction of the concept map. Leat and Nichols (2000b, 2003) also found that groups that were willing to rework their first constructed web of relationships integrated more information into their explanation. Unlike previous studies on geographical relational thinking in secondary education, we did not find significant correlations between educational level, grade or previous achievement in geography and relational thinking. A possible explanation could be that the studies we reviewed gathered their data from individual students whereas our study used data from small groups. There is a lot of evidence that group interaction, the strategy groups employ or the quality of the talk in the discussion affect the outcome and therefore could also account for part of the variance of the total score (see for example: Barron, 2000; Goos, Galbraith, & Renshaw, 2002; Mercer, Wegerif, & Dawes, 1999; Mortimer & Scott, 2003; Ruiz-Primo, Shavelson, et al., 2001). A qualitative analysis of the group discussions might deepen our understanding of geographical relational thinking in small groups.

4.6.1 Limitations and future research

This research study consisted of a relatively large-scale quantitative investigation of relational thinking by students working together in collaborative groups to solve a mystery. A limitation of our study was the small sample size. Sixty-eight groups in a multiple regression analysis with more than two predictors does not allow firm generalisations to be made. Either large-scale quantitative research or qualitative research on group discussions of students working in small groups to solve mysteries would be future directions to be taken to gain further insights into students' geographical relational thinking.

We found a first indication that different designs of mysteries may result in different levels of relational thinking. We recommend more research on the relationship between the particular design of a mystery and the level of relational thinking produced. This could give insights into how to design mysteries more effectively to foster relational thinking.

More in general we recommend research on students' relational thinking in secondary geography education in a wide range of countries, not only to get more evidence on students' relational thinking skills, but also to facilitate international comparison and to learn from each other's practices.

4.7 Conclusion and recommendations

The difficulties that many groups from the six schools in the Netherlands had with geographical relational thinking point to the need for more exercises in Dutch geography lessons that practise this thinking skill. The absence of any significant relationship between educational level, school year or students' geography grades and indicators of relational thinking might mean that relational thinking needs to receive much more attention in Dutch geography classes. Although our study was limited to Dutch secondary geography education, the (scarce) evidence we found on students' relational thinking in the earth sciences, geography and biology suggests that relational thinking, often mentioned as a core geographical thinking skill (e.g. IGU-CGE, 2016; Lambert, 2004; Uhlenwinkel, 2013), also might need more attention in geography lessons in other countries.

To foster students' relational thinking in geography lessons, exercises and assignments that present complex regional problems to students and not only linear relationships are needed. This would be a promising shift of focus, because most studies we reviewed provided evidence that an explicit focus on relational thinking fosters students' thinking in interconnected causal relationships. Complex relational thinking exercises can also develop a "systems disposition", a set of attitudes to be focused on thinking in complex relationships (DeVane, Durga, & Squire, 2010). Asking students to represent their relational thinking as a causal scheme or concept map has proved to be very helpful in getting them to think in terms of webs of relationships. Renshaw and Wood (2011) found that the concept map assignment reoriented students from linear causality chains to thinking in terms of interdependences. A whole-class discussion guided by a skilled geography teacher after each exercise in relational thinking can enhance students' understanding. Explicit attention for relationships *and* cross-links in this whole-class discussion would be helpful for a holistic, interrelated understanding of the mystery.

A step-by-step approach could use less complex exercises like the Jakarta mystery first to practise relational thinking and then move on to more difficult mysteries like the Rio one. These could be used as the next step in geographical relational thinking to focus explicitly on interconnections to understand a regional problem. When students have more competence in thinking in systems of relationships, analogical reasoning as described by Richland and Simms (2015) is a promising tool for geography teachers. It can deepen students' insights into how the world works, in the deeper structures of our existence on the planet.

Learning to think in systems of relationships is expected to foster students' relational understanding of their existence in this world, and would therefore be helpful in the choices they have to make to live responsible lives. If geography lessons contribute to the development of individual capabilities by delivering this powerful knowledge to our students, teaching geography will be a very relevant undertaking in secondary education.

Chapter 5

Strategies used by small student groups to understand a geographical mystery*

Abstract

Relational thinking in geography is often complex, due to the interdisciplinary character of the subject and the many relationships between human and natural systems. We explored the strategies of twelve small groups of students in upper secondary education in the Netherlands as they attempted to understand a regional problem presented as a mystery. Four different relating strategies were found. The six low-performing groups on the mystery assignment employed different relating strategies from the six high-performing groups, who primarily used a webbing strategy. The findings suggest that a webbing strategy, focused on the establishment of multi-causal relationships, is more successful in tackling complex assignments in geography such as understanding regional problems.

* This chapter is based on: Karkdijk, J., Van der Schee, J. A., & Admiraal, W. F. (2021). Strategies used by small student groups to understand a geographical mystery. *J- READING. Journal of research and didactics in geography, 10*(1), 5-21.

5.1 Introduction

Thinking about geography, the question could arise: why should we teach it? In our globalised and interconnected world, with big issues such as international migration, climate change, inequality and environmental destruction, what has the school subject which teaches "the world" to offer to young people? It teaches students how people at specific locations on the planet, in different circumstances, try to make their living and how their interrelationships with people at other places and their interaction with the environment form places and landscapes (IGU-CGE, 2016). Geography can help students to understand the world by studying it in a specific way, using key geographic concepts. This is called "the geographic advantage" (Hanson, 2004) or "thinking geographically" (Jackson, 2006; Lambert, 2017). Lambert (2017) states that relational thinking lies at the heart of thinking geographically (p. 28). He uses a slightly modified version of Jackson's (2006) four pairs of key concepts that capture relational thinking in geography: space and place, scale and connection, proximity and distance and people and environment (p. 28). Relational thinking is also at the heart of the Geographical Analysis Model of Van der Schee (2000). This model describes relational thinking from the perspective of regional change. The model is structured by the relationship between man and nature and the unit of analysis is the region. It analyses regional change with vertical relationships (between elements of the natural system and the human system within a region) and horizontal relationships (between regions at different scales). The German Geographical Society DGfG (2012) identifies the systems concept as the core concept of the school subject and has developed a cube model, which describes the human-environmental system containing the human and the physical subsystems. The DGfG mentions the relationships within and between these (sub)systems as "the central object of study" (p.10) at different scales (from the local to the global). Thus, relational thinking in geography is connected to systems thinking but is structured by the use of key geographical concepts, has a particular focus on the analysis of relationships between man and nature and uses the region as a localised unit at different scales (Favier & van der Schee, 2014a; Cox, Elen, & Steegen, 2019).

Clearly, relational thinking in geography is hard for students, because they have to analyse the interconnections between economic, political, socio-cultural and physical aspects at different scales. Understanding the interconnected problems of our world cannot be achieved by thinking in simple causalities: students need to think in interwoven, complex causalities (DeVane, Durga, & Squire, 2010; Fögele, 2017). Research in the earth sciences and geography has shown that students in secondary education have difficulties with complex, multi-causal relational thinking (Assaraf & Orion, 2005; Cox, Elen, & Steegen, 2017; Favier & Van der Schee, 2014b; Kali, Orion, & Eylon, 2003). Consequently, there are large differences between students in their geographical relational thinking. Mehren, Rempfler, Buchholz, Hartig, and Ulrich-Riedhammer (2018) validated a three-stage competence model of systems thinking in geography. Students on a low level (competence stage 1) identified only some elements that were hardly linked. Established relationships were mono-causal or linear. At stage 2, relationships became more complex as linear thinking predominated. At stage 3, networked elements and relationships were identified, illustrating complex relational thinking.

This study examines the different relational strategies that were used by groups with low performance in relational thinking and by groups that performed highly in relational thinking when they had to explain a multi-causal regional problem that was presented as a mystery.

5.2 The mystery as a thinking strategy

The mystery is one of the most powerful strategies from the Thinking Through Geography program (Leat, 2001, p.51), which was designed to foster thinking in geography lessons, although this thinking is more generic than thinking geographically (Morgan, 2017). Mysteries are known internationally and used in different school subjects (geography, history, economics). They are designed to promote relational thinking and reasoning in small groups in order to gain insight into a real-life problem that seems odd at first glance (Leat, 2001; Schuler, 2005). In geographical mysteries, these problems are localised in specific regions. In addition to the whole class introduction to the mystery and some additional materials like maps or photographs, students are given 16-30 strips that each contain a piece of information: background data, personal information about a person or people involved, concrete information on the problem situation, the geographical context and some red herrings. Students have to make connections (multiple causal relationships) between relevant information strips in order to understand the regional situation and to explain or solve the problem presented in the mystery (Leat, 2001).

Leat and Nichols (2000b) describe several stages in the process of tackling a mystery. The first stage is the *display stage*: students read the strips and try to understand the information. The second stage is the *setting stage*, where groups categorise the data and form sets and subsets of the information strips. All groups form a reject pile of strips that contain information they consider irrelevant or do not understand. Low-performing groups have larger reject piles than high-performing groups. The third stage is the *sequencing and webbing stage*. In this stage, groups relate sets and individual strips in order to understand the mystery. This relating can have a linear causal pattern (sequencing) or a non-linear pattern (webbing). According to Leat and Nichols, some groups are willing to break up their first relational chains or webs and to reconsider the reject pile to search for a better understanding and integration of more strips in a modest or a radical *reworking stage*. Leat and Nichols (2000b) found that many low-performing groups do not rework their data at all, because they fear losing the relationships they have already established. The final stage, *the abstract stage*, is only reached by some high-performing groups as they continue the discussion after the physical manipulation of strips. Leat and Nichols note that not all groups go through all five stages in an orderly manner: some groups miss a stage, compress a stage or work with several stages simultaneously (p. 106).

Leat and Nichols suggest a relationship between students' thinking activities as they move from one stage to another and the levels of the SOLO taxonomy. The SOLO taxonomy (Biggs & Collis, 1982) has five levels of increasing structural complexity of students' responses to an assignment: (1) the *pre-structural level*, where no single relevant piece of information is connected with the question; (2) the *uni-structural level*, where one piece of information is used to give a correct answer to the question; (3) the *multi-structural level*, where two or more pieces of information are correctly related to answer the question but are unrelated with each other; (4) *the relational level*, where two or more pieces of information are interrelated with each other and also correctly related to answer the question; and (5) *the extended abstract level*, where two or more interrelated pieces of information are correctly connected with the question and where abstract principles and hypothesising beyond the given data are used. Leat and Nichols (2000b) suggest that the move from the display stage to the setting stage corresponds with a shift from the uni-structural level to the multi-structural response level. The move from the setting stage to the sequencing and webbing stage means a shift from the multi-structural level to the relational response level.

Leat and Nichols' study suggests that the mystery is a promising strategy for fostering relational thinking in geography lessons. Research into the effects of the use of mysteries on relational thinking of secondary school students (Karkdijk et al., 2013) showed a significant positive effect. A multilevel regression analysis revealed that students in Dutch upper secondary education who used mysteries in geography lessons established significantly more correct geographical relationships than students who followed the regular geography curriculum.

5.3 Research aim and research question

A previous study on the use of mysteries and geographical relational thinking (Karkdijk, Van der Schee, & Admiraal, 2019) found large differences between groups in relational thinking: the majority of the groups used unconnected relationships to explain the mystery question (responses on the multi-structural SOLO level) and only a minority of the groups gave a coherent, interconnected answer (on the relational level of the SOLO taxonomy). Except in the work of Leat and Nichols, we have found no information about strategies groups employ to understand a mystery and the effectiveness of these strategies for relational thinking. The aim of this study was to describe which strategies small student groups employed to understand a geographical mystery. More evidence on these strategies could be helpful to understand differences between groups in their relational thinking when working on assignments which demand multi-causal, complex relating skills (such as mysteries). This evidence could also help teachers to foster multi-causal relational thinking in group work on these assignments and in a whole-class debriefing afterwards. For the current study, we adopted a qualitative approach to answer the following research question: *What are the differences between low-performing and high-performing groups in their strategies to understand a geographical mystery?*

5.4 Method

5.4.1 Participants

The participants were students of twelve classrooms of six secondary schools in the Netherlands. Three schools were located in the highly urbanised western region of the country (in Hilversum, Gouda and Rotterdam) and three in less urbanised regions (one in Hoorn, in the north-western region, and the two other schools were located in Goes and Middelharnis, in the south-western region). Twelve professional geography teachers were involved in the research project. A total of 35 groups of three students, 15 to 17-year-olds, in upper secondary education were observed working collaboratively on the mystery. Twelve of these 35 groups were analysed on their strategies for understanding a mystery.

5.4.2 The mystery

The mystery was designed for students in upper secondary education by the first author, based on information about severe landslides in Rio in spring 2010. The mystery presents Fabio Pereira, a favela dweller, who was offered a new dwelling on the outskirts of Rio by the government because his neighbourhood was hit by severe landslides in 2010. The mayor had ordered the most threatened favelas to be removed for the safety of the residents. However, Fabio refused to move. The mystery question was why Fabio Pereira refused to move from his threatened house to a safer apartment offered to him by the government. The mystery strips contained information on: Fabio's house that was built by his father thirty years ago; the strong bonds in the migrant community of his neighbourhood, the central location of his present neighbourhood and the peripheral location of the new apartment offered; the high real estate prices in the central city; projects to build luxury apartments in parts of the city where the government had also removed favelas for safety reasons; the start of the rehousing very shortly after the landslides and without consulting the inhabitants; the approaching football World Cup and Olympics in Rio; and Rio as a tourist city. With the help of this information, students had to analyse the geography and society of Rio in order to answer the mystery question. The most important inference students had to make from the data strips was Fabio's distrust of government intentions with regard to the rehousing plan, because his neighbourhood could easily be transformed into a more profitable part of the city. Thus, the assignment focused mainly on vertical relationships within the region.

All 35 groups had to construct a concept map of their explanation of the mystery, because a concept map provides not only a representation of the relational thinking of students (Mehren et al., 2018; Shavelson, Ruiz-Primo, &Wiley, 2005), but collaborative concept mapping also has the potential to elicit reasoning on relationships (Van Boxtel et al., 2002). The concepts had to be connected by labelled arrows to establish propositions (two concepts connected by a labelled arrow). The students had to formulate the concepts by themselves. Each arrow in the diagram had to represent a causal relationship between two concepts.

5.4.3 Data collection and analysis

The 35 groups worked outside class and all of the group work was recorded with a video camera. No time constraints were set on the group work. Most groups finished their work in about 60 minutes. To determine their level of performance, each group's discussion was verbally transcribed and analysed in conjunction with their concept map on established relationships. To investigate differences in strategies to explain the mystery, we observed the films of the group work of 12 selected groups.

5.4.3.1 Group performance

The transcriptions of the group discussion and the concept map of each of the 35 groups were both analysed for established correct *and* relevant propositions. We used a criterion map for this analysis (see Appendix C). Each proposition was scored and a total score was obtained. A correct and relevant proposition scored two points, an incomplete or unclear proposition one point and incorrect or irrelevant propositions got zero points. In this way a total proposition score was obtained for each group. For a detailed description of our coding system and the construction of the criterion map, we refer to Karkdijk, Admiraal, and Van der Schee (2019).

All propositions with a score of one or two points were represented ("standardised") in the format of the criterion map, to allow comparison of the structural complexity of the groups' explanations. Structural complexity was expressed as a level of the SOLO taxonomy of Biggs and Collis (1982), using the "standardised" concept maps and the transcriptions of the group discussions. In order to operationalise the taxonomy, we followed precisely the description of Biggs and Collis' SOLO taxonomy (1982) and Stimpsons' approach (1992) which corresponds with Biggs and Collis' description. Following Biggs and Collis, we also used transitional levels (see Table 5.1). For a detailed description of our operationalisation of the SOLO levels, see Karkdijk, Van der Schee and Admiraal (2019).

5.4.3.2 Selection of 12 student groups

We decided to compare the strategies of the groups with the lowest proposition scores with those of the groups with the highest proposition scores. Lowest and highest proposition scores were defined as one standard deviation or more from the mean (20.8). Because of the skewed distribution of the scores, we had to add one group in the highest category (standard deviation of 0.9342) in order to compare the same number of groups. Table 5.2 shows that groups with the lowest proposition scores (the first six groups) also had SOLO levels up to the multi-structural level, indicating that their explanations of the mystery consisted of isolated relationships. The six groups with the highest proposition scores had outcomes on the relational level or higher, indicating a coherent explanation consisting of interrelationships.

Table 5.1 SOLO levels used in this study.

Level nr	SOLO level
1	Uni-structural U
2	Transitional U/M
3	Multi-structural M
4	Transitional M/R
5	Relational R
6	Transitional R/EA
7	Extended abstract EA

5.4.3.3 Strategies employed by groups

In this study, we defined strategies as the acts of the groups in their group work as they tried to understand the mystery. We used Leat and Nicols' stages (2000b) for the analysis of these acts to explain the mystery (see Tables 5.2-5.5), because they were clearly recognisable, although each particular group had its own sequence in the progress of the group work. We did not operationalise the abstract stage, because all groups started the construction of the concept map immediately after finishing the physical manipulation of the strips.

Concerning the *display stage*, the strategies for reading the strips and familiarising themselves with the data were analysed. How they organised the reading of the strips was important, because it may have meant that not all group members familiarised themselves with all strips, with the result that individual members started with incomplete information. The different strategies for organising the display stage will be described and discussed in the Findings.

Concerning the strategies employed in the *setting stage*, we first observed whether a group started with a priori categories to form sets or whether the categories came from the information provided. In the first case there could be a risk of blinding themselves with an a priori categorisation through which the data were interpreted. Second, we focused on the joint categorisation of the strips into sets as an indication of the level of collaboration. Had the sets been determined by all members or did individual members form sets for themselves without discussing it with the others? If the latter concerned a single set while all other sets had been determined collaboratively, we categorised this as “almost”. Third, we analysed whether all group members were aware of the distribution of the strips among the sets. If not, some group members would have incomplete information about the content of the sets (for example in the reject pile). If this was the case with one to five strips (less than twenty per cent of all strips), we categorised this as “almost”. Fourth, we looked at the size of the reject pile. The number of strips in the reject pile varied in the course of the group work. We decided to count the number of strips in the reject pile when the groups had finished the setting stage, because this meant that they were not available in the most important part of the relational thinking process.

Leat and Nicols describe the *sequencing and webbing stage* as a separate stage, where groups employ a relating strategy in order to find a solution to the mystery problem. We defined the relating strategy *as the thinking strategy of the group, when they decided how to connect strips in order to find an answer to the mystery question.* Most of the time it was possible to discern a stage in the group work after the categorisation of the strips where the group members were clearly focused on establishing connections between sets or between individual strips of sets. However, already in the setting stage, or sometimes right from the start, comments were often made by group members about how to connect the strips. We therefore decided to take into account the whole group work to characterise the relating strategies. Because the categories to characterise the relating strategies emerged from the analysis, they will be described and discussed in the Findings.

We operationalised the *reworking stage* as the stage before the construction of the concept map, where the group decided to reorganise the established relationships between strips, and to use strips from the reject pile in order to gain a better understanding of the mystery. This stage was easily recognised, because students announced the stage verbally (e.g. “we have to do it differently” or “we have to look at the reject pile again for useful strips”) and acted by reconsidering the reject pile and modifying some or all sets. Sometimes the reworking remained modest, when strips were moved from one set to another and strips from the reject pile were included. In other cases, the reworking was more radical, when two or more new sets were established or two or more sets were removed.

The second author acting as a second rater observed four groups on their relating strategy, the most subjective and complex part of the analysis. The results of the analysis were consistent with those of the first rater.

5.5 Findings

First, we describe the differences in strategies employed in the different stages by the twelve selected research groups, as represented in Tables 5.2-5.5. Second, we present case studies of two groups to illustrate their reasoning processes and their relating strategies as they tried to understand the mystery. A third case study, of group 24, can be found in Appendix D.

5.5.1 Differences in strategies

5.5.1.1 The display stage

As can be seen in Table 5.2, we observed two main strategies for reading the strips and familiarising themselves with the data: (1) the strips were divided between group members who read them individually; (2) the strips were spread out on the table to be read by all. Groups that used the first strategy risked group members continuing the group work with incomplete information, especially if the reading was done in silence, as was the case in groups 40 and 34. This could be a hindrance for understanding. However, one group with good results (34) also used this strategy. Moreover, although not all strips had been read, Table 5.2 shows that most groups embarked on a discussion of the information provided right

Table 5.2 Differences in strategies in the display stage.

Group No.	Total propositio n score	SOLO level	Strategies in the display stage		
			Strips divided between group members to be read individually	Strips spread out on the table to be read by all	Group discussion while reading strips
Low-performing groups					
22	8	3	x		x
24	9	3		x	x
40	9	2	x		
39	10	3		x	x
12	11	2		x	x
66	11	3		x	x
High-performing groups					
34	30	5	x		
65	32	5		x	x
33	35	5		x	x
41	36	5		x	x
18	37	5		x	x
16	49	6		x	x

from the start and that low- and high-performing groups had nearly similar strategies in the display stage.

5.5.1.2 The setting stage

Table 5.3 Differences in strategies in the setting stage.

Group No.	Strategies in the setting stage						
	Were the categories made and imposed on data a priori?		Were all (sub)sets designed by all?		Were all strips assigned to a (sub)set by all?		Reject pile
	Yes, but (later) also from data	No, all from data	No	Almost/ yes	No	Almost/ yes	Number of strips
Low-performing groups							
22	x			x	x		9
24	x			x	x		11
40		x		x	x		7
39		x		x	x		4
12	x			x	x		2
66	x			x	x		9
High-performing groups							
34	x			x	x		3
65		x		x	x		0
33		x		x	x		0
41		x		x		x	5
18	x			x		x	7
16	x			x		x	0

Table 5.3 reveals no major differences between low- and high-performing groups in their strategies for organising the setting stage. Most groups started with a priori categories to form sets, for example the categories “useless” or “not useless”; or used the economic, political and physical dimensions as a priori categories, but all groups also used the data for categorisation. In all groups the categories and resulting sets were determined jointly, but in most groups not all group members knew all the strips a set contained. However, the three highest performing groups showed a higher level of collaboration and communication: each member was aware of all strips in each set. On average, the reject piles of the low-performing groups were somewhat higher than those of the high-performing groups. The three high-performing groups were able to include all strips in their explanation.

5.5.1.3 The sequencing and webbing stage: four relating strategies

The major difference between low- and high-performing groups was in their *relating strategies* (Table 5.4), observed not only in the sequencing and webbing stage, but also in the

Table 5.4: Differences in relating strategies.

Group No.	Relating strategies in the sequencing and webbing stage and in the other stages			
	Looking for a logical/chronological chain of events	Reasoning from categories	Relating separate strips one by one to Fabio's decision	Webbing
Low-performing groups				
22		x	x	x (very limited)
24		x	x	x (limited)
40	x		x	x (very limited)
39	x	x	x	x (very limited)
12	x		x	x (limited)
66	x		x	x (limited)
High-performing groups				
34				x (main)
65	x			x (main)
33				x (main)
41			x	x (main)
18			x	x (main)
16				x (main)

other stages (see Method). We observed four relating strategies that were employed to understand Fabio's decision to stay in his threatened house: *(1) looking for a logical and/or chronological chain of events; (2) reasoning from categories; (3) relating separate strips one by one to Fabio's decision; and (4) establishing causal relationships between the strips, the webbing strategy.*

Groups that were *looking for a logical and/or chronological chain of events* tried to reconstruct the story of the landslides in Rio and the events that happened in Fabio's life. They usually made a sequence of the strips and tried to tell the story in order to understand the question of the mystery. Students argued that if the strips are ordered so that you know what happened first, next and last, as well as the logical consequences of these events, you would be able to understand the significance of the individual strips, Fabio's circumstances, and his decision.

The strategy *reasoning from categories* meant that after the group had made categories and formed sets, these sets became the organisational frame for their relational thinking. For example, they could try to summarise the information of all strips within a set in one sentence and then to relate these summarising sentences, or to reason about relationships between the strips within a set or about relationships between sets.

Trying to understand the mystery by *relating separate strips one by one to Fabio's decision* was observed when groups tried to understand the significance of each separate strip to the decision Fabio took. For example, Fabio's house being built by his father thirty years ago was easily connected to Fabio's decision. However, this method was not easily applicable to most strips, for instance to the strips containing information that Rio has very high land values and real estate prices. These kind of strips were labelled as irrelevant by the groups that used this strategy.

Groups that looked for causal relationships between strips before answering the mystery question, the *webbing* strategy, tried to understand the geography and society of Rio with the aid of connections between multiple strips. For example, connecting strips with information on the location of Fabio's present house near the centre of Rio, the location of his new apartment offered by the government on the outskirts of Rio, the long and costly distance between this new apartment and Rio's city centre, and the high land values in Rio gave a group insight into the significance of location in Rio. This insight, connected with luxury apartments being built on formerly released areas near the city centre, explained Fabio's distrust in the government which could have intentions other than helping poor favela residents. In fact, the webbing strategy meant postponing answering the mystery question until a satisfying understanding of the regional and personal situation had been reached.

Table 5.4 shows that all high-performing groups used the webbing strategy as their main strategy while low-performing groups used the other strategies but hardly employed the webbing strategy. In some low-performing groups, there was almost no webbing to observe; we categorised this as "very limited". Other low-performing groups started off well in webbing but for various reasons they did not continue. In these cases, we categorised the webbing strategy as "limited". Sometimes webbing was hindered by the many strips in the reject pile (groups 22, 24 and 66) and/or by the strategy of reasoning from categories (group 24). In the latter case, relating was limited to strips within a set or to relationships between sets. Consequently, free linking between individual strips was partly blocked. In other low-performing groups webbing was hindered as a result of poor communication between students (groups 12 and 66) or very low concentration (group12).

Table 5.4 also shows that low-performing groups used more strategies than high-performing groups. Three high-performing groups only used webbing. Groups (65, 41 and 18) started with another strategy, but turned to the webbing of individual strips to understand the society and (political) geography of Rio.

5.5.1.4 The reworking stage

Table 5.5 shows that four low-performing groups used a *reworking stage*, in order to increase their understanding. Only one high-performing group reworked their arrangement of strips.

Table 5.5 Differences in the reworking stage.

Group No.	Strategies in the reworking stage		
	Did the group rework the data before constructing the concept map?		
	No	Yes, modest	Yes, radical
Low-performing groups			
22	x		
24		x	
40	x		
39			x
12			x
66		x	
High-performing groups			
34	x		
65	x		
33	x		
41	x		
18		x	
16	x		

5.5.2 Relating strategies presented: two case studies

In this section case studies present the relating strategies and the reasoning process of a low-performing group (40) and a high-performing group (65). Appendix D contains a third case study (group 24). The strategies and reasoning process are illustrated by excerpts from the group discussion. Each excerpt is preceded by a description of the relevant part of the group work and some comments on the utterances within the excerpt.

5.5.2.1 Group 40

This was a group of two boys and one girl, 15/16-year-olds. After the strips were spread out on the table, they decided that each group member would read one third of the strips in silence. After the reading, student 1 proposed that they make a logical order of the strips (lines 56 and 65). When student 2 proposed that they find out which strips provided good reasons for Fabio to stay first (lines 60 and 68), she was corrected by student 1 (lines 61and 69). Thus, at the start of the discussion, the strategy “relating individual strips one by one to Fabio’s decision” was rejected in favour of the strategy “looking for a logical and/or chronological chain of events”.

56	S1	Well, order it in a way that seems logical to you.
57	S3	I think it starts with this one.
58	S1	I think it starts with strip number one.
59	S3	I don’t think so. The question is about him, so it would make sense if it starts with this and then carries on saying.
60	S2	I would also consider this one.. This is also a good reason why he just needs to leave, I think.
61	S1	Yes, needs to leave, but just start with the beginning. Because why is it said...? Then this would be a good start.
(…)		
65	S1	Yes, shall we have a look at which ones we can put in order then?
66	S2	Yeah
67	S3	I think we should … that one….
68	S2	Yes, or just take the cards which have good reasons, according to us.
69	S1	What do you mean, reasons? Shouldn’t you first find the correct order so that you know what is going on and then focus on the reason. Because you can pick the reasons why the school is closed but you do not know that he has four children.
70	S2	Right, that makes sense.

They then started with a chronological order of strips that contained a date or a time indication. The other strips were checked one by one to find out whether they contained a reason for Fabio not to move. Collaborative reasoning on the meaning of the strips or their relationship with other strips (the webbing strategy) was very limited. A promising start connecting tourism in Rio and the distance of Fabio’s new apartment on the outskirts of Rio to the city centre, was immediately ended by a lack of insight into the location of Fabio’s house (lines 120-124), caused by a superficial or partial reading of the strips. Separate strips judged as useless were put in the reject pile without any discussion (lines 127-131). The strategy “relating individual strips one by one to Fabio’s decision” was therefore used alongside the strategy of looking for a chain of events.

120	S2	Well, perhaps with that important tourist city because that way he might have more income.
121	S1	Yes, but he lives in Morro dos Prazeres, right?
122	S2	No, he lives in Rio.
123	S3	No, he lives in Rio, but Rio is big. They use that name for parts of Rio, I thought.
124	S2	Ah all those names. I'm getting so confused. So, where does he live? In Morro dos…
125	S1	Morro dos Prazeres, that's where he lives. He's got his house there. That's probably just a part of Rio.
126	S2	Okay.
127	S3	Fucking film, man. I really don't need that.
128	S1	Calm down, man.
129	S3	This is about the film. This one should be included as to why he wants to leave.
130	S2	City of God… This has got nothing to do with the topic. (incomprehensible)
131	S3	This should also be included, I think. Why he needs to leave. This has got nothing to do with it. Nor this one. This is about travelling.

Their lack of insight into the location of Fabio's current house and the location of the new apartment also led them to draw wrong conclusions. For example: student 3 reasoned that Fabio had to move to the northeast of Brazil (line 282). However, he was corrected by student 1, who had read a strip with information on Fabio's new apartment on the outskirts of Rio (line 291). Student 3 checked the reject pile and re-used the strip with information on the long distance between the outskirts of Rio and its city centre (line 295). They added this as an argument for Fabio to the concept map already under construction (line 298).

282	S3	Look, northeast Brazil is an area gripped by drought, poor soil and unemployment. So they need to leave somewhere, go somewhere, but the place he needs to go to still has poor soil and unemployment.
283	S1	Where do you see that?
284	S3	Six, it is only less dangerous.
285	S1	But northeast Brazil is the place he comes from.
286	S3	No way is he going to live 100 kilometres away.
287	S1	Yes, but he is from northeast Brazil. That's probably why he has left there.
288	S3	Yeah, and he will also move to another place in northeast Brazil, I think. Do you know how much money it costs to move everyone to another place?
289	S1	Is that what you think or what you know?
290	S3	Logical thinking.
291	S1	I think it would make more sense to eh… People will be accommodated in several flat areas close to the outskirts of town. Which happens to be written here.
		(…)

295 S3 This is also a good reason: from the outskirts, from Rio's city centre, it's an hour's drive.

296 S1 Yes, good one.

297 S3 Lots of travelling to the city centre.

298 S2 Alright, I'll start a new one then, okay? Yes, long distance to travel to the city centre.

This group work was characterised by two strategies to find an answer to the mystery question: making a chain of events and trying to relate separate strips one by one to Fabio's decision to stay. The discussion shows an almost total absence of collaborative reasoning. Relating strips to one another in order to understand the geography and society of Rio (webbing), was almost completely absent. As a result, their total proposition score was relatively low (8) and their concept map was on the multi-structural level with only minimal explanation (Figure 5.1).

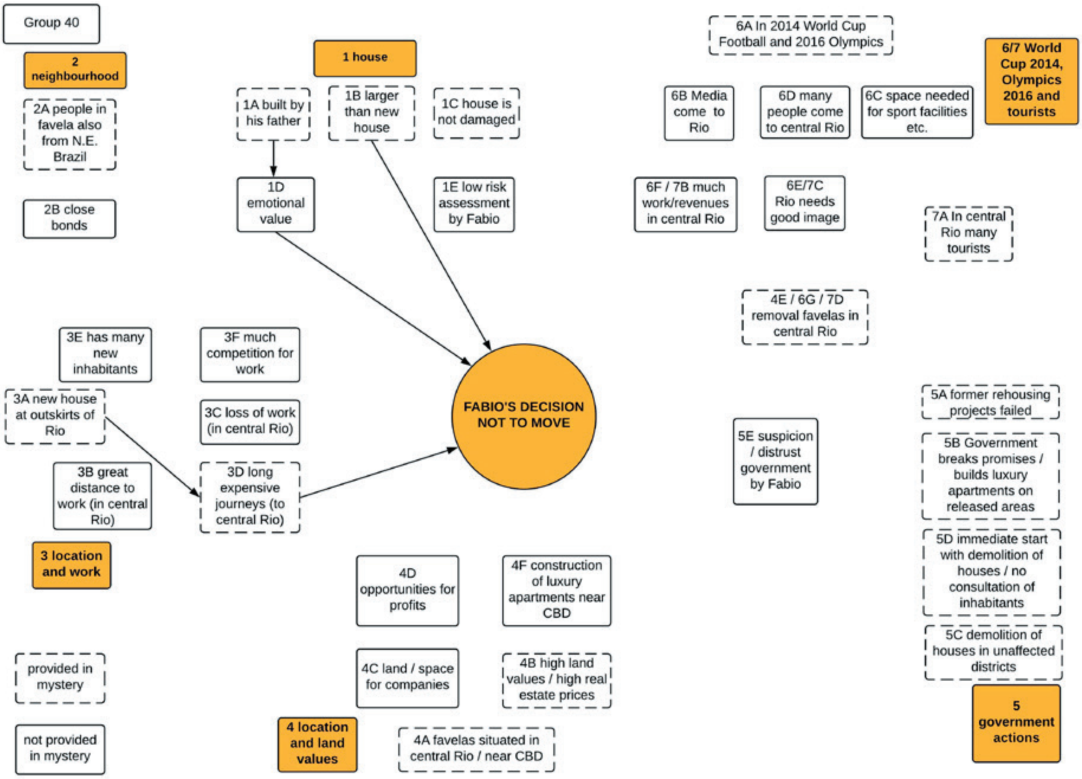


Figure 5.1 Standardised concept map of group 40.

5.5.2.2 Group 65

Group 65 consisted of three boys aged 16/17 years old. The group started by spreading out the strips on the table. Although the students read the strips individually, they discussed nearly all of them together. Right from the start they focused on relationships between strips, the webbing strategy, (lines 6, 8), as well as possible connections between several individual separate strips and Fabio's decision (line 3).

1	S2	Let's just start with card 1.
2	S3	Yes, that doesn't really matter.
3	S2	This just says that it is hit by heavy rainfall; that doesn't seem like a valid reason to me.
4	S1	Because of the heaviest rainfall.
5	S3	Yes, that's just information, that is just information.
6	S2	You could say, for instance, that this is the cause, and as a result, all those can.
7	S3	The government commands…
8	S1	Yeah, you have to make the right connections between those.

Within two minutes extensive collaborative reasoning characterised the discussion (lines 33-44). The connection of two strips lead to the inference that the government is not trustworthy: favela dwellers will simply be driven out for the building of luxury apartments (lines 41-44). The webbing strategy had become the dominant one.

33	S1	The mayor wants to plant forests on the released demolition site.
34	S1	Because of the forest, landslides will be less likely to occur, leaving the soil a little firmer.
35	S3	Yes, that's true.
36	S2	But that will only happen after everything has been planted.
37	S1	And that takes quite long, because those forests still need to grow.
38	S2	Yeah, that's written here as well, in 1982, such a village was cleared, because of a serious risk.
39	S3	And did that turn out well?
40	S2	It doesn't really mention whether it was successful.
41	S1	Look check this one: Some years ago, at several "extremely dangerous" sites, favelas were removed by the government and replaced with high-rise towers with luxurious apartments.
42	S2	Instead of planting forests. They just removed them! Yes.
43	S3	So, you really cannot always believe them.
44	S1	No.

After that (third minute), they continued the webbing by adding new strips to this connection:

54	S1	The demolition has already started here, while any eh… landslides had not even happened yet.
55	S2	So, in fact, it is just another, um… they just got another destination.
56	S3	Yes.
57	S2	That has to do with those luxurious apartments then.
58	S1	Botafogo is a small favela, south of Rio, in which no house has been destroyed (while the demolition has already started there).
59	S3	It could also be an excuse, he should know.

60	S2	That has to do with those apartments.
61	S1	Yes.

A few minutes later, this web was again further expanded:

110	S3	Land and real estate prices in Rio are among the highest in the world.
111	S2	Yes, that does match with those luxurious apartments.

These excerpts are a good illustration of what is meant by webbing. Although at the start the students also tried to connect strips directly and in a linear manner to Fabio's decision, their focus in the display stage was already primarily on establishing relationships between strips. After all strips had been read and discussed, the webbing continued. Sentences that contain linking expressions are noteworthy (lines 231, 233, 235, 241).

231	S2	Yes, ah, you could add this to that one then. The mayor says he wants to plant forests on the released demolition site. Yes.
232	S1	Says. But then things will be built.
233	S2	Yes, you could, for instance, connect that to this one, those high housing prices, he had wanted to build luxurious apartments. Yes, that's what he says.
234	S1	So that can be placed here.
	S3	But I thought that, on one occasion, it didn't happen right? Uhm, he'd lied or something, that there wouldn't be a forest after all.
235	S1	This one should also be added here somewhere, that they'd already started demolishing, without having seen any landslides. That has got to do with it as well. That the government just wanted to build and all.
236	S3	Yes.
237	S3	Yeah, residential towers have been built as well.
238	S3	That is in the government's interest.
239	S2	False promises had been made and, in fact, they just earned a lot of money.
240	S3	In 2014, the Football World Cup final was played in Rio and, in 2016, Rio will host the Olympic Games.
241	S2	Yes, here: Rio is the most important tourist city in Brazil, so we can connect that to this one yeah.
242	S3	They want to make their city look good.

After 18 minutes' work, they had formulated their main conclusion, based on their understanding of what is happening in Rio (line 269):

265	S2	But we haven't really drawn a conclusion as to why he wanted to stay there so much.
266	S3	Yeah, maybe because he just doesn't trust the government.
267	S2	Yes, that more or less sums it up, I think… reading all this text over here.
268	S1	Yeah, that seems to be our main… He
269	S3	Because he obviously knows his city, and he knows what has happened.

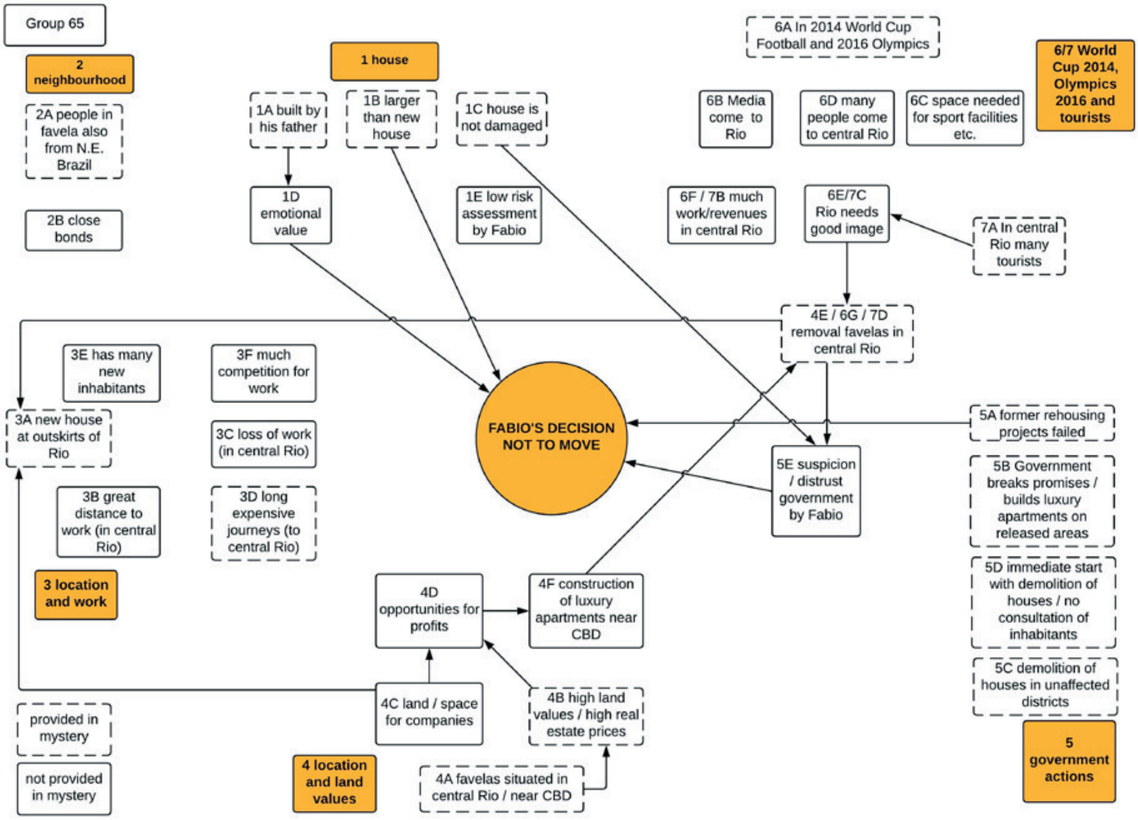


Figure 5.2 Standardised concept map of group 65.

Group 65 was characterised by highly concentrated, collaborative reasoning by all three group members. Clearly focused on causal relationships, their understanding of the society and political geography of Rio grew as the group's work continued (line 269). Their total proposition score was 32 and their concept map was on the relational level (Figure 5.2).

5.6 Discussion

In this qualitative research we analysed which strategies low-performing and high-performing groups employed to tackle a geographical mystery. Our major finding was that low-performing and high-performing groups differed in the relating strategies they employed. Whereas low-performing groups used two or more of the relating strategies: looking for a logical and/or chronological chain of events, reasoning from categories and relating separate strips one by one to Fabio's decision and used the webbing strategy hardly at all, high-performing groups used the webbing strategy as their main strategy. This suggests webbing to be the most promising relating strategy for understanding the multi-causal character of the mystery. In this mystery landslides and their devastating impact on favelas were triggers, while the interrelations between political, economic and personal/emotional factors acted as background factors that had to be analysed in conjunction with each other. Webbing involves postponing the answer to the mystery question and first searching for causal relationships between the strips in order to understand the interconnections between these background factors. As a student from the best-performing group (16) said after they analysed the mystery: “In my head the relationships look like a spider’s web”. The three other relational strategies had serious flaws when it came to analysing the mystery thoroughly. Relating separate strips directly to Fabio’s decision could not lead to a comprehensive explanation, because only a few simple causalities could be found in this way. Making a chronological sequence of events, thus trying to tell the story of what happened to Fabio, also produced disappointing results, because Fabio’s story and the landslides were only a story-line in the mystery, while most background factors could not be organised as a chain of events. Finally, categorisation and forming sets was helpful to get an overview of the information provided, but when students continued their thinking with the resulting sets, the risk of blinding themselves was obvious. Thinking in relationships between strips within sets or in relationships between sets hindered webbing between strips of different sets, necessary for the discovery of important causalities. In contrast to the observations of Leat and Nichols (2000), low-performing groups reworked their findings more often than high-performing groups did. They had to do this because of the disappointing results of their discussion so far.

Our findings on differences between low- and high-performing groups in their relating strategies and standardised concept maps (Figures 5.1, 5.2 and the group discussion and standardised concept map of group 24 in Appendix D) suggest a resemblance with the three-stage competence model in systems thinking of Mehren et al. (2018). At the lowest competence level, the mystery was explained with mono-causal relationships or short linear relationships by trying to connect separate strips directly to Fabio’s decision. At a somewhat higher competence level, linear thinking became more predominant as students tried to make a logical and/or chronological thinking line, a chain of events, or gave separate linear causal explanations using sets. The highest competence level was reached by groups that identified networked elements and relationships by using the webbing strategy.

The suggestion of Leat and Nichols (2000b) that the move from the setting to the sequencing and webbing stage resembles a progression from the multi-structural SOLO level to the relational SOLO level is debatable. Most of our observed groups reached the sequencing and webbing stage, but not all had an answer on the relational SOLO level. For in the sequencing and webbing stage, a group can make a chain of events (sequencing) or can make connections only within or between sets, relating strategies that will result in a response on the uni-structural or the multi-structural SOLO level. Only a webbing strategy used in this stage will result in a coherent answer on the relational level.

5.6.1 Implications

Our analysis of relating strategies revealed the difficulties that low-performing groups had with complex relational thinking in geography: they did not know how to tackle a multi-causal problem. This indicates the need for explicit attention to be paid to relating strategies when setting assignments that, like mysteries, ask for multi-causal thinking. A teacher could give a hint to a group which is stuck to reconsider their relating strategy. Another possibility would be to ask the groups to explain their chosen relating strategy when they present their solutions in a whole-class debriefing. In discussions with students that employed other strategies, a comparison of the relative effectiveness of the strategies could be made and hopefully a more promising relating strategy would then be revealed to students in low-performing groups.

We also recommend the frequent use in geography lessons of assignments that focus on multi-causal, complex reasoning, like mysteries, thinking with scenarios, simulation games on climate change or water management, etc. This could deliver a "systems disposition" on the part of students. DeVane et al. (2010) state that "given the right scaffolding and structure in a learning community, participants (…) can develop a "systems disposition" towards different problem contexts. Such a disposition is not a universal heuristic for inquiry (…), but it is a set of attitudes toward systems" (p. 15). Although assignments may be quite different in terms of theme or region, students would be inclined to look for coherence, interconnections, networked relationships. The additional use of concept maps or causal diagrams as representational tools can be very fruitful in fostering complex, multi-causal thinking (Cox, Steegen, & Elen, 2018).

5.6.2 Limitations

This paper describes our observations of the strategies of only 12 groups as they tackled a geographical mystery. We found four different relating strategies and suggest a correlation with the level of group performance. More evidence is needed to support these suggestions.

Our finding that the three best groups demonstrated good communication and collaboration and our observation of poor communication in several low-performing groups indicate that communication could also be a relevant difference between high- and low-performing groups. However, this was beyond the scope of this study.

5.7 Conclusion

Teaching geography with a clear focus on the analysis of complex regional problems and change offers students the necessary understanding of the relational character of people's lives in specific places of the world. A regular use of complex, multi-causal assignments in geography lessons and explicit attention to students' relating strategies could be useful in helping students to develop an attitude, an inclination to look for the relationships between different dimensions and scales of a regional problem. Such a "systems disposition" could help them to avoid jumping to quick, mono-causal conclusions on complicated regional problems and to look first for a deeper, relational understanding. This makes secondary geography worth teaching.

Chapter 6

Conclusions and discussion

6.1. Introduction

The aim of this dissertation was to obtain evidence on students' geographical relational thinking, in order to gain insight into possible ways that teachers might help their students master this skill. Using the Geographical Analysis Model of Van der Schee (2000), geographical relational thinking was defined as analysing, explaining, evaluating and/or predicting the horizontal and vertical relationships and the interactions between them, on different scales, that cause regional change. The research focus was on the ability to think in interrelated, multi-causal relationships in order to give coherent explanations for a phenomenon or an event framed as a mystery question, and the strategy of the mystery was used to foster and diagnose students' relational thinking. After an effect study into the use of mysteries on students' individual geographical relational thinking, a more extensive study was conducted on students' geographical relational thinking in the group work part of a mystery. The research took place in regular classrooms in Dutch upper secondary education.

The main research question was:

How does the use of mysteries in secondary geography education support geographical relational thinking in students?

The use of "support" in the question implied that we used mysteries to foster geographical relational thinking skills as well as to diagnose their geographical relational thinking. The main research question was divided into these seven research questions:

1. What is the effect of the use of mysteries on students' geographical relational thinking?

2. Which geographical relationships do students in small groups establish to solve a mystery?

3. How coherent are the solutions to the mystery posed?

4. *How does geographical relational thinking in terms of the SOLO taxonomy differ between groups?*

5. *How does geographical relational thinking in groups differ between the two mysteries?*

6. How can differences between groups in geographical relational thinking be explained by characteristics and collaborative behaviour of the groups?

7. *What are the differences between low-performing and high-performing groups in their strategies to understand a geographical mystery?*

In order to answer these seven research questions, four studies were conducted. The first study had a quasi-experimental quantitative design, using a pre-test and post-test to find an answer to the first research question. Data were gathered at seven secondary schools. The second study was a quantitative descriptive study, in which the second and third research questions were answered. Data were gathered at six secondary schools. The third study, a quantitative correlational study, used the same database in order to answer the fourth, fifth and sixth research questions. The last question was answered in the fourth study using a part of the same database used for the second and third studies and had a qualitative, descriptive design.

In the next section of this final chapter, the most important findings of the four studies on the seven research questions are summarised and the answer to the main research question is formulated as the main conclusions. Section 6.3 reflects on the methods used, resulting in the limitations of this research and recommendations for future research. Section 6.4 contains the theoretical and methodological contributions of this dissertation and section 6.5, the final section, presents the most significant implications.

6.2 Main findings and conclusion

6.2.1 The first study

The first research question to be answered before launching a subsequent larger research project on mysteries and geographical relational thinking, was: What is the effect of the use of mysteries on students' geographical relational thinking? The teachers in experimental groups used three or four mysteries in their lessons. The pre-test and post-test focused on variables linked to the Millennium Goals. The tests were devised to measure causal relational thinking in geography, operationalised as the ability to construct a relational scheme consisting of three or more interrelated concepts. We found that the means of correctly established relationships in the pre-tests and post-tests differed significantly between schools, so a multilevel regression analysis was necessary to explain the variance in differences between pre-test and post-test scores at student level. Although the differences between schools were striking, they were beyond the scope of our research. We only analysed the differences between students in their progression between pre-test and post-test. The progression in the number of established relationships was significant higher for students in the experimental condition than for students in the control condition. We found that the repeated use of mysteries in combination with the representation of the constructed explanations as relational schemes had a moderate positive effect on fostering students' relational thinking skills in geography. This underpinned the use of mysteries for further research into students' geographical relational thinking.

6.2.2 The second study

Study 2 analysed the relationships students established to explain the Jakarta or the Rio mystery. The first research question of this study was: Which geographical relationships do students in small groups establish to solve a mystery? Students used a wide range of relationships in order to explain the mysteries. However, the more concrete, less complex relationships dominated. These were apparently easy to establish and were almost or entirely nearly direct and linear. Cross-links were by far in the minority, indicating that constructing a coherent explanation is a higher relational thinking level for many students.

The second research question that directed this study was: How coherent are the solutions to the mystery posed? As a large proportion of the groups (about 40%) did not establish any correct cross-links in their explanation, their explanations were not coherent, but consisted of separate more or less linear relationships. However, a comparable proportion of the groups was able to formulate more coherent explanations by using one or more cross-links and branches in their concept maps. Therefore, groups differed greatly in their ability to think in interrelated multi-causal relationships.

Summarising we conclude that the study revealed that many groups were not able to construct a coherent, interrelated explanation of the mystery.

6.2.3 The third study

The third study focused on the differences between small groups in their explanations of the mysteries and how these differences are to be understood. The first research question concerned differences between groups expressed as differences in levels of the SOLO taxonomy: How does geographical relational thinking in terms of the SOLO taxonomy differ between groups? This part of the study overlapped with the second study, and so did the conclusion that groups differed greatly in their ability to think and reason in interrelated causal relationships. This was illustrated by three groups that operated on the multi-structural level, the relational level and the transitional level between these levels. The group on the multi-structural level reasoned in separate, mostly linear relationships, whereas the group on the relational level demonstrated an ability to construct a more integrated, web-like explanation of the mystery.

The second research question of this study was to explore possible differences between the two mysteries in the quality of geographical relational thinking that they provoked: How does geographical relational thinking in groups differ between the two mysteries? We found that the design of a mystery seemed to influence the complexity of the geographical relational thinking provoked. In the Rio mystery, which was designed with fewer concepts that were easier to connect linearly with the question of the mystery than the Jakarta mystery was, students were asked to think more in interrelated relationships. As a result, the mean total proposition score on the Rio mystery was significantly lower than on the Jakarta mystery.

The third research question of this study was: How can differences between groups in geographical relational thinking be explained by characteristics and collaborative behaviour of the groups? The striking result of our research was that we did not find any significant correlations between the group characteristics educational level, educational year, or previous achievements in geography and geographical relational thinking. The only significant predictor that explained a small proportion of the variance of three indicators of relational thinking was the amount of on-task discussion the students had on the problem of the mystery before they started to construct the concept map. A more extensive discussion on the different aspects of the problem presented in the mystery question and the relationships between these aspects apparently resulted in a better understanding of the relationships within the problem than quickly starting with the construction of the concept map.

6.2.4 The fourth study

The fourth study was designed to obtain more insights into the differences between groups in their geographical relational thinking by analysing the strategies they employed to tackle a mystery. The research question was: What are the differences between low-performing and high-performing groups in their strategies to understand a geographical mystery? We found the most remarkable differences between low-performing and high-performing groups in the relating strategies they employed. Four different relating strategies were found: (1) looking for a logical and/or chronological chain of events; (2) reasoning from categories/sets; (3) relating separate strips one by one to the mystery question; and (4) establishing causal relationships between the strips, the webbing strategy. All high-performing groups used the latter as their main strategy. The low-performing groups used one or more of the other relating strategies as their main relating strategy, which apparently resulted in a more partial, incoherent or less coherent explanation of the mystery. The webbing strategy avoided them reaching a quick answer on the mystery, by first searching for causal relationships between individual strips in order to understand the interconnections between different pieces of information (triggers, personal conditions and background factors). After establishing different interconnections, a more coherent explanation of the mystery was possible and more likely.

The results of this study also made clear that the suggested coincidence of the move from the setting to the sequencing and webbing stage with a progression from the multi-structural SOLO level to the relational SOLO level, as suggested by Leat and Nichols (2000b, 2003), is debatable. Establishing unconnected, separate linear relationships in the sequencing and webbing stage means that an answer is being constructed on the multi-structural SOLO level, instead of a coherent, interrelated answer that characterises the relational SOLO level.

6.2.5 Main conclusions

The mystery, a powerful strategy for fostering and diagnosing relational thinking (Leat, 2001; Leat & Nichols, 2003), was used in this research to investigate students' geographical relational thinking, in order to give teachers possible insights into strategies to foster this thinking skill among their students. The main research question was: How does the use of mysteries in secondary geography education support geographical relational thinking in students?

The main conclusions, based on our findings, are:

(1) The repeated use of mysteries in combination with groups' representation of their explanation of the mystery as a concept map (causal diagram) can have a significant positive effect on students' relational thinking in geography.

(2) The quality of geographical relational thinking of a large proportion of small groups working on a mystery seems generally low, although the differences between student groups appear to be considerable:
   - the more concrete, fairly easy to establish relationships seem to dominate the explanations;
   - where a large proportion of small student groups appears to be able to establish only more or less linear, unconnected relationships in order to explain a complex, interrelated problem, a comparable proportion seems to be able to construct a more coherent explanation by using cross-links.

(3) Small student groups seem to differ in their ability to explain a complex problem such as presented by a mystery in group effort and not in particular group characteristics like educational level, educational year or geography grades. The two differentiating elements of group effort that result in a more extensive and more coherent explanation, might be:
   - an extensive on-task group discussion to understand the connections between the different aspects of the mystery problem before the answer to the mystery question is constructed; and
   - the use of the webbing strategy as the main relating strategy employed to understand the mystery.

With respect to the aim of this research, we conclude that teachers can foster students' geographical relational thinking by:

(a) practising strategies (like mysteries) with more complex assignments to think in interrelated, multi-causal relationships, because a large proportion of the student groups thought in more or less directly, unconnected linear relationships;

(b) using these strategies in a progressive order from assignments with more concrete concepts easy to connect with the mystery problem to assignments with more abstract background information that demands thinking in interrelationships before an answer to the mystery problem can be formulated;

(c) giving explicit attention to cross-links or interrelationships in the whole-class debriefing session; and

(d) giving explicit attention in the debriefing session to the employed relating strategies and the power of the webbing strategy to tackle a complex interrelated problem.

6.3 Reflections on the methods used

6.3.1 Study 1

With respect to study 1, we found no validated geographical relations test, so we had to design a compact test to measure geographical relational thinking. We designed a test with three identical assignments, which asked students to establish and explain relationships between two variables with the use of other related variables that were provided in the instructions on the tests. The tests measured the construct of interrelated, multi-causal thinking and students needed almost a lesson to complete it. A more comprehensive test measuring geographical relational thinking containing different kinds of assignments could have provided a more solid underpinning of our first main conclusion. Furthermore, a repeated, delayed post-test would have made clear whether the positive effect of the intervention would last for more than just one or two weeks.

Because we used mysteries in combination with a special kind of concept map (a causal diagram) as the representation format for the mystery explanation, it was not possible to draw a conclusion on the effect of mysteries themselves on geographical relational thinking. There is evidence that concept mapping activities help students organise their relational thinking and in this way foster it (Cox, Steegen, & Elen, 2018; Hwang, Yang & Wang, 2013; Renshaw & Wood, 2011). However, a comparison of the effect on geographical relational thinking of the use of mysteries with different representational formats was beyond the scope of this first study.

In the light of these limitations future research on the effect of the use of mysteries on geographical relational thinking is recommended, because more evidence is still needed.

6.3.2 Studies 2,3 and 4

With respect to the method employed to obtain the database for studies 2, 3 and 4, the focus was on the ability to establish multi-causal, interrelated relationships in order to give coherent explanations of a phenomenon or an event framed as a mystery question. This choice was made because students struggle with relating variables correctly as found in our analysis of Dutch National Exams in 2009 and 2010 (Karkdijk & Van der Schee, 2012), a finding backed up by the researcher's experience as a geography teacher that students have great difficulties connecting causally more than two variables to explain regional change. However, a second focus on *different kinds* of causal relationships could have deepened the evidence provided by the studies from this research. Then we would have had insights not only into whether students are able to show geographical relational thinking, but also into whether they can do this with different causal relationships such as horizontal and vertical relationships, direct and indirect relationships, personal factors, background factors, triggers, etcetera.. This would have contributed to a more comprehensive overview of possibilities and difficulties students have with geographical relational thinking. Some evidence has been found on the difficulties that students have with establishing horizontal relationships, interactions and indirect relationships (Favier & Van der Schee, 2014a, b; Van der Schee, 1987, 2000), but more evidence on students' ability to interconnect triggers, personal factors and background factors is necessary.

There has been some criticism of the dynamic method used to construct the criterion maps of the two mysteries. Benninghaus et al. (2019) developed a tool for teachers to analyse and quickly assess "influence diagrams" (comparable to concept maps) in order to assess students' systems thinking. This tool is an expert-based reference diagram (criterion map). They acknowledge that the assessment of influence diagrams as the explanation of a mystery is extremely difficult, precisely because of the complexity of a mystery problem. In order to tackle this problem, they used insights from eight experts to construct an expert-based reference map with core interrelationships that were accepted by most experts. They question the enrichment of our original criterion maps (also constructed by experienced geographers) with extra propositions derived from group concept maps, because students are not experts. And they also state that a single criterion map seems to strict for a complex and open assignment such as a mystery is (p. 5). However, a difference from their mystery assignment is that we allowed and encouraged groups to also use extra concepts (not provided in the mystery) to explain the mystery in their own way and in their words. This resulted in additional, unforeseen but possible, correct and relevant propositions that had to be added to the original criterion map. In this way, we created a criterion map that had many possible pathways to explain the complex and open mystery question.

An intriguing, methodological statement made by Benninghaus et al. (2019) is that a more densely connected influence diagram is not necessarily better than a sparse one for students who only have to give a causal explanation of the mystery problem (p. 6). We agree that, generally spoken, it is not necessary that groups establish as many relationships as possible to explain a mystery problem correctly. However, in this particular research, a count of correct and relevant propositions to determine the quality of an answer was meaningful, for to give an interrelated, coherent explanation, students had to formulate propositions using information on six different factors *and* to formulate cross-links between propositions within the factors. This implies that a considerable number of propositions was necessary in order to make the requested explanation as complete as possible.

A final reflection relates to findings to explain the differences between groups in geographical relational thinking. In addition to the qualitative research on the strategies employed to tackle a mystery, a study on group communication could have provided more evidence to explain these differences, for example by using the differentiation in exploratory, cumulative and disputational talk (Mercer, 1996; Mercer, Wegerif, & Dawes, 1999). More research on the relationship between the quality of the talk in small groups and the level of geographical relational thinking is therefore recommended.

6.4. Theoretical and methodological contributions

This dissertation contributes to insights into a key geographical thinking skill that has scarcely been examined at student level. To be able to think in terms of interrelated, multi-causal relationships to understand regional change is not only necessary for understanding geography, but can be seen as a capability that young people need in order to live and act consciously in an interconnected, globalised world (IGU-CGE, 2016; Koninklijk Nederlands Aardrijkskundig Genootschap KNAG, 2017). In the light of this necessary thinking skill, our research has captured the interest of the Dutch Council for Education. Referring to our first study, the Council presents mysteries as an example of teaching 21st century thinking skills in an integrated manner in geography lessons (Onderwijsraad, 2014, p. 22).

This dissertation provides evidence on the current level of geographical relational thinking of secondary school students, as well as on how to increase their ability to think in multi-causal relational thinking in geography. Aside from our definition of geographical relational thinking referred to by several authors (Cox, Elen, & Steegen, 2017; Varró & Van Gorp, 2020), the main contributions made to the discipline of educational geography are fivefold.

First, evidence is provided on the effect of the use of mysteries (in combination with the use of causal diagrams) on geographical relational thinking. This evidence has been almost completely absent up to now and the effect we found underpins the use of this strategy to foster students' relational thinking in geography.

Second, insight is added to the scarce evidence available on the level of geographical relational thinking of students: although some of the groups were able to construct a more or less coherent explanation of the mystery problem, a considerable proportion of the groups reasoned in unconnected linear and direct relationships.

Third, evidence is found on the five stages of the group work on a mystery of Leat and Nichols (2000b, 2003). These are only described by Leat and Nichols, but have never been operationalised before, as far as we know. With the use of our operationalisation based on the characteristics described by Leat and Nichols, the stages were clearly discernible in the group work of low and high performing groups and this made our analysis possible. This underpins Leat and Nichols' five-stage model of group work on mysteries.

Fourth, insight is delivered into the different relating strategies groups employ to understand a mystery. Leat and Nichols (2000b, 2003) describe the "sequencing and webbing stage" where groups establish linear and web-like relationships. The current study revealed four relating strategies groups employed to establish relationships in order to tackle the mystery. Only the webbing strategy led to a coherent explanation on the relational SOLO level. The other three strategies resulted in linear, unconnected relationships and in an incoherent explanation. This insight is different from the suggestion of Leat and Nichols (2000b, 2003) that the sequencing and webbing stage corresponds with the relational level of the SOLO-taxonomy of Biggs and Collis (1982). An operationalisation of the SOLO taxonomy in line with Biggs and Collis (1982) will distinguish this level when an answer consists of branches and interrelationships (cross-links). Groups that are only able to establish unconnected linear relationships in the sequencing and webbing stage operate on the multi-structural level.

Fifth, the four relating strategies add useful information to our knowledge of how groups try to understand a complex multi-causal problem which is what a mystery is. Only the webbing strategy resulted in a coherent explanation. This insight into the unequal effectiveness of the four relating strategies helps teachers understand how to instruct their students to tackle a complex, interrelated problem or how to guide a fruitful debriefing session on a strategy like a mystery and so adds to their Pedagogical Content Knowledge.

In this way, this dissertation contributes to bridging the gap between scientific educational research and teaching practice in secondary schools mentioned by Vanderlinde and Van Braak (2010). They found that teachers were sceptical about the practical relevance of educational research. Research should provide evidence on what to do in their daily teaching practice to foster the education of their students. Our research was conducted by a secondary geography teacher and data were gathered in regular secondary geography classes. The dissertation provides evidence on the effect of using mysteries in class on students' geographical relational thinking, on the quality of this thinking and on the relating strategies employed to tackle a mystery. This evidence offers teachers possible ways to foster these skills and therefore has the desired practical relevance. This teacher research has also extended the knowledge base on teaching and learning in secondary education.

6.5 Implications

Relational thinking is an explicit goal (CEVO, 2007) in the Dutch national geography curriculum for higher general secondary education (havo) and pre-university education (vwo). Students are expected to be able to analyse geographical issues by using horizontal and vertical relationships as described in the Geographical Analysis Model of Van der Schee (2000). However, this research on the use of mysteries revealed that many groups performed poorly on geographical relational thinking and that performance on this thinking skill was not significantly related to educational level, educational year or geography grades. This means that interrelated, multi-causal relational thinking with horizontal and vertical relationships needs more explicit attention at the level of geography lessons in class. But how could this be done?

The evidence found on the positive effect of the use of mysteries in combination with the use of causal diagrams on relational thinking shows that the regular use of this strategy and comparable assignments on interrelated problems will be fruitful. A start with not too difficult multi-causal relational assignments with more concrete information that is easy to connect to the problem question, followed by assignments with more abstract information that needs to be interrelated before an answer on the problem question is possible, offers possibilities for practising multi-causal, interconnected thinking step by step. Through this practice, accompanied by effective teacher scaffolding such as that mentioned in conclusions a-d, students will develop a "systems disposition", an inclination to structure information of a complex problem by establishing causal relationships between elements and by looking for interrelationships between these relationships (DeVane et al., 2010).

Besides using assignments that ask for relational thinking on complex problems, there are more possibilities for implementing geographical relational thinking in geography lessons. An inspiring example for educational geographers and teachers of how to structure geography lessons to foster geographical relational thinking can be found in secondary history education. Van Drie and Stoel (2017) conducted a design study on how to foster students' causal historical reasoning and teachers' Pedagogical Content Knowledge on how to teach it. The causal reasoning of students was first analysed and a rubric was constructed to assess causal reasoning on six categories: (1) using more and different kinds of causes; (2) establishing relationships; (3) using historical concepts; (4) drawing conclusions; (5) using arguments and (6) understanding more explanations. The rubric was also used to design lessons focused on particular elements of causal reasoning. Six design principles resulting from their study deserve attention, especially in relation to this research on geographical relational thinking with mysteries. These design principles are (p. 32-35):

(1) lessons are based on the diagnosis of students' causal reasoning;
(2) elements of causal reasoning are mentioned explicitly as learning goals;
(3) the central question of the lesson is complex and explanatory;
(4) explicit instruction is given on different aspects of causal reasoning;
(5) the teaching strategy to answer the central question is open, activating and stimulates interaction between students;
(6) different learning paths are offered to answer the central question.

This design study is an inspiring example for educational geographers and geography teachers of how to implement geographical relational thinking in their lessons. Educational geographers and geography teachers could also work in co-operation to design lessons that foster geographical relational thinking using the findings of their history colleagues. The rubric could be adapted to geography lessons and the design principles are useful. A few comments on the design principles will illustrate their usefulness in geography lessons with students working on a mystery.

The use of mysteries fits well with these six design principles. Concerning the first design principle, mysteries offer excellent means for diagnostic assessment, because the teacher has the opportunity to listen to students' reasoning when they move the information strips in the group work part and when solutions are presented and discussed in the whole class debriefing session. This information can be used formatively, to design a lesson (or another mystery) to improve students' relational thinking. Explicitly mentioning the reason for organising a mystery lesson (design principle 2) offers the opportunity to organise a whole-class discussion on the learning gains concerning relational thinking in the debriefing session. With respect to design principle 4, three aspects deserve explicit attention. First, the relational strategies employed should be compared and discussed in whole class, as the results of our research showed considerable differences between groups in the effectiveness of the relating strategies they employed. Second, the interrelationships between different aspects of the problem must be emphasised, as our research revealed that the majority of the groups were not able to construct a coherent explanation of the mystery question. Third, attention needs to be given how to verbalise established relationships correctly, because "students who master the grammar of relational thinking are better able to recognise geographical relationships" (Favier & Van der Schee, 2014b, p. 178).

Using the findings of our first study and the available evidence, another design principle could be added to the six principles of Van Drie and Stoel (2017). The regular (and correct) use of conceptual frameworks like concepts maps and causal diagrams in assignments on multi-causal thinking seems very fruitful, because these frameworks help students to organise their relational thinking and they foster relational thinking (Cox, Elen, & Steegen, 2017; Favier & Van der Schee, 2014b, Kreeftenberg, 2017, Novak & Canas, 2008, Renshaw & Wood, 2011). This could be a seventh design principle for lessons to practise causal reasoning: students visualise their relational thinking in a conceptual framework.

Designing lessons to foster geographical relational thinking by educational geographers and geography teachers collaborating in professional learning communities will be fruitful in offering the geography teachers ways to implement relational thinking in their lessons in a structured manner. This is needed, because this is an indispensable thinking skill of powerful geographical knowledge. Powerful knowledge, a concept introduced by Young (2008), can be explained in two ways (Maude, 2017, 2018): by the characteristics of this knowledge or by its results, what it can do for those who obtain it. The latter perspective on powerful knowledge is described by Young (2008, p.14) as: "knowledge (that) refers to what the knowledge can do or what intellectual power it gives to those who have access to it. Powerful knowledge provides more reliable explanations and new ways of thinking about the world, and acquiring it can provide learners with a language for engaging in political, moral, and other kinds of debates". According to Maude, a geographical concept that offers students a new, powerful way of thinking, is holistic thinking, because "an awareness of interconnections should prompt students to think broadly and deeply in their geographical investigations, and to look for holistic and integrated explanations of phenomena" (2017, p. 31). This awareness resembles the development of the systems disposition mentioned by DeVane et al. (2010), and is practised by regularly working with assignments on complex, interrelated geographical problems. Of course, multi-causal, interrelated thinking is also practiced in other school subjects, like history, but geography as a subject offers students unique forms of relational thinking. First, geography is concerned with relationships between man and environment, bridging the natural and human sciences. Second, geography studies relationships between regions and their effects on man and environment. Third, geography relates scales, for example by investigating how global processes affect man and environment at specific places: places vary as a result of different physical, cultural, demographical, political or economic circumstances and as a result of different relative locations. This geographical relational thinking, underpinned with knowledge about how the world works, the human and physical phenomena and processes, is recognised by geography teachers in different European countries as the subject's contribution to the educational aim of promoting "informed and aware" citizens (Uhlenwinkel, Béneker, Bladh, Tani, & Lambert, 2017, p. 336).

Learning to think in interrelated, multi-causal relationships to understand, explain, evaluate or predict regional change is expected to foster students' relational understanding of our dynamic world and also their awareness of their own relational existence in it. This will offer them the needed capabilities to live responsible lives. With geography lessons contributing to the development of individual capabilities by delivering this powerful knowledge to students, teaching geography will be a very relevant undertaking in secondary education.

References

Adey, P. S., & Shayer, M. (1994). *Really raising standards: Cognitive intervention and academic achievement.* London, United Kingdom: Routledge.

Admiraal, W., Buijs, M., Claessens, W., Honing, T., & Karkdijk, J. (2017). Linking theory and practice: Teacher research in history and geography classrooms. *Educational Action Research*, *25*(2), 316-331.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of educational objectives.* Boston, MA: Allyn & Bacon.

Applis, S. (2014). Global learning in a geography course using the mystery method as an approach to complex issues. *Review of International Geographical Education Online (RIGEO)*, *4*(1), 58-70.

Arnold, R. D., & Wade, J. P. (2015). A definition of systems thinking: A systems approach. *Procedia Computer Science, 44*, 669–678.

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.

Baker, T. R., Battersby, S., Bednarz, S. W., Bodzin, A. M., Kolvoord, B., Moore, S., … & Uttal, D. (2015). A research agenda for geospatial technologies and learning. *Journal of Geography*, *114*(3), 118-130.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences*, *9*(4), 403-436.

Benninghaus, J. C., Mühling, A., Kremer, K., & Sprenger, S. (2019). Complexity in education for sustainable consumption: An educational data mining approach using mysteries. *Sustainability*, *11*(3), 722.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York, NY: Academic Press.

Brown, H. (2018). Infrastructural ecology as a planning paradigm: Two case studies. *International Journal of Sustainable Development and Planning, 13*(2), 187-196.

Centrale Examencommissie Vaststelling Opgaven CEVO. (2007). *Aardrijkskunde HAVO. Syllabus centraal examen.* Utrecht, The Netherlands: Centrale Examencommissie Vaststelling Opgaven vwo, havo, vmbo.

Chee, Y. S. (2010). Studying learners and assessing learning: A process-relational perspective on the learning sciences. *Educational Technology, 50*(5), 5-9.

Cito. (2009). *Toets en Itemanalyse. Havo aardrijkskunde 2009 tijdvak 1.* Arnhem, The Netherlands: Cito, Psychometrisch Onderzoekscentrum.

Cito. (2010). *Toets en Itemanalyse*. *Havo aardrijkskunde 2010 tijdvak 1.* Arnhem, The Netherlands: Cito, Psychometrisch Onderzoekscentrum.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, MI: Lawrence Erlbaum Associates, Publishers.

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

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

Cox, M., Elen J., & Steegen, A. (2019). Fostering students geographic systems thinking by enriching causal diagrams with scale. Results of an intervention study. *International Research in Geographical and Environmental Education, 29*(2), 112-128.

DeVane, B, Durga, S., & Squire, K. (2010). Economists who think like ecologists': Reframing systems thinking in games for learning. *E–Learning and Digital Media 7*(1), 3-20.

Faiola, A., Lopes, M., & Mooney, C. (2019) The price of "progress" in the Amazon. *The Washington Post,* June 28, 2019. Retrieved from: https://www.washingtonpost.com/world/2019/06/28/how-building-boom-brazilian-amazon-could-accelerate-its-deforestation/?arc404=true

Favier, T. T. (2011). *Geographic Information Systems in inquiry-based secondary geography education.* Enschede, The Netherlands: Ipskamp.

Favier, T. T., & Van der Schee, J. A. (2014a). The effects of geography lessons with geospatial technologies on the development of high school students' relational thinking. *Computers and Education, 76,* 225-236.

Favier, T. T., & van der Schee, J. A. (2014b). Evaluating progression in students' relational thinking while working on tasks with geospatial technologies. *Review of International Geographical Education Online, 4*, 155-181.

Fögele, J. (2017). Acquiring powerful thinking through geographical key concept. In C. Brooks, G. Butt & M. Fargher (Eds.), *The power of geographical thinking* (pp. 59-73). Cham, Switzerland: Springer.

Geographical Association GA. (2012). *Thinking Geographically* [pdf]. Retrieved from http://www.geography.org.uk/gtip/mentoring/geography/learning/conceptsandthinkinggeographically/

German Geographical Society DGfG (2012), *Educational standards in geography for the intermediate school certificate.* Bonn, Selbstverlag Deutsche Gesellschaft für Geographie, 2012. Retrieved from: https://vgdh.geographie.de/wp-content/docs/2014/10/geography_education.pdf

Golledge, R. G. (2002). The nature of geographic knowledge. *Annals of the Association of American geographers*, *92*(1), 1-14.

Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics, 49*(2), 193–223.

Hanson, S. (2004). Who are “we”? An important question for geography’s future. *Annals of the Association of American Geographers*, *94*(4), 715–22.

Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London, United Kingdom: Routledge.

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.

Hoekveld, G. A. (1971). De Geografische beschouwingswijze. In G.A. de Bruijne, G.A. Hoekveld & P.A. Schat, *Geografische Verkenningen 1, op zoek naar een geografisch wereldbeeld.* (pp. 11-46). Bussum, The Netherlands: Romen.

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*, 242-258.

Hopwood, N. (2009). UK high school pupils’ conceptions of geography: Research findings and methodological implications. *International Research in Geographical and Environmental Education, 18*(3), 185-197.

Huynh, N. T., & Sharpe, B. (2013). An assessment instrument to measure geospatial thinking expertise. *Journal of Geography*, *112*(1), 3-17.

Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students’ learning performance in natural science courses. *Computers and Education, 69*, 121–130.

International Geographical Union, Commission on Geographical Education IGU-CGE. (1992). *International charter on geographical education.* Retrieved from: https://www.igu-cge.org/wp-content/uploads/2018/02/1.-English.pdf

International Geographical Union, Commission on Geographical Education IGU-CGE. (2016). *2016 International charter on geographical education* [pdf]. Retrieved from https://www.igu-cge.org/charters.htm

Jackson, P. (2006). Thinking geographically. *Geography, 91*, 199-204.

Johnson, D. W., & Johnson, R. T. (2009). Energizing Learning: The Instructional Power of Conflict. *Educational Researcher, 38*(1), 37-51.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63-85.

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.

Karkdijk, J., Admiraal, W. F., & Van der Schee, J. A. (2019). Small-group work and relational thinking in geographical mysteries: A study in Dutch secondary education. *Review of International Geographical Education Online (RIGEO)*, *9*(2), 402-425.

Karkdijk, J., & Van der Schee, J. A. (2012). Leren denken met aardrijkskunde: Mysteries in de klas. *Geografie*, *21*(5), 38-41.

Karkdijk, J., Van der Schee, J. A. & Admiraal, W. F. (2013). Effects of teaching with mysteries on students' geographical thinking skills. *International Research in Geographical and Environmental Education*, *22*(3), 183-190.

Karkdijk, J., Van der Schee, J. A. & Admiraal, W. F. (2019). Students' geographical relational thinking when solving mysteries. *International Research in Geographical and Environmental Education, 28*(1), 5-21.

Koninklijk Nederlands Aardrijkskundig Genootschap KNAG. (2017). *Visie op het aardrijkskunde-onderwijs voor het primair en voortgezet onderwijs naar aanleiding van Curriculum.nu*. Retrieved from: https://geografie.nl/sites/geografie.nl/files/paragraph/attachment/file/KNAG%20Visiedocument%20Curriculum.nu_.pdf.

Kreeftenberg, G. M. (2017). *Geo-ICT als middel voor het relationeel denken van leerlingen binnen het aardrijkskundeonderwijs* (Master's thesis).

Lambert, D. (2004). *The power of geography* [pdf]. Retrieved from http://www.geography.org.uk/download/NPOGPower.doc

Lambert, D. (2014). Curriculum thinking, “capabilities” and the place of geographical knowledge in schools (Special Contribution). *Journal of Japanese Educational Research Association for the Social Studies*, *81*, 1-11.

Lambert, D. (2017). Thinking geographically, Chapter 2. In M. Jones (Ed.), *The handbook of secondary geography* (pp. 20–29). London, United Kingdom: Geographical Association.

Lambert, D., & Balderstone, D. (2010). *Learning to teach geography in the secondary school (2nd ed.).* London, United Kingdom: Routledge.

Lambert, D., Solem, M., & Tani, S. (2015) Achieving human potential through geography education: a capabilities approach to curriculum making in schools. *Annals of the Association of American Geographers 105*(4), 723-735.

Leat, D. (1997). Cognitive acceleration in geographical education. In D. Tilbury & M. Williams, (Eds.), *Teaching and Learning Geography* (pp. 143-153). London: Routledge.

Leat, D. (2001). *Thinking through geography* (2nd ed.). Cambridge, United Kingdom: Chris Kington Publishing.

Leat, D, & Nichols A. (2000a). Observing pupils’ mental strategies: signposts for scaffolding. *International Research in Geographical and Environmental Education, 9*(1), 19-33.

Leat, D., & Nichols, A. (2000b). Brains on the table: Diagnostic and formative assessment through observation. *Assessment in Education: Principles, Policy & Practice*, *7*(1), 103-121.

Leat, D., & Nichols, A. (2003). *Mysteries make you think*. Sheffield, United Kingdom: Geographical Association.

Leat, D., Van der Schee, J. & Vankan, L. (2005). New strategies of learning for geography: A tool for teachers’ professional development in England and The Netherlands. *European Journal of Teacher Education, 28*(3), 327-342.

Lee, J., & Bednarz, R. (2009). Effect of GIS learning on spatial thinking. *Journal of Geography in Higher Education*, *33*(2), 183-198.

Lezak, S. B., & Thibodeaux, P. H. (2016). Systems thinking and environmental concern. *Journal of Environmental Psychology*, 46, 143-153.

Lipman, M. (1991). *Thinking in Education*. Cambridge, United Kingdom: Cambridge University Press.

Massey, D. (2014). Taking on the world. *Geography 99*(1), 36-39.

Maude, A. (2017). Applying the concept of powerful knowledge to school geography. In C. Brooks, G. Butt & M. Fargher (Eds.), *The power of geographical thinking* (pp. 27-40). Cham, Switzerland: Springer.

Maude, A. (2018). Geography and powerful knowledge: A contribution to the debate. *International Research in Geographical and Environmental Education*, *27*(2), 179-190.

McClure, J., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching, 36*(4), 475-492.

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and instruction, 6*(4), 359-377.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal, 25*(1), 95-111.

Mehren, R., Rempfler, A., Buchholz, J., Hartig, J., & Ulrich-Riedhammer, E. M. (2018). System competence modelling: 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.

Morgan, J. (2017). Are we thinking geographically? In M. Jones, & D. Lambert (Eds.), *Debates in geography education (2nd ed.*) (pp. 287-297). London, United Kingdom: Routledge.

Morgan, J., & Lambert, D. (2005). *Geography. Teaching School Subjects 11-19*. London, United Kingdom: Routledge.

Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms.* Maidenhead/Philadelphia: Open University Press.

Moseley, D., Baumfield, V., Elliott, J., Gregson, M., Higgins, S., Miller, J., & Newton, D.P. (2005). *Frameworks for Thinking*. Cambridge: Cambridge University Press.

Murre, P. M. (2019). *Meer door minder. Het geven van een vakles vanuit christelijk perspectief.* Gouda, The Netherlands: Verloop drukkerij.

Nichols, A., & Kinninment, D. (2001). *More Thinking Through Geography*. Cambridge, United Kingdom: Chris Kington Publishing.

Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. Florida: Institute for Human and Machine Cognition (IHMC).

Novak J. D., & Gowin, D.B. (1984). *Learning how to learn.* New York, NY: Cambridge University Press.

Nuthall, G. A. (2005). The cultural myths and realities of classroom teaching and learning: A personal journey. *Teachers College Record, 107* (5*)*, 895-934.

Onderwijsraad. (2014). *Een eigentijds curriculum.* Retrieved from: file:///Users/jk/Downloads/Een-eigentijds-curriculum-1.7.pdf

Perkins, D. N., & Grotzer, T. A. (2001). Models and moves: The role of causal and epistemic complexity in students' understanding of science. Retrieved from: http://citeseerx.ist.psu.edu/viewdoc/summary?

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.

Renshaw, S., & Wood, P. (2011). Holistic understanding in geography education (HUGE) - An alternative approach to curriculum development and learning at Key Stage 3. *The Curriculum Journal, 22*(3)*,* 365-379.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2)*,* 177–192.

Roberts, M. (2010). What is "evidence-based practice" in geography education? *International Research in Geographical and Environmental Education*, *19*(2), 91-95.

Ruiz-Primo, M. A., & Shavelson, R. J., Li, M., & Schultz, S. E. (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educational Assessment, 7*(2)*,* 99-141.

Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores of two concept-mapping techniques. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *38*(2), 260-278

Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. *School Science and Mathematics, 102*(1), 33-44.

Shavelson R., Ruiz-Primo M., & Wiley, E. (2005). Windows into the mind. *Higher Education, 49*(4), 2005, 413-430.

Schuler, S. (2005). Mysterys als Lernmethode für das globales Denken. *Praxis Geographie*, *4*, 22-27.

Smith, J. (2015). Geographies of interdependence. *Geography, 100*(1), 12-19.

Srinivasan, M., McElvany, M., Shay, J. M., Shavelson, R. J., & West, D. C. (2008). Measuring knowledge structure: Reliability of concept mapping assessment in medical education. *Academic Medicine*, *83*(12), 1196–1203.

Stickler, C. M., Coe, M. T., Costa, M. H., Nepstad, D. C., McGrath, D. G., Dias, L. C. P., … & Soares-Filho, B. S. (2013). Dependence of hydropower energy generation on forests in the Amazon Basin at local and regional scales. *Proceedings of the National Academy of Sciences*, *110*(23), 9601-9606.

Stimpson, P. (1992). Assessment in geography: An evaluation of the SOLO taxonomy. In H. Schrettenbrunner & J. Van Westrhenen (Eds.), *Empirical research and geography teaching* (pp. 157-177). Utrecht; Amsterdam, The Netherlands: Koninklijk Nederlands Aardrijkskundig Genootschap; Centrum voor Educatieve Geografie Vrije Universiteit.

Szymanski, M. (2008). *Globales Lernen - das Mystery als Lernmethode für vernetztes Denken im Geographieunterricht.* Dresden, Germany: Technische Universität Dresden.

Turns, J., Atman, C. J., & Adams, R. (2000). Concept maps for engineering education: A cognitively motivated tool supporting varied assessment functions. *IEEE Transactions on Education, 43*(2)*,* 164-173.

Uhlenwinkel, A. (2013). Spatial thinking or thinking geographically? On the importance of avoiding maps without meaning. In T. Jekel et al. (Eds.), *GI_Forum 2013. Creating the GISociety* (294–305). Berlin, Germany: Wichmann.

Uhlenwinkel, A. (2017). Geographical thinking: is it a limitation or powerful thinking? In C. Brooks, G. Butt & M. Fargher (Eds.), *The power of geographical thinking* (41-53). Cham, Switzerland: Springer.

Uhlenwinkel, A., Béneker, T., Bladh, G., Tani, S., & Lambert, D. (2017). GeoCapabilities and curriculum leadership: Balancing the priorities of aim-based and knowledge-led curriculum thinking in schools. *International Research in Geographical and Environmental Education*, *26*(4), 327-341.

Van Boxtel, C., Van der Linden, J., Roelofs, E., & Erkens, G. (2002). Collaborative concept mapping: Provoking and supporting meaningful discourse. *Theory Into Practice, 41*(1), 40-46.

Van der Schee, J. A. (1987). *Kijk op kaarten, een empirisch onderzoek naar het gebruik van geografische denkvaardigheden bij het analyseren van kaarten door leerlingen uit het vierde leerjaar van het vwo.* Amsterdam, The Netherlands: Koninklijk Nederlands Aardrijkskundig Genootschap/Geografisch en Planologisch Instituut van de Vrije Universiteit Amsterdam.

Van der Schee, J. A. (2000). Helping children to analyse a changing world: Looking for patterns and relationships in space. In M. Robertson, & R. Gerber (Eds.), *The child's world: Triggers for learning* (pp. 214–231). Melbourne, Australia: Australian Council for Educational Research.

Van der Schee, J. A. (2009). Aardrijkskunde, wat is dat voor vak? In G. van den Berg, A. Bosschaart, R. Kolkman, I. Pauw, J. A. van der Schee, & L. Vankan (Red.), *Handboek vakdidactiek aardrijkskunde* (pp. 7-30). Amsterdam, The Netherlands: Landelijk Expertisecentrum Mens- en Maatschappijvakken..

Van der Schee, J., Leat, D., & Vankan, L. (2006). Effects of the use of thinking through geography strategies. *International Research in Geographical & Environmental Education*, *15*(2), 124-133.

Van der Schee, J., Vankan, L., & Leat, D. (2003). The international challenge of more thinking through geography. *International Research in Geographical and Environmental Education*, *12*(4), 330-343.

Van der Schee, J., Trimp, H., Béneker, T., & Favier, T. (2015). Digital geography education in the twenty-first century: Needs and opportunities. In M. Solari, A. Demirci, & J. A. van der Schee, *Geospatial Technologies and Geography Education in a Changing World* (pp. 11-20). Tokyo, Japan: Springer.

Van Drie, J. & Stoel, G. (2017). *De feiten voorbij. Bevorderen van causaal redeneren in de geschiedenisles*. Amsterdam, The Netherlands: Landelijk Expertisecentrum Mens- en Maatschappijvakken.

Van Voorst, R. S. (2014). *Get ready for the flood. Risk-handling styles in Jakarta, Indonesia.*

Vanderlinde, R., & Van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal*, *36*(2), 299-316.

Vankan, L.J.A.E., & Van der Schee, J. A. (2004). *Leren denken met aardrijkskunde.* Nijmegen, The Netherlands: Stichting Omgeving en Educatie.

Varró, K., & Van Gorp, B. (2020). Fostering a relational sense of place through video documentary assignments. *Journal of Geography in Higher Education*, *45*(1), 63-86.

Weeden, P., & Lambert, D. (2006). *Geography inside the black box: assessment for learning in the geography classroom.* London, United Kingdom: LG Assessment.

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

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching, 42*(2),166–184.

Young, M. (2008). From constructivism to realism in the sociology of the curriculum. *Review of Research in Education*, *32*(1), 1-28.

Zwart, R. C., Smit, B., & Admiraal, W. F. (2015). Docentonderzoek nader bekeken: Een reviewstudie naar de aard en betekenis van onderzoek door docenten. *Pedagogische Studiën*, *92*(2), 131-148.

Summary

Teaching geography in secondary education is teaching how people at specific locations on our planet try to make a living and how their interrelationships with people in other places and their interaction with the environment form regions and landscapes. This means that in order to understand our world, geography students have to learn to think relationally. They have to establish relationships between human activities in different places and between man and nature, from the local to the global, to be able to understand regional change. Relationships between human society and nature within a region are called vertical relationships. Relationships between regions are called horizontal relationships. Geographical relational thinking in this dissertation is described as: analysing, explaining, evaluating and/or predicting the horizontal and vertical relationships and the interactions between them, on different scales, that cause regional change. The scarce evidence available reveals that this multi-causal, interrelated thinking in geography is difficult for secondary school students.

In the late 1990s, the mystery was designed as one of the thinking strategies of the Thinking Through Geography programme. The strategy of the mystery has its focus on cause and effect. In order to explain a complex problem or event, formulated as a triggering, open question, students receive 16-30 strips containing information on the problem or the event, but without causal relationships. In small group work, students read the strips, move and categorise them to make groups and subgroups, establish linear or web-like causal relationships and formulate their explanation as an answer to the mystery question. In the whole-class debriefing session, groups present their explanation of the mystery question and also the strategy they used to understand the mystery and to construct the explanation. The strategy of the mystery can be used to foster students' relating skills and also as a diagnostic and formative instrument to assess and foster their thinking in geography. Although mysteries are widely used in geography lessons, very limited research has been conducted to obtain more evidence on their effect on students' relational thinking in geography.

The importance in secondary geography of learning to think relationally, the difficulties students have with relational thinking, the scarce evidence on students' relational thinking skills together with the almost complete absence of research on the strategy of the mystery to assess and foster relational thinking in geography lessons, provided the rationale for this

dissertation. The aim was to obtain evidence on students' geographical relational thinking, in order to gain more insight into possible ways that teachers could help their students to master this skill. Therefore, mysteries were used to elicit students' geographical relational thinking when working on a mystery in small groups, to analyse differences between small student groups in their relational thinking and to evaluate the effects of using mysteries on students' geographical relational thinking. The main research question of this dissertation was: *How does the use of mysteries in secondary geography education support geographical relational thinking in students?* The focus of the research was on the ability to think in interrelated, multi-causal relationships in order to give coherent explanations for a phenomenon or an event framed as a mystery question. The main research question was approached by four studies, each with its own research questions.

The first study

The research question of the first study was: What is the effect of the use of mysteries on students' geographical relational thinking? This research was designed as a quasi-experimental study, with a pre-test-post-test control group design. Seven secondary schools in the Netherlands participated. At each school one class was assigned to the experimental condition (where teaching with three or four mysteries was added to the regular geography curriculum) and another class of the same educational level and year to the control condition (the same number of lessons, the regular geography curriculum, but without mysteries). A total of 221 students, aged 15-17, in upper higher general secondary education (HAVO) and pre-university education (VWO), participated: 105 in the experimental classes and 116 in the control classes. The pre-test and post-test focused on variables linked to the Millennium Goals. The tests were devised to measure causal relational thinking in geography, operationalised as the ability to construct a relational scheme consisting of three or more interrelated concepts. We found that the progression in the number of established correct relationships was significant higher for students in the experimental condition than for students in the control condition. Thus, the repeated use of mysteries in combination with the representation of the constructed explanations as relational schemes had a positive effect on students' relational thinking skills in geography.

In the next three studies the relational thinking qualities of small groups working on a mystery were considered. The same database was used for these studies. Two geographical mysteries, one on the devastating landslides in Rio de Janeiro and another on the regular floods in

Jakarta, acted as the learning activity to elicit relational thinking and reasoning. The research was carried out in six schools in the Netherlands. In total 205 students aged 15-18 years in upper higher general secondary education (HAVO) and pre-university education (VWO), participated. The students collaborated in 69 small groups to understand and explain one of the two mysteries. Thirty-five groups worked on the Rio mystery and 34 on the Jakarta one. Students had to represent their explanations of the mystery as a concept map with labeled cause and effect arrows.

The second study

In this study, the relationships groups established to explain the Jakarta or the Rio mystery were analysed. The first research question was: Which geographical relationships do students in small groups establish to solve a mystery? Students used a wide range of relationships in order to explain the mysteries. However, the more concrete, less complex relationships dominated and many were almost or entirely directly and linearly formulated. Cross-links were by far in the minority, indicating that constructing a coherent explanation is a higher relational thinking level for many students.

The second research question that directed this study was: How coherent are the solutions to the mystery posed? A large proportion of the groups (about 40%) did not establish any correct cross-link in their explanation and delivered explanations that were not coherent, but consisted of separate, more or less linear relationships. Almost the same number of groups were able to formulate more coherent explanations by using one or more cross-links and branches in their concept maps. Therefore, groups differed greatly in their ability to think in interrelated multi-causal relationships.

The third study

The third study focused on the differences between small groups in their explanations of the mysteries and how these differences are to be understood. The first research question concerned differences between groups expressed as differences in levels of the SOLO taxonomy. The SOLO taxonomy expresses the structure of learning outcomes: how interrelated is the presented answer on the assignment? This research question was: How does geographical relational thinking in terms of the SOLO taxonomy differ between groups? This part of the study overlaps with the second study, and so does the conclusion that groups differed greatly in their ability to think and reason in interrelated causal relationships.

The next research question of this study was designed to explore possible differences between the two mysteries in the quality of geographical relational thinking that they provoked: How does geographical relational thinking in groups differ between the two mysteries? We found that the design of a mystery seemed to influence the complexity of the geographical relational thinking provoked. In the Rio mystery, which was designed with fewer concepts that were easier to connect linearly with the question of the mystery than the Jakarta mystery, students were asked to think more in interrelated relationships. As a result, the mean total proposition score on the Rio mystery was significantly lower than on the Jakarta mystery.

The third research question of this study was: How can differences between groups in geographical relational thinking be explained by characteristics and collaborative behaviour of the groups? We did not find any significant correlations between the group characteristics educational level, educational year or previous achievements in geography and geographical relational thinking. The only significant predictor that explained a small proportion of the variance of three indicators of relational thinking was the amount of on-task discussion the students had on the problem of the mystery before they started to construct the concept map.

The fourth study

The fourth study was designed to obtain more insights into the differences between groups in their geographical relational thinking by analysing the strategies they employed to tackle a mystery. The research question was: What are the differences between low-performing and high-performing groups in their strategies to understand a geographical mystery? We found the most remarkable differences between low-performing and high-performing groups in the relating strategies they employed. Four different relating strategies were found: (1) looking for a logical and/or chronological chain of events; (2) reasoning from categories/sets; (3) relating separate strips one by one to the mystery question; and (4) establishing causal relationships between the strips, the webbing strategy. All high-performing groups used the latter as their main strategy. The low-performing groups used one or more of the other relating strategies as their main relating strategy, which apparently resulted in a more partial, incoherent or less coherent explanation of the mystery. The webbing strategy apparently resulted in a more coherent explanation of the mystery.

Conclusions

The main conclusions are:

(1) The repeated use of mysteries in combination with groups' representation of their explanation of the mystery as a concept map (causal diagram) can have a significant positive effect on students' relational thinking in geography.

(2) The quality of geographical relational thinking of a large proportion of small groups working on a mystery seems generally low, although the differences between student groups appear to be considerable:
   - the more concrete, fairly easy to establish relationships seem to dominate the explanations;
   - where a large proportion of small student groups appears to be able to establish only more or less linear, unconnected relationships in order to explain a complex, interrelated problem, a comparable proportion seems to be able to construct a more coherent explanation by using cross-links.

(3) Small student groups seem to differ in their ability to explain a complex problem such as presented by a mystery in group effort and not in particular group characteristics like educational level, educational year of geography grades. The two differentiating elements of group effort that result in a more extensive and more coherent explanation, might be:
   - an extensive on-task group discussion to understand the connections between the different aspects of the mystery problem before the answer to the mystery question was constructed; and
   - the use of the webbing strategy as the main relating strategy employed to understand the mystery.

With respect to the aim of this research, we conclude that teachers can help students to foster students' geographical relational thinking by:

(e) practising strategies (like mysteries) with more complex assignments to think in interrelated, multi-causal relationships, because a large proportion of the student groups thought in more or less directly, unconnected linear relationships;

(f) using these strategies in a progressive order from assignments with more concrete concepts easy to connect with the mystery problem to assignments with more abstract background information that demands thinking in interrelationships before an answer to the mystery problem can be formulated;

(g) giving explicit attention to cross-links or interrelationships in the whole-class debriefing session; and

(h) giving explicit attention in the debriefing session to the employed relating strategies and the power of the webbing strategy to tackle a complex interrelated problem.

Implications

The Dutch national geography curriculum of higher general secondary education (HAVO) and pre-university education (VWO) expects students to be able to analyse geographical issues by using horizontal and vertical relationships. In the light of the findings and conclusions of this dissertation, geographical relational thinking needs more explicit attention at the level of geography lessons in class. This could be done by implementing in lessons the four conclusions (a – d) on how to foster geographical relational thinking in class. Regular use in class of assignments on multi-causal, interrelated thinking will not only support practice in the establishment and correct formulation of relationships, but will also contribute to students' inclination to structure information on a complex problem by establishing causal relationships between elements and by looking for interrelationships between these relationships.

A broader approach is desirable, however, to implement relational thinking in geography lessons. Educational geographers and geography teachers could emulate the efforts in the subject of history. In a design study conducted to increase the Pedagogical Content Knowledge of history teachers on teaching causal relational thinking, researchers and history teachers developed a rubric to assess causal reasoning and also design principles for lessons to explicitly foster causal relational thinking. Educational geographers and geography teachers could also collaborate to design lessons that foster geographical relational thinking using the findings of their history colleagues.

Structural attention to geographical relational thinking in geography lessons is needed, because learning to think in interrelated, multi-causal relationships to understand, explain, evaluate or predict regional change will foster students' relational understanding of our dynamic world and also their awareness of their own relational existence in it. This will offer them the needed capabilities to live responsible lives.

Samenvatting

Lesgeven in aardrijkskunde in het voortgezet onderwijs is leerlingen leren hoe mensen op verschillende locaties op aarde proberen te voorzien in hun bestaan. En hoe daarbij hun relaties met mensen op andere plaatsen, samen met de interactie die ze hebben met hun omgeving, regio's en landschappen vormen. Dit betekent dat leerlingen, willen ze de wereld om hen heen gaan begrijpen, in relaties moeten leren denken. Ze moeten verbanden leren zien tussen menselijke activiteiten in verschillende regio's alsook tussen mens en natuur, en dat op verschillende ruimtelijke schaalniveaus, om regionale veranderingen te kunnen begrijpen. Relaties tussen de menselijke samenleving en de natuur in een bepaalde regio heten verticale relaties en relaties tussen verschillende regio's heten horizontale relaties. Geografisch relationeel denken, waarover het in dit proefschrift gaat, is omschreven als: het analyseren, verklaren, waarderen en/of voorspellen, op verschillende ruimtelijke schaalniveaus, van de horizontale en verticale relaties en de interacties daartussen die zorgen voor regionale verandering. Het weinige onderzoek dat er is gedaan naar dit meervoudig, causaal relationeel denken zoals dat bij het vak aardrijkskunde gevraagd wordt, laat zien dat leerlingen dit moeilijk vinden.

Aan het einde van de jaren negentig van de vorige eeuw is het mysterie ontworpen als een van de denkstrategieën van het programma *Leren denken met aardrijkskunde*. De werkvorm mysterie is gericht op het leggen van oorzaak-gevolgrelaties. Om een ingewikkeld probleem of een ingewikkelde gebeurtenis te kunnen verklaren, krijgen leerlingen op 16 tot 30 losse strookjes met informatie, zonder oorzaak-gevolgrelaties. In het groepswerk lezen leerlingen de strookjes, categoriseren ze, verschuiven ze, leggen lineaire of meer webvormige verbanden en formuleren vervolgens hun uitleg van het mysterie. In de klassikale nabespreking presenteren sommige groepen hun uitleg van het mysterie en hun strategie om tot deze verklaring te komen. De werkvorm mysterie kan worden gebruikt om het leren relateren bij leerlingen te bevorderen, maar ook als een diagnostisch of formatief instrument om het niveau van het geografisch relationeel denken bij leerlingen te bepalen en hen hiermee verder te helpen. Hoewel mysteries in allerlei landen worden gebruikt in aardrijkskundelessen, is er toch heel weinig onderzoek gedaan naar het effect ervan op het geografisch relationeel denken van leerlingen.

Het belang van relationeel denken bij aardrijkskunde in het voortgezet onderwijs, de moeilijkheden die leerlingen hebben hiermee, het weinige onderzoek dat er is gedaan naar het geografisch relationeel denken van leerlingen en naar het belang van de werkvorm mysterie voor deze denkvaardigheid, vormde de rationale voor deze dissertatie. Het doel van het onderzoek was om meer te weten te komen over het geografisch relationeel denken van leerlingen, opdat er manieren kunnen worden gevonden hoe docenten hun leerlingen kunnen helpen deze denkvaardigheid beter onder de knie te krijgen. Daarvoor werden mysteries gebruikt om het geografisch relationeel denken van leerlingen tijdens het groepswerk te activeren, om verschillen tussen groepen leerlingen wat betreft hun relationeel denken te kunnen analyseren en om het effect van het gebruik van mysteries op het geografisch relationeel denken te kunnen bepalen. De hoofdvraag van het onderzoek voor dit proefschrift was: *Hoe bevordert het gebruik van de werkvorm mysterie in aardrijkskundelessen in het voortgezet onderwijs het geografisch relationeel denken van leerlingen en maakt dit inzichtelijk?* De focus van het onderzoek lag daarbij op het kunnen denken in onderling verbonden, meervoudig causale relaties om een samenhangende verklaring te kunnen geven voor een als en mysterie gepresenteerd verschijnsel of gebeurtenis. Deze hoofdvraag is beantwoord aan de hand van vier studies, ieder met zijn eigen onderzoeksvragen.

De eerste studie

De onderzoeksvraag van de eerste studie was: wat is het effect van het gebruik van mysteries op het geografisch relationeel denken van leerlingen? Dit was een quasi-experimenteel onderzoek met een *pre-test-post-test control group design*. Zeven scholen in Nederland deden mee aan het onderzoek. In iedere school was er een experimentele klas en een klas uit hetzelfde leerjaar die als controlegroep diende. Beide klassen kregen dezelfde leerstof uit dezelfde methode, maar in de experimentele klas werden ook drie of vier mysteries uitgevoerd. Er deden in totaal 221 leerlingen uit de bovenbouw van havo en vwo mee: 105 ervan zaten in een experimentele klas en 116 in een klas die diende als controlegroep. De pre-test en de post-test waren afgeleid van de Millenniumdoelen en waren ontworpen om causaal relationeel geografisch denken te meten aan de hand van door leerlingen te maken relatieschema's met tenminste drie concepten. Uit de analyse bleek dat leerlingen uit de experimentele klassen een significant grotere vooruitgang maakten in het aantal juist gelegde relaties dan leerlingen in de controlegroepen. De herhaalde inzet van mysteries gecombineerd met de presentatie van de uitwerking ervan als een conceptmap bleek een significant positief effect te hebben op het geografisch relationeel denken van leerlingen.

In de volgende drie studies werd de kwaliteit van het geografisch relationeel denken van leerlingen onderzocht. Voor deze drie studies werd dezelfde database gebruikt. Twee geografische mysteries werden gebruikt om het relateren en redeneren van leerlingen te activeren, een over de verwoestende aardverschuivingen in Rio de Janeiro van 2010 en een andere over de frequente overstromingen van delen van Jakarta. Het onderzoek werd uitgevoerd op zes scholen in Nederland. In totaal 205 leerlingen uit de bovenbouw van havo en vwo deden mee aan het onderzoek. Ze werkten samen in 69 kleine groepjes aan een van de mysteries, 35 groepen aan het Rio-mysterie en 34 groepen aan het Jakarta-mysterie. De uitwerking van het mysterie moesten ze presenteren als een concept map met gelabelde oorzaak-gevolgpijlen.

De tweede studie

Voor deze studie zijn de relaties die groepen legden om het Rio- of Jakarta-mysterie te verklaren, geanalyseerd. De eerste onderzoeksvraag van deze studie was: welke geografische relaties leggen leerlingen, samenwerkend in kleine groepjes, om het mysterie te verklaren? Leerlingen gebruikten veel verschillende relaties om de mysteries te verklaren, maar de meer concrete, eenvoudiger relaties overheersten en de relaties werden vaak direct en lineair gelegd. Dwarsverbanden, zogenaamde crosslinks, werden maar weinig gelegd en dat geeft aan dat het geven van een echt samenhangende verklaring een volgende stap is in het relationele denken van leerlingen.

De tweede onderzoeksvraag van deze studie was: hoe samenhangend zijn de oplossingen die gegeven zijn voor de mysteries? Een groot gedeelte van de groepen (40%) legde niet een correcte crosslink in hun verklaring, maar gaf verklaringen die onsamenhangend waren, bestaande uit van elkaar gescheiden, meer of minder lineaire relaties. Een ongeveer even groot aantal groepen was wel in staat een min of meer samenhangende verklaring te geven door het gebruik van crosslinks en branches in hun conceptmaps. Groepen verschilden dus sterk in hun vermogen om te denken in onderling verbonden, meervoudig causale relaties.

De derde studie

De derde studie zoomde in op de verschillen tussen de groepen wat betreft hun uitleg van de mysteries en op de mogelijke verklaring voor deze verschillen. De eerste onderzoeksvraag betrof de verschillen tussen groepen in de uitleg van de mysteries, uitgedrukt als verschillen in niveau van de SOLO-taxonomie van de verklaringen. De SOLO-taxonomie geeft de mate van interne samenhang, de structuur, van de gegeven uitleg weer. De onderzoeksvraag was: hoe verschillen groepen in hun geografisch relationeel denken, uitgedrukt als niveau van de SOLO-taxonomie? Deze onderzoeksvraag overlapt met de laatste van de vorige studie en zo ook de conclusie: groepen verschilden sterk in hun vermogen om te denken en te redeneren in onderling verbonden causale relaties.

De volgende onderzoeksvraag van deze studie was gericht op verschillen tussen de twee mysteries in de kwaliteit van het geografisch relationeel denken dat ze opriepen: hoe verschilt het geografisch relationeel denken in de groepen tussen de twee mysteries? Het bleek dat het ontwerp van een mysterie de complexiteit van het geografisch relationeel denken van leerlingen beïnvloedt. Het Rio-mysterie had minder concepten die gemakkelijk rechtstreeks waren te verbinden met de vraag van het mysterie dan het Jakarta-mysterie, en vroeg van leerlingen meer te denken in samenhangende relaties. Het gemiddeld aantal juist geformuleerde relaties in de conceptmaps was voor het Rio-mysterie dan ook significant lager dan voor het Jakarta-mysterie.

De derde onderzoeksvraag van deze studie was: hoe kunnen de verschillen tussen groepen in geografisch relationeel denken worden verklaard met behulp van groepskenmerken en groepsgedrag? Er zijn geen significante correlaties gevonden tussen het geografisch relationeel denken in de groepen en de groepskenmerken schoolsoort, schooljaar of tot nu toe behaalde cijfers voor aardrijkskunde. De enige predictor die een klein deel van de variantie van de drie indicatoren voor geografisch relationeel denken verklaarde, was de hoeveelheid *on-task* gesproken woorden in het groepswerk vóór de start van het maken van de concept map.

De vierde studie

De vierde studie was ontworpen om meer inzicht te krijgen in het geografisch relationeel denken van groepen door het analyseren van de strategieën die groepen inzetten om het mysterie te ontrafelen. De onderzoeksvraag was: welke verschillen zijn er tussen de groepen die het mysterie het slechtste uitleggen en groepen die het mysterie het beste uitleggen, in hun gehanteerde strategieën om het mysterie te begrijpen? De meest duidelijke verschillen tussen de in dit opzicht slechtste en beste groepen werden gevonden in hun relateerstrategieën. Er werden vier relateerstrategieën gevonden: (1) proberen een logische of een chronologische volgorde te vinden in de gebeurtenissen op de informatiestrookjes; (2) redeneren vanuit gemaakte categorieën/sets van strookjes; (3) proberen individuele strookjes een voor een te koppelen aan de vraag van het mysterie; en (4) het leggen van causale relaties tussen de strookjes, de webbing strategie. De best presterende groepen gebruikten allemaal de laatste strategie als belangrijkste om het mysterie te begrijpen. De slechtst presterende groepen gebruikten een of meer van de eerste drie relateerstrategieën als de belangrijkste om het mysterie te begrijpen. De webbing strategie resulteerde blijkbaar in de meest samenhangende verklaring van het mysterie.

Conclusies

De belangrijkste conclusies zijn:

(1) het herhaalde gebruik van mysteries gecombineerd met het laten uitwerken van de uitleg ervan als een concept map (causaal diagram) kan het geografisch relationeel denken van leerlingen significant verbeteren;

(2) de kwaliteit van het geografisch relationeel denken van een groot deel van groepen leerlingen die aan een mysterie werken, lijkt over het algemeen laag, hoewel er grote verschillen zijn tussen groepen:
   - de meer concrete, niet te moeilijk te leggen relaties lijken te domineren in de verklaringen;
   - een groot deel van de groepen lijkt alleen maar in staat te zijn om min of meer lineaire, niet-verbonden relaties te leggen als verklaring van een complex probleem, hoewel een ongeveer even groot gedeelte in staat is een meer samenhangende verklaring te geven door dwarsverbanden, crosslinks, te gebruiken;

(3) groepjes leerlingen verschillen van elkaar in groepsgedrag wat betreft het kunnen uitleggen van een complex probleem zoals een mysterie, maar groepskenmerken zoals schoolsoort, schooljaar of behaalde cijfers voor aardrijkskunde lijken er niet toe te doen. De twee onderdelen van groepsgedrag die leiden tot een meer complete, samenhangende uitleg lijken te zijn:

- een omvangrijke on-task groepsdiscussie om de relaties tussen verschillende aspecten van het mysterie te begrijpen voordat een antwoord op de vraag van het mysterie wordt opgesteld; en
- het gebruikmaken van de webbing strategie als de belangrijkste relateerstrategie om het mysterie te begrijpen.

Wat betreft de beoogde doelen van dit onderzoek kan worden geconcludeerd dat docenten hun leerlingen kunnen helpen bij het verbeteren van hun geografisch relationeel denken door:

(a) ze te laten oefenen met werkvormen zoals mysteries die vragen in samenhangende, meervoudig causale relaties te denken, want een groot deel van de groepen lijkt te denken in min of meer directe, niet-verbonden lineaire relaties;
(b) deze werkvormen zoals mysteries in een opklimmende moeilijkheidsgraad te gebruiken, van opdrachten met wat meer concrete concepten die gemakkelijk rechtstreeks te relateren zijn aan de op te lossen vraag naar die waar meer abstracte achtergrondinformatie in onderlinge samenhang bij betrokken moet worden om tot een antwoord te komen.
(c) in de nabespreking expliciet te aandacht te geven aan dwarsverbanden, crosslinks, tussen de gelegde relaties en
(d) in de nabespreking expliciet aandacht geven aan de gebruikte relateerstrategieën en de kracht van de webbing strategie om een complex, samenhangend probleem te kunnen ontrafelen.

Implicaties

Het havo- en vwo-examenprogramma aardrijkskunde in Nederland verwacht dat leerlingen horizontale en verticale relaties kunnen gebruiken om een vraagstuk te kunnen analyseren. Gezien de uitkomsten van het onderzoek van deze dissertatie, zou geografisch relationeel denken veel meer aandacht moeten krijgen in aardrijkskundelessen. Dit kan worden gedaan door wat er gezegd is onder a tot en met d van bovenstaande conclusies toe te passen in de lessen. Het regelmatig gebruiken van werkvormen die een beroep doen op het denken in samenhangende, meervoudig causale relaties zal leerlingen oefenen in geografisch relationeel denken, maar zal ook bijdragen tot een gericht zijn op het structureren van informatie door de causale samenhangen tussen onderdelen van een probleem en de dwarsverbanden hiertussen helder te krijgen.

Een bredere benadering om geografisch relationeel denken in aardrijkskundelessen te implementeren is echter ook wenselijk. Hiervoor kunnen onderwijsgeografen en aardrijkskundedocenten gebruik maken van wat er op dit vlak bij geschiedenis ontwikkeld is. In een ontwerpstudie om de *Pedagogical Content Knowledge* van geschiedenisdocenten ten aanzien van het lesgeven in causaal relateren te bevorderen, hebben onderzoekers en docenten geschiedenis een rubric ontwikkeld om het causaal redeneren van leerlingen te kunnen beoordelen. Ook zijn er ontwerpprincipes geformuleerd om het causaal denken te bevorderen. Onderwijsgeografen en docenten aardrijkskunde kunnen samenwerken om lessen te ontwerpen die het geografisch relationeel denken van leerlingen bevorderen, gebruikmakend van de opbrengsten van hun collega's bij geschiedenis.

Het structureel aandacht geven aan geografisch relationeel denken in aardrijkskundelessen is nodig, want door te leren denken in onderling samenhangende, meervoudig causale relaties om veranderingen in de leefgebieden van mensen te leren begrijpen, te beoordelen en te voorspellen, zullen leerlingen inzicht verwerven in een steeds veranderende wereld en in hun eigen daarmee sterk verweven bestaan. Dat inzicht zal hen toerusten om in deze wereld verantwoord te leven en te handelen.

Papers in this dissertation and contributions of co-authors

The chapters 2-5 are based on four papers in scientific journals of research in geographical education. For each paper the contribution of the author and the co-author is described.

1. Chapter 2 is based on: Karkdijk, J., Van der Schee, J. A., & Admiraal, W. F. (2013). Effects of teaching with mysteries on students' geographical thinking skills. *International Research in Geographical and Environmental Education*, *22*(3), 183-190.

   Jan Karkdijk was the first author of this paper and Joop van der Schee and Wilfried Admiraal were the supervisors and the co-authors. Jan Karkdijk approached the participating teachers from secondary schools, designed three mysteries, the pre-test and post-test and wrote the instructions for the teachers and the students. Joop van der Schee designed the Benghazi mystery and reviewed the other mysteries. Jan Karkdijk and Wilfried Admiraal collaboratively analysed and interpreted the data. The manuscript was reviewed and revised repeatedly by the two supervisors.

2. Chapter 3 is based on: Karkdijk, J., Van der Schee, J. A., & Admiraal, W. F. (2019). Students' geographical relational thinking when solving mysteries. *International Research in Geographical and Environmental Education, 28*(1), 5-21.

   Jan Karkdijk was the first author of this paper and Joop van der Schee and Wilfried Admiraal were the supervisors and the co-authors. Jan Karkdijk approached the participating teachers from the secondary schools, designed both mysteries and gathered, analysed and interpreted all data. Joop van der Schee reviewed the Jakarta mystery. The manuscript was reviewed and revised repeatedly by the two supervisors.

3. Chapter 4 is based on: Karkdijk, J., Admiraal, W. F., & Van der Schee, J. A. (2019). Small-group work and relational thinking in geographical mysteries: A study in Dutch secondary education. *Review of International Geographical Education Online (RIGEO)*, *9*(2), 402-425.

   Jan Karkdijk was the first author of this paper and Wilfried Admiraal and Joop van der Schee were the supervisors and the co-authors. The database was the same as the database of the previous paper. Wilfried Admiraal carried out the multilevel variance components analyses and Jan Karkdijk analysed and interpreted the other statistical analyses. The manuscript was reviewed and revised repeatedly by the two supervisors.

4. Chapter 5 is based on: Karkdijk, J., Van der Schee, J. A., & Admiraal, W. F. (2021). Strategies used by small student groups to understand a geographical mystery. *J-READING. Journal of research and didactics in geography, 10*(1), 5- 21.

   Jan Karkdijk was the first author of this paper and Joop van der Schee and Wilfried Admiraal were the supervisors and the co-authors. The same database was used as for both previous papers. Jan Karkdijk gathered, analysed and interpreted all data. The manuscript was reviewed and revised repeatedly by the two supervisors.

Dankwoord

In 2009 kreeg ik de gelegenheid van mijn school, het Calvijn College te Goes, de master Academisch Meesterschap aan de Universiteit van Amsterdam te doen. Gaandeweg de master werd mijn belangstelling voor en plezier in het doen van onderzoek in onderwijs sterker en zocht ik naar een mogelijkheid dit na het beëindigen van de master te continueren. In november 2011 was ik een van degenen aan wie door het NWO een Promotiebeurs voor Leraren werd toegekend. Gedreven door interesse en ambitie startte ik in augustus 2012. Echter, enkele maanden later werd duidelijk dat de ernstige ziekte waarmee mijn lieve dochter Marieke kampte, doorzette. Haar situatie verslechterde en ik legde het werk aan het onderzoek neer. In de periode na haar overlijden in mei 2013 stond ik voor de keus om het onderzoek weer op te pakken of niet. De ambitie was inmiddels verdwenen, maar de interesse was er nog wel. In augustus 2014 heb ik de herstart gemaakt, geïnspireerd door de moed en het doorzettingsvermogen die mijn dochter tot op het eind van haar leven heeft laten zien, en vanwege de mogelijkheid die ik met de promotiebeurs had gekregen, een kans die anderen ook graag hadden gehad. Voor het kunnen doen en voltooien van dit onderzoeksproject komt daarom mijn dank toe aan de Heere God, Die mij hiervoor niet alleen de mogelijkheid maar ook de kracht gegeven heeft.

Mijn dank gaat ook uit naar het bestuur van mijn school in de persoon van de oud-voorzitter van het College van Bestuur, Bert Vogel, die mij in de gelegenheid stelde en ook aanmoedigde tot het aanvragen en aanvaarden van de Promotiebeurs voor Leraren, en die in de jaren daarna ook steeds zijn interesse voor de resultaten van het onderzoek toonde. Ook zijn opvolger, Jan Bakker, dank ik hartelijk voor de aangeboden mogelijkheden en ook stimulans het traject daadwerkelijk af te ronden.

Bijzondere woorden van dank zijn voor mijn promotoren Joop van der Schee, emeritus hoogleraar onderwijs-geografie aan de Vrije Universiteit in Amsterdam en Wilfried Admiraal, hoogleraar onderwijswetenschappen aan de Universiteit van Leiden. Al tijdens de master Academisch Meesterschap hadden we contact en hielpen jullie mij de aanvraag voor de promotiebeurs in te dienen. Ik dank jullie hartelijk voor al het begrip, het meeleven, de wijze adviezen, de mild getoonzette correcties, de stimulerende opmerkingen en zeker ook de humor waarmee jullie de begeleiding gestalte hebben gegeven. Ik heb onze contacten en jullie manier van begeleiding echt enorm gewaardeerd en ik heb er ook heel wat van geleerd.

Joop, als ik maar iets sneller was geweest, was je de eerste promotor gebleven, de plaats die je toekwam. Nu werd op het einde van het traject Martijn Meeter, hoogleraar onderwijswetenschappen aan de Vrije Universiteit in Amsterdam, als promotor toegevoegd. Martijn, hartelijk dank voor de welwillendheid waarmee je dit traject bent ingestapt en je positief kritische vragen over het proefschrift dat al in concept gereed lag, waardoor het werd versterkt. Daarnaast ben ik de leden van de leescommissie zeer erkentelijk voor het beoordelen van dit proefschrift. Op deze plaats wil ik ook mijn dank uitspreken aan Jacquelien Bulterman. Jacquelien, we hadden contact over allerlei zaken, van geloof tot aan het doen van onderzoek en je hebt me erg gesteund al die jaren door je hartelijke meeleven met onze persoonlijke situatie en je meedenken met de opzet en inhoud van mijn onderzoek.

Mijn dank gaat ook uit naar leerlingen en collega's die inhoudelijk hebben bijgedragen aan het tot stand komen van dit onderzoek. In de eerste plaats aan al die leerlingen die door hun deelname aan het onderzoek of door het moeizaam transcriberen van filmopnames dit onderzoek mogelijk maakten. Vervolgens ook naar de vele collega-aardrijkskundedocenten die bereid waren hun lessen in te vullen ten bate van mijn onderzoek, hetzij door een aantal mysteries in de klas te doen, een voor-en natest af te nemen dan wel door ruimtes in hun school te reserveren waar ik groepjes leerlingen kon filmen. In het bijzonder wil ik Marnix Bloemhof, Francelle Trouwborst, Nico Jan Valk en Safrien van de Leemkolk bedanken die als *second raters* de betrouwbaarheid van mijn onderzoeksresultaten verhoogden. Graag bedank ik ook de collega-onderzoekers die meedachten met mijn onderzoek en nuttige tips en denkduwtjes gaven, zoals Tim Favier en Iris Pauw. Patricia Grocott, collega en vriend Henk van Dongen en collega Laura Brouwer dank ik voor de nauwgezette redactie van mijn Engels of de vertaling van Nederlandse teksten naar het Engels.

Uiteraard dank ik van harte mijn ouders en schoonouders, familieleden, vrienden en collega's die de afgelopen jaren door hun vragen naar de voortgang van mijn onderzoek en het stimuleren daartoe hebben meegeholpen dat dit proefschrift tot een goed einde is gekomen. In het bijzonder wil ik daarbij mijn dochter Joanne en haar man Pieter en mijn dochter Esther en haar vriend Marcus danken. Jullie hebben niet opgehouden door belangstelling te tonen, aan te moedigen, of ook door mee te lezen met concepthoofdstukken mij te stimuleren dit proefschrift ook daadwerkelijk af te ronden. En dat geldt in het bijzonder voor jou Madeleen, die als mijn geweldige vrouw steeds met belangstelling voor het werk mij moed insprak, hoop gaf, hielp relativeren en stimuleerde. Je trouw en liefde in voor- en tegenspoed waren en zijn een dragende kracht in mijn leven. Daarom draag ik jou dit proefschrift op.

Appendix A

Pre-test and post-test, instruction and questionnaire for participating teachers

1. Teacher instruction pre-test and post-test

The pre-test is set up in such a way that students can get to work without instruction from the teacher. The test takes about 40 - 45 minutes: one lesson.

Present the test as an examination:

1. **The test will be taken in the classroom, under the supervision of a teacher.**
2. **Give students separate test papers on which they write their name and class.**
3. **Hand out the test.**
4. **The students take the test in silence, without mutual consultation (prohibit cheating).**

At the end of the test, please take the papers, put them in an envelope per class and indicate which class is your experiment class.

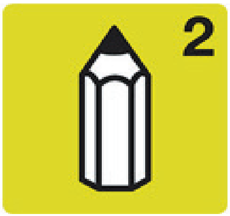
2. Student instruction pre-test and post-test

VN-Milleniumgoals

Millenium goal 1 **Poverty cut by half and less people suffering from hunger (poverty, hunger)**

In 2015, the percentage of people living in extreme poverty must have been reduced by at least half compared to 1990. Extreme poverty means that someone has less than $1.25 a day to spend. In 1990, 1.8 billion people lived in extreme poverty, representing 41.7 percent of the world's population.
Also, the percentage of people suffering from hunger must be halved by 2015. That percentage was 20 percent in 1990. At the time, one third of all children under five were malnourished.

Millenium goal 2 **All children can attend school (education participation)**

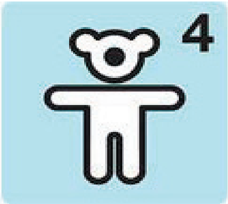


In developing countries, millions of children are not yet attending school. Education provides children an opportunity to develop further, which increases their chances in life. This will ultimately benefit society as a whole. It is therefore necessary to ensure that all children around the world can attend and complete primary education by 2015.

Millenium goal 3 **Equality of men and women (position of the woman)**

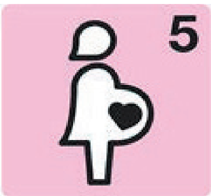
The third millennium goal is about women's rights. As laid down in international human rights treaties, men and women formally have the same rights. However, in practice, the documentation proves not to be sufficient in preventing subordination of women. Therefore, the third millennium goal objective aims to have just as many girls as boys going to school in primary, secondary and higher education by 2015.

Millenium goal 4 **Less infant mortality (infant mortality)**



In 2015, the percentage of children dying before their fifth year must have been reduced by two-thirds compared to 1990, when 12.4 million children died worldwide.
Many children die from diseases such as diarrhoea, measles, pneumonia or malaria that could have been prevented or cured.

Millenium goal 5 **Improvement of the mothers' health (mother mortality, contraceptives, abortion, number of children per woman)**



Every year, hundreds of thousands of women die from the consequences of their pregnancy. Bleeding, infections and high blood pressure are common causes of death which can be prevented with good medical attention and maternity care. The fifth millennium goal wants maternal mortality to be reduced with 75 percent in 2015 compared to 1990.
Maternal mortality is often the result of illegal and dangerous abortions. Availability of contraceptives and access to a safely performed abortion is therefore vital, as is the right of women to make their own choices about their sexuality and whether or not to have children. This is called reproductive health. By 2015, all women must have access to this at the latest.

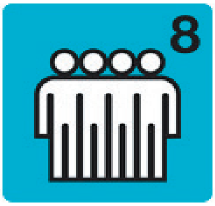
Millenium goal 6 **Fighting HIV / AIDS, malaria and other deadly diseases (Aids, malaria)**

It has been agreed that the spread of HIV/AIDS must have been stopped in 2015. In addition to the spread of HIV/AIDS, malaria and other major diseases such as TB must also be stopped in 2015.
Malaria is caused by mosquito bites. Among other things, this disease can be stopped by having children sleep under insecticide-treated nets (mosquito nets).

Millenium goal 7 **More people in a sustainable environment (environment, deforestation, soil depletion, water, hygiene)**

Air pollution, deforestation and depletion of agricultural land are direct threats to people's living conditions and health. A sustainable environmental policy is therefore of vital importance.
The seventh millennium goal states that, compared to 1990, the percentage of people without access to safe drinking water and sanitation must be halved by 2015.

Millennium goal 8 Global partnership for development (development aid, debt, trade barriers)



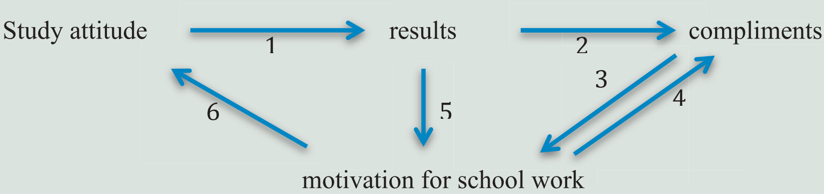
The eighth millennium goal talks about the way in which rich countries can help developing countries achieve the other millennium goals. For example, they can do this by giving more development aid, relieving debts and removing trade barriers.

Assignments for the Millennium Goals

The eight Millennium Goals are not separate goals, but together they form the basic conditions for dignified living conditions for everyone. The factors to which the Millennium Goals refer are closely related and influence each other. For each Millennium Goal, these factors are listed in brackets.

The assignments below are about the mutual relationships between these factors. You must always clarify the relationship between the two factors mentioned, by making use of other related factors. You have to choose them yourself from the factors listed in brackets for all Millennium Goals. Show the relationships using an arrow diagram and then explain each relationship (see example).

Example of an arrow diagram showing relationships with an explanation.



1: A good study attitude, such as doing homework faithfully and on time, gives good results.
2: If you have good results, you probably get compliments too.
3: Getting compliments motivates you to keep doing your schoolwork well.
4: Students who appear motivated are more likely to receive compliments.
5: Good results (grades) reward your work and motivate you to keep working.
6: Without being motivated for your schoolwork, you will not have or get a good study attitude.

Questions pre-test:

1. Indicate the relationship between **poverty** and the **position of women in society** with the help of other related factors in the form of an arrow diagram and then explain the relationships established.

2. Indicate the relationship between **poverty** and **the environment** with the help of other related factors in the form of an arrow diagram and then explain the relationships established.

3. Indicate the relationship between **participation in education** and the **number of children per woman** with the help of other related factors in the form of an arrow diagram and then explain the relationships established.

Questions post-test:

1. Indicate the relationship between **poverty** and **infant mortality** in society with the help of other related factors in the form of an arrow diagram and then explain the relationships established.

2. Indicate the relationship between **poverty** and **sustainable environment** with the help of other related factors in the form of an arrow diagram and then explain the relationships established.

3. Indicate the relationship between a **sustainable living environment** and the **fight against deadly diseases** with the help of other related factors in the form of an arrow diagram and then explain the relationships established.

3. Questionnaire about the research

1. Do you have the impression that the pre-test has been taken seriously by the students?

2. What mysteries did you have the students do?

3. Have you been actively present in the classroom with the mysteries?

4. Do you have the impression that the students were seriously working on the mysteries?

5. Did you have enough time for a good debriefing?

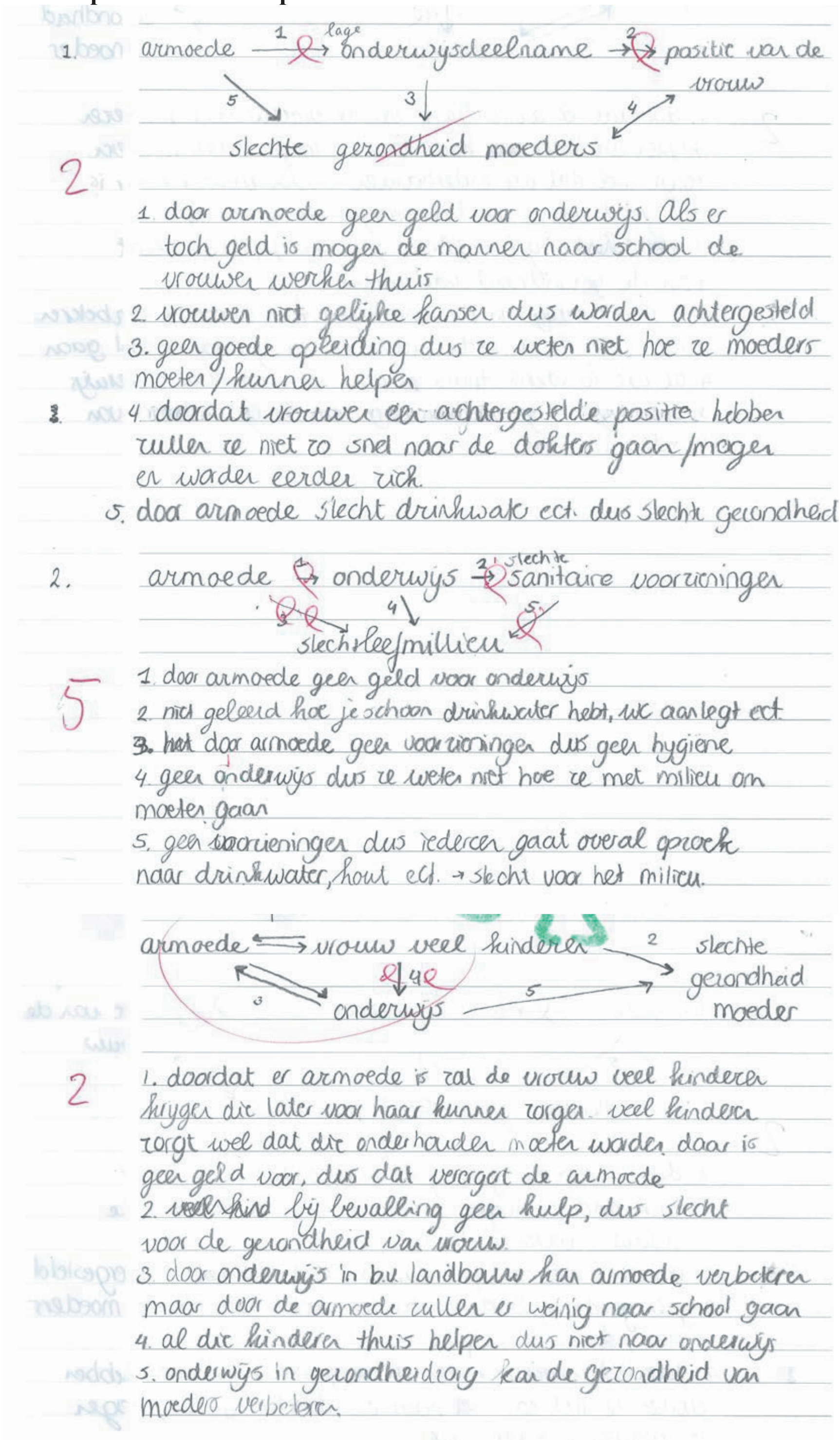
6. Did the debriefing go well?

7. Do you have the impression that the post-test was taken seriously by the students?

8. What other comments would you like to make as a result of the mysteries administered?

Think of:
- its subjects
- the level of difficulty
- the time required
- ………..

4. Example of a student's pre-test



1. armoede →(1, lage) onderwijsdeelname →(2) positie van de vrouw
(5) slechte (3) gezondheid (4) moeders

2

1. door armoede geen geld voor onderwijs. Als er toch geld is mogen de mannen naar school de vrouwen werken thuis
2. vrouwen niet gelijke kansen dus worden achtergesteld
3. geen goede opleiding dus ze weten niet hoe ze moeders moeten/kunnen helpen
4. doordat vrouwen een achtergestelde positie hebben zullen ze niet zo snel naar de dokter gaan/mogen en worden eerder ziek.
5. door armoede slecht drinkwater ect. dus slechte gezondheid

2. armoede → onderwijs →(2, slechte) sanitaire voorzieningen
(3) (4) (5) slecht leefmilieu

5

1. door armoede geen geld voor onderwijs
2. niet geleerd hoe je schoon drinkwater hebt, wc aanlegt ect
3. het door armoede geen voorzieningen dus geen hygiene
4. geen onderwijs dus ze weten niet hoe ze met milieu om moeten gaan
5. geen voorzieningen dus iedereen gaat overal opzoek naar drinkwater, hout ect. → slecht voor het milieu.

armoede ⇄ vrouw veel kinderen →(2) slechte gezondheid moeder
(3) (4) onderwijs →(5)

2

1. doordat er armoede is zal de vrouw veel kinderen krijgen die later voor haar kunnen zorgen. veel kinderen zorgt wel dat die onderhouden moeten worden daar is geen geld voor, dus dat vergroot de armoede
2. veel kind bij bevalling geen hulp, dus slecht voor de gezondheid van vrouw.
3. door onderwijs in b.v. landbouw kan armoede verbeteren maar door de armoede zullen er weinig naar school gaan
4. al die kinderen thuis helpen dus niet naar onderwijs
5. onderwijs in gezondheidzorg kan de gezondheid van moeders verbeteren.

5. Correction model pre- and post-test Millennium Development Goals

The aim is to count the number of correctly established relationships, per question and in total.

Students have indicated these relationships in a relationship diagram. Each relationship is indicated by an arrow.
By counting the number of arrows indicating a correct relationship it is possible to determine correctly established relationships, but:

- If an established relationship seems doubtful, one should read the explanation of that relationship. When convinced, the relationship counts.
- If a student explains a relationship and mentions an additional connection within that relationship (e.g. via an intermediary variable), it counts. All additional factors mentioned in the notes count!
- Only count relationships that are in the relationship chain between the two requested variables. So no lateral digressions that do not lead to the other variable nor relationships with factors that come out of the blue.
- Do not count relationships established by inserting development aid/international cooperation.
- Intermediate factor wrong but continuation right: subtract point.
- Do not count very vague factors (such as: 'poor facilities', sustainable things').
- No food leads to hunger: no extra point.

Each properly established relationship: 1 point.
Any permanently questionable or incorrect relationship: 0 points.
Determine the number of points per question and the total number by counting the number of correctly established relationships.

Specifics/examples:

1. Poverty and child mortality
- Relationships via education to infant mortality count, but do not count relationships via education back to poverty.
- If two factors are mentioned together between poverty and child mortality (e.g. hunger and no medicine), count this relationship chain twice: 4 points.
- Poverty leads to too little money: do not count as an extra relationship.
- High number of children means high infant mortality: do not count.

2. Poverty and (sustainable) environment
- Relations via deforestation, depletion, air pollution, no sanitary, no hygiene, slums, no money for environmental measures or no money for drinking water supplies: count.
- Two elements mentioned at the same time: relationship chain counts double.
- Deforestation leads to soil depletion: counts.
- When started with a specific realisation of a (non-) sustainable environment: 1 point.
- Relationship through education back to poverty: does not count.

3. Sustainable environment and combating deadly diseases
- Relationships via bad water, poor hygiene, all kinds of pollution: count.
- When started with a specific realisation of a (non-)sustainable environment: 1 point

Appendix B:

The Jakarta and the Rio Mystery: introduction, task and strips

1. Mystery: Rio de Janeiro

Introduction:

Almost every year, Brazil is in the news because of mudslides and landslides victims. Especially people who live in favelas ("slums") situated on steep slopes are victims. In April 2010, Rio de Janeiro was hit by extremely heavy rains and parts of slums were destroyed by landslides. The mayor of Rio announced a plan to demolish the most dangerous parts of the favelas and help relocate people to safer areas. The place where Fabio Pereira's house stands has also been labelled too dangerous to stay in and he has been asked to leave. However, Fabio refuses to leave.

Task:

Why does Fabio Pereira not want to leave his home, labelled as life-threatening?

Answer this question using the information on the slips provided by the teacher.

Present your answer in the form of a causal arrow diagram on a flap.

From 5 April to 8 April 2010, Rio was hit by the heaviest rainfall in fifty years.	The with water-saturated, unstable soil on steeper slopes began to slide and slip off.
On the 8th of April, the mayor of Rio signed a regulation to relocate eight favelas located in the most dangerous places, including Favela Morro dos Prazeres and Laboriaux, the upper part of Favela Rocinha.	Houses that fell due to landslides took low-lying houses down with them. More than 15.000 people were made homeless and more than 250 people died.
Laboriaux was created in 1982 when people were driven from another part of Rocinha because of "extremely high" danger of landslides.	People are being resettled in safe, more levelled areas on the outskirts of the city.
Travelling by public transport takes hours from the outskirts of the city to the centre of Rio.	Rapid relocation started in various favelas on 8 and 9 April.
The mayor says he wants to plant forests on the areas that have become vacant due to demolition.	Two people died in Laboriaux; thirty in Morro dos Prazeres.
Northeast Brazil is an area afflicted by drought, poor soil and unemployment.	Many people in Rocinha and Morro dos Prazeres are migrants from northeast Brazil.
The house of Fabio Pereira, father of four children, is located in Morro dos Prazeres, has three floors and was built by his father 30 years ago.	The favela Cidade de Deus, from the famous 2002 film City of God, started as a relocation project in the 1960s and is now one of Rio's most dangerous favelas.

Botafogo is a small favela on the south side of Rio, where no house has been destroyed by landslides, but where demolition has already started.	The government offers people who cooperate in relocation a one-year allowance of $ 100-300 per month, about a third of a monthly salary.
Land and property prices in Rio are among the highest in the world.	The 2014 FIFA World Cup final was held in Rio and the 2016 Olympic Games will be held there.
On several "extremely dangerous" sites, where favelas used to be cleared away by the government, residential towers with luxury apartments are now located.	Houses are marked with paint for demolition without consulting the leaders of the local community.
The eight favelas in need of demolishing are on slopes near Rio's Central Business District (CBD).	Fabio's parents migrated from northeast Brazil to Rio decades ago.
Rio de Janeiro is Brazil's main tourist city.	The worst houses located on the higher slopes were particularly hit by landslides.
The undamaged house of Fabio Pereira will have to be demolished as well because it is located in a very unsafe place, according to the government.	The government offers Fabio Pereira an apartment with one room and two bedrooms as a replacement home.
The school in Laboriaux has been closed for weeks. Children are roaming the streets of the neighbourhood, missing the daily free meal.	12.000 people live in Morro dos Prazeres and 4000 in Laboriaux.

2. Mystery: Jakarta

Introduction:

Every year, parts of Jakarta are hit by severe floods. Some years, even the CBD of Jakarta is hit by the floods. What is the cause of these floods?

Quote from a senior official in the Public Works Department responsible for managing water in Jakarta:
They build their homes on the river banks! Yes, of course they will continuously suffer from floods. Without those foolish people in the slums, we would not even have floods in Jakarta so we could focus on other important things.

Task:

Are Nani and the other inhabitants of the slums on the riversides in Jakarta the cause of the floods?

Answer this question using the information on the slips provided by the teacher.

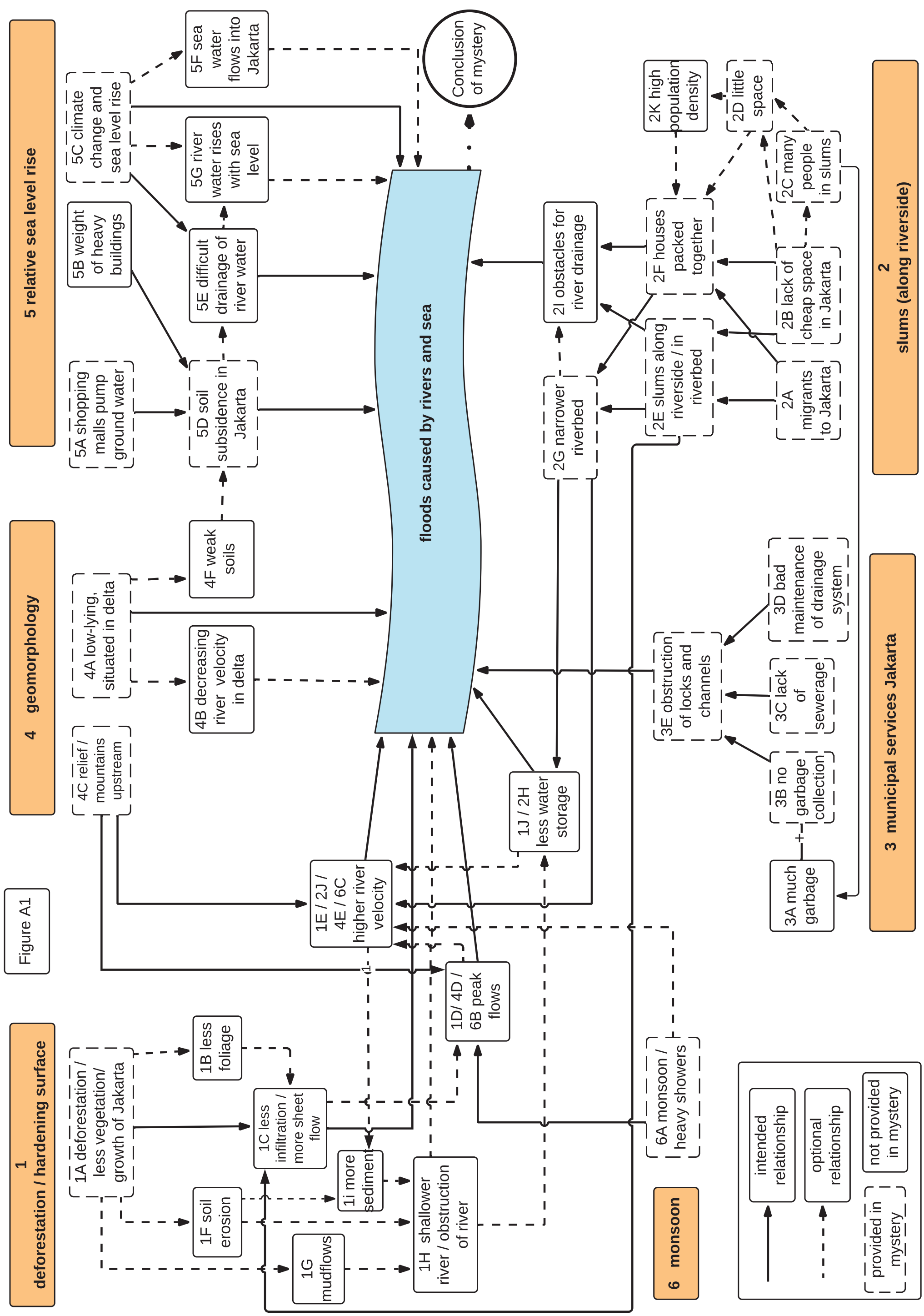
Present your answer in the form of a causal arrow diagram on a flap.

Jakarta is situated in a river delta on the coast of North Java.	Jakarta drops 10 cm per year compared to the sea level.
Every year, tens of thousands of migrants come to Jakarta in search of work and a better life.	This century, global warming and (absolute) sea level rise will continue.
In Jakarta, very cheap or free land can only be found in the worst places: close to the rivers.	Most of the more than 3 million inhabitants of the slums in Jakarta live illegally along the riversides, under the government.
Jakarta lacks a properly functioning garbage collection service.	Residents of slums do not have their own toilet or sewerage system.
Bantaran Kali floods three or four times a year.	New shopping malls, hotels and residential towers in downtown Jakarta are pumping up enormous amounts of groundwater.
Nani lives in Bantaran Kali and sells homemade fried rice and cups of coffee to her (regular) customers in the neighbourhood.	Nani has invested money in tiles for the ground floor.
Under the large residential towers, shopping centres, hotels and shopping malls of downtown Jakarta, the ground is sinking severely.	Like most of her neighbours, Nani does not want to move to cheap government rental apartments that are still to be built.

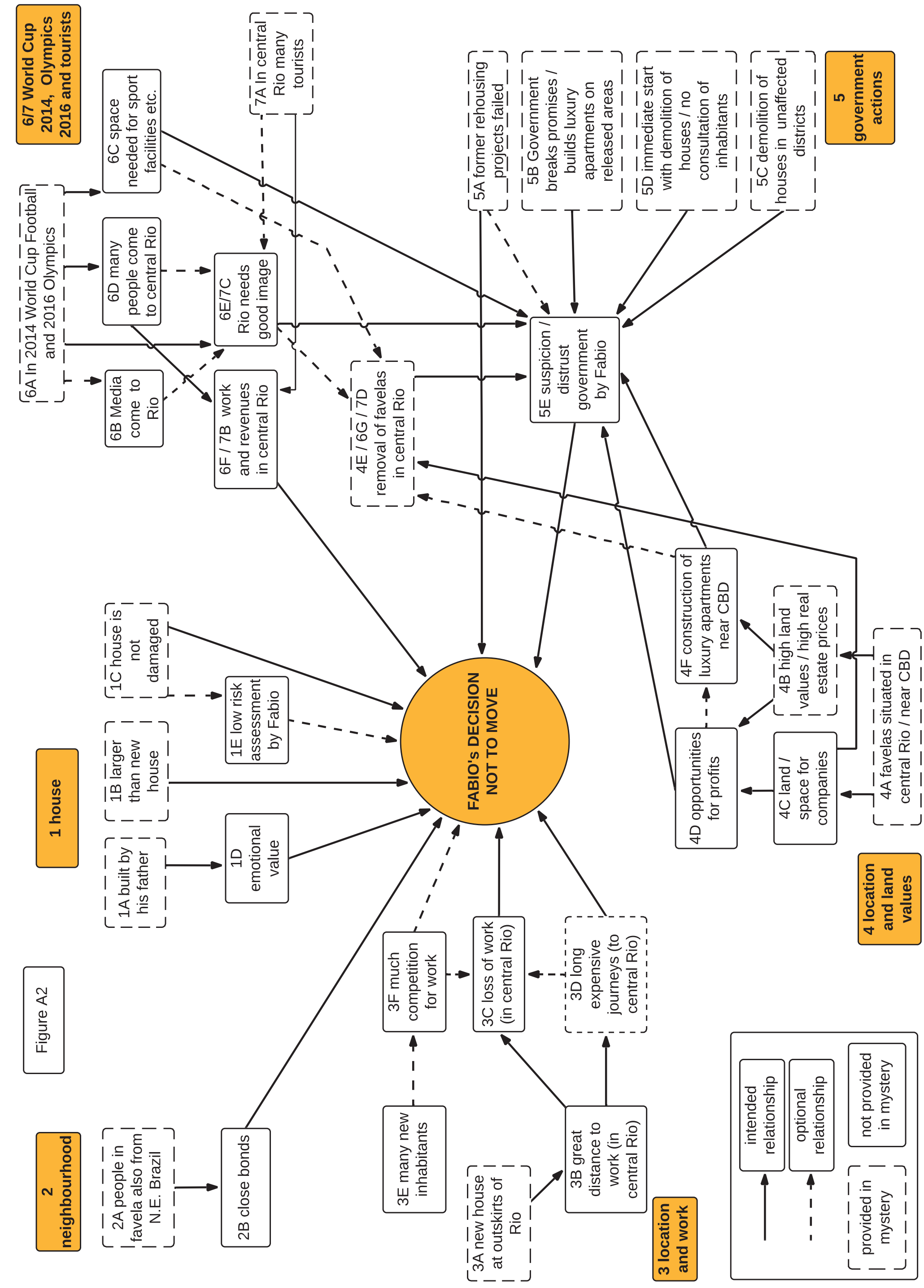
Nani has built a second floor on top of her house where she can keep valuables and food dry.	The slum Bantaran Kali is situated on the bank of the Ciliwung and is home to more than a thousand people.
The Ciliwung has its source in the mountains of Puncak, south of Jakarta, and carries excess water through the city towards the sea.	In the southern catchment area of the Ciliwung, a lot of forest has been cut down for new city districts, industrial areas and tea plantations.
In Bantaran Kali, the houses are densely built up to the water; the alleys are very narrow.	The width of the Ciliwung in Jakarta has decreased significantly in recent years: from about 50 meters to about 10 meters.
Officially, 30% of the surface of Jakarta should consist of foliage (parks, forests, etc.). This was only 9% in 2014.	In the rainy season, the monsoon falls in short, heavy showers from late afternoon to evening.
The system of drainage canals and rivers to drain excess water from Jakarta is poorly maintained.	The 2013 government plan provides for the dredging and widening of rivers and canals in Jakarta.
The government wants to offer slum dwellers along riversides alternative housing by building cheap rental apartments elsewhere in the city.	In the event of a flood, the ground floor of Nani's house is submerged by up to one meter of water. The water stinks and is black.
Various locks and drainage canals are often completely blocked in Jakarta.	According to the 2013 government plan, 70.000 slum residents living illegally on the river banks must leave.

Appendix C

Criterion maps of the Jakarta and Rio mystery



Criterion map of the Jakarta Mystery



Criterion map of the Rio Mystery

Appendix D

Group 24 as a supplementary case study (chapter 5)

Group 24

Group 24 was a group of three girls in their fourth year of higher general secondary education. At the start of their group work the strips were spread out orderly on the table. Student 2 proposed to use the same strategy as they did earlier to solve a mystery:

22	S2	Shall we just do the same as we did last time? We then first looked at issues that deal with the matter, you know.
23	S3	Yes.
24	S1	And yes, several…
25	S2	You use two groups, one of which has nothing to do with it, you can just throw that away. And then the other group that is useful, in politics, social dimension etc., you know..

So they started reading the strips one by one, discussing their usefulness for answering the mystery question. Useful strips were placed on the left side of the table, the others on the right. Most strips were read and discussed together. Halfway the selection, student 3 decided to use an atlas to look for the location of Rio. Student 1 and 2 continued the selection process, discussing together. When they had finished they put 11 strips, most of which were unknown to student 3, back in the envelope. Then, student 3 joined the group discussion again which continued with categorising and making sets of the remaining 17 strips using the political, economic and social dimension. The "geographical" dimension was added later. Each of the 17 strips was read aloud and categorised after common decision. After the resulting sets were made, they continued their discussion on how the sets and strips should be connected. Student 2 made a proposal to summarize each set in one sentence (line 187), which was not clearly understood by the others. Student 1 explained it, but proposed an alternative strategy focused on finding causal relationships between individual strips from different sets (line 192), which was accepted.

187	S2	But we also, the previous time we also used groups, you know, you create groups. And then you kind of create this and that together. You come up with a sort of summary in one sentence.
188	S1	Yes.
189	S2	You get it?
190	S3	No.
191	S2	No, and I really don't know how I am supposed to explain this.
192	S1	No, but you just look at the main focus of all these notes together. And write it down. Or you can try to find out whether you can connect one of those, from these and those. That way you'll be able to see a possible cause-effect.

This webbing strategy of student 1 was taken over by the other group members and a first causal relationship was established by student 3 (line 212), but not discussed with the group.

211	S1	These ones could go together. Like this, let me think.
212	S3	I reckon these two are also somehow related because look here is the, yes that's when he got his apartment with a room and two bedrooms as a replacement, but if they were to collaborate they would get a hundred to three hundred dollars a month and that is a third of a monthly salary. So why would you leave? If you get less.

As they continued, connections based on association (line 222- 226) became dominant and no new causal relationships were established or discussed.

222	1	Ja misschien wel, of niet. En hier dit, dit gaat hier over de sloop.
223	2	Maar dit heeft toch helemaal niks met bossen te maken.
224	1	Nee maar deze misschien hier ook bij of niet?
225	2	Omdat het over zijn huis gaat.
226	1	Ja nee dit is meer ff kijken hoor. Dit is herhuisvesting.

222	S1	Yes, could be, or not. And this here, this is not about the demolition.
223	S2	But this doesn't have anything to do with the forests, right?
224	S1	No, but these should be added here as well, or not?
225	S2	Because it has to do with his house.
226	S1	Yes no, this has got to do with, let me see, This is relocation.

As a result, 7 new sets were made, mainly based on mystery topics or on chronology (landslide, his house, demolition, rehousing). Then the strategy of student 2 she mentioned in line 187 reappeared, in order to construct the causal diagram (line 257). Each set was summarised in one sentence which was written down on a piece of paper to be used in the causal diagram.

256	S1	Okay, what's next.
257	S2	Let's say this must be summarised into one sentence. And this. And then you sort of have, you have seven sentences which could eventually help create a causal diagram. With cause and effects.

The group continued their discussion by trying to capture each set in one sentence and consequently connect them. This is illustrated by some examples of those summarising sentences and how they are connected (lines 298, 305 and 336). Intermediate discussions on how to formulate the sentences are deleted. Although the connections were expressed by the students as causalities, only the conclusion contained real causalities (line 363); the other connections only organised a chain of events.

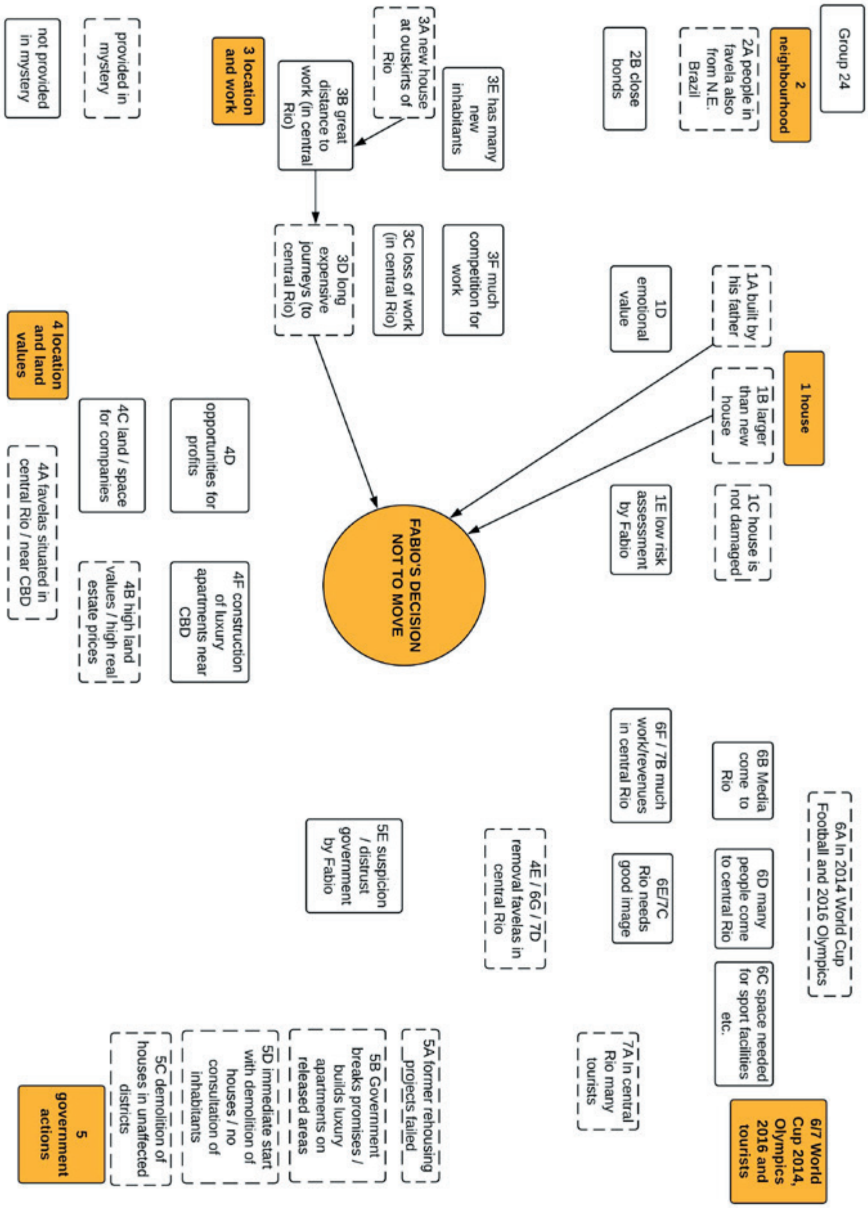
298	S2	But should you put: as a result of poor housing, houses will be demolished, or: as a result of poor housing, houses will be demolished instead of destroyed.
305	S1	Okay. You could then for example, erm, this results in: the government will offer residents new housing.
336	S1	Perhaps it could be that you, for instance, do the housing and then that one. Say, this results in: new housing will take place, erm, this allows new apartments with two bedrooms and a compensation of a hundred to three hundred dollars per month.
363	S1	Erm, and then, you can draw the conclusion, then after that, for instance, this eh causes too little space to live and too little money for Fabio and then another arrow, that's why he doesn't want to leave.

After having summarised their remaining four 'geographical' sets, they only found reasons for Fabio to move! What to do? Student 3 proposed to reconsider the strips in the reject pile (a reworking) and the others agreed. As a result of the selection process of students 1 and 2 at the start, student 3 was not familiar with every strip on the pile yet, so she read them for herself. She valued each strip one by one on its usefulness as a reason for Fabio to stay. And she discovered a useful one (line 491) which gave new insights into the mystery question (line 492-501).

491	S3	Oh then he is far away from…, He will be placed on the outskirts of the city.
492	S2	Maybe he doesn't want to live that far from Rio.
493	S1+3	Yes.
494	S1	Because of his, you know, because of his selling, I don't know if he sells something.

495	S2	Maybe he doesn't want to live far from Rio because his parents immigrated to Rio decades ago.
496	S1	Yes, or because of his work.
		(…)
501	S3	Yes, true, so I also think just because of his work. Because uhm, maybe there is not so much work over there and he will need to…

Again this last excerpt of the discussion shows some webbing, triggered by the discovery of a useful strip. However, the webbing of individual strips remained limited, because it was hindered by the strategy of making connections between summaries of sets mainly based on association. The reworking through reconsideration of the reject pile illustrates that the removal of a large part of the strips at the start also hindered understanding of the mystery. Last but not least, the focus on separate strips in order to explain Fabio's decision, recognizable throughout the group work, hindered insight into the geography and society of Rio, necessary for a better understanding of Fabio's decision. As a result, this group also has a relatively low total proposition score (9) and an output on the multistructural level, as can be seen in their ('standardised') concept map.



Standardised concept map of group 24.