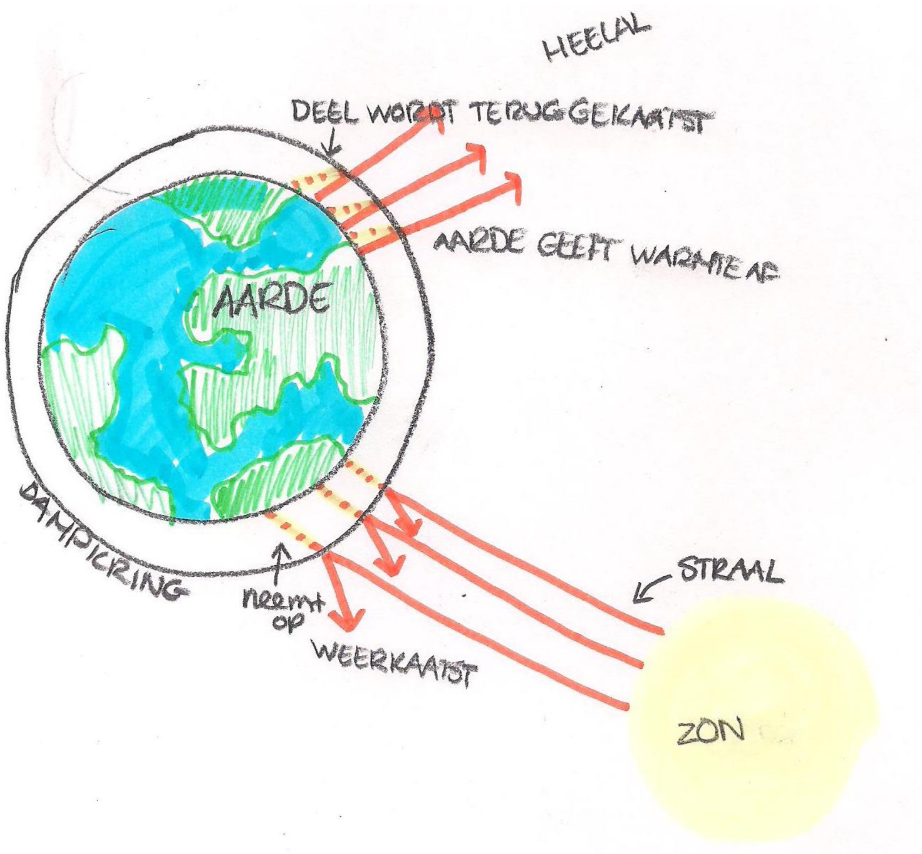
Universiteit Utrecht



To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

Master thesis Franka Kerklaan

Student number: 3384179

Master: Science Education and Communication

Utrecht University

Supervisor: Prof. Dr. W.R. van Joolingen

Second supervisor: Dr. N. P. G. Rutten

Date: 24-08-2015

Abstract

This study investigates the extent to which students' drawings of a physical system, the Energy of the Earth and its atmosphere, can be interpreted to provide insight into their understanding and mental models of the system. The study builds on research by Van Joolingen and Kenbeek (2013) and used the same scoring system and factor analysis as an extra method to get to know possible mental models or mindsets of students. The scoring system was tested on data from Tubingen and all datasets were compared. The information from the drawings was complemented with interviews and videorecordings of a selection of students. These extra methods did not lead to more information, leading to the conclusion that all that is to be gained from drawings is the same as the oral explanation of students. The factor analysis did reveal different factors in all datasets. Implications for further research are discussed.

To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

Content

1. Introduction

A key part of science learning is to facilitate students reflecting on their existing mental models about scientific concepts, to revise these models and to build new ones (Greca & Moreira, 2000). Students need to learn about basic concepts and their relations but also about the way these concepts are conceptualized and used in solving scientific problems to be able to build their own mental models. Science education involves teaching and learning about the way scientific knowledge is constructed and extended. Gilbert (2005) stated that another goal of an educational experience for students is to develop an understanding of some principle or concept. Students have to be able to apply that information to resolve an extended range of problems in a variety of situations (Gilbert, 2005). How much and in what way an individual understands can be seen as the mental model he or she has. Mental models are conceptual organizations of information in memory. They are internalized, organized knowledge structures that are used to solve problems. Mental models capture a type of memory that instructors want students to build (Gilbert, 2005). Students are constantly revising their mental models based on new knowledge, ideas, concepts and experiences. Also, students' mental models are personal, idiosyncratic and often unstable (Greca & Moreira, 2000).

External representations can be used to enable students to express their mental models (Reiss & Tunnicliffe, 2001; Gobert&Buckley, 2000). Representations can lead to a deeper understanding (Ainsworth, 1999). External representations include action, speech, written description, and other material depictions (Gobert & Buckley, 2000). An example of such an expressed model is a drawing. Letting students draw, aids immediate and longer-term comprehension, because it provokes careful observation and extensive mental manipulation of the presented content. In the task of drawing, participants are asked to search for and make explicit key conceptually relevant information (Mason et al, 2013). However, if we want to make sure students learn from drawing, we need to find a way to understand what students actually draw in order to get a view on their learning. To be able to probe students' understanding according to their drawings, a reliable analysis tool is needed. From an earlier study, in the domain of "Newton's cradle", Mason and colleagues concluded that richer and more accurate drawings were associated with better comprehension of an animation (Mason et al, 2013).

Students and drawings

Ainsworth et al. (2011) stated that drawing is a way to better understand complex science systems. It can make students more engaged. Students learn to represent and reason in science. Thus, making students draw can be used as a learning strategy and gives the learners an opportunity to communicate better (Ainsworth et al, 2011).

Self-generated drawing is an example of a constructive learning activity (Chi, 2009); the benefits of these kind of activities are attributed to their requirement of learners to transform the original learning material into a different output. The drawing strategy appears to be a useful strategy in fostering mental model building. In tests, a drawing strategy can help learners to construct a representation that goes beyond the text (Leopold & Leutner, 2012). Also, drawing might be seen as an externalization of a concept or idea (Brooks, 2009). Van Meter and Garner (2005) present the 'Generative theory of drawing construction'. This theory states that students who are asked to draw a picture while reading a text have to engage in three cognitive processes. First, the students have to select the relevant information from the presented text. Second, they have to build up an internal verbal model of the text information by organizing the selected information. And finally, the students have to construct an internal nonverbal representation of the text information. This representation is connected with the verbal representation and with relevant prior knowledge (Van Meter & Garner, 2005). Subjects might enjoy the drawing and take a certain care in the production of it. For researchers, another advantage is a rich mass of data, which can be obtained at comparative ease (Reis & Tunnicliffe, 2001).

Interpretation of drawings

For understanding students' mental models, it is necessary that they are expressed externally. However, the expressed model and the student's mental model may differ (Reis & Tunnicliffe, 2001). Reis and Tunnicliffe studied the drawings by students who had been asked to draw their own intestines. The researchers found that students, at least to some extent, had drawn their assumption of what the researchers wanted to see. To overcome this problem and to get a more comprehensive view on a student's mental model, a combination of different methods is needed. Different methodologies reveal different things about understanding: a multi-dimensional complexity (Reis & Tunnicliffe, 2001). Using interviews next to analyzing drawings can be helpful. But in the case of interviews, it is uncertain whether the construction we find represents precompiled theories which are stored in long-term memory or whether they are constructed by the children on the spot under the influence of the questions (Vosniadou & Brewer, 1992).

Students' understanding and mental models of the Energy of the Earth and its Atmosphere

The scientific model used in this study is the model of the Energy of the Earth and its Atmosphere. This model is also known as the natural greenhouse effect. The model of current scientific consensus about the greenhouse effect is that a defined layer of greenhouse gases somewhere in the atmosphere impacts changes in temperature by holding or trapping in radiation or heat from the sun (McNeill & Vaughn, 2012). Several studies have been conducted about students' understanding and mental models of the greenhouse effect and climate change. One of the more elaborate studies was performed by Shepardson et al (2011), who found five mental models about the greenhouse effect among 12-13 year olds. The first model consists of the mistake in the word 'greenhouse'; students who held this mental model drew the earth as a greenhouse. The second model contains greenhouse gases which cause ozone depletion or formation, which either allows more of the sun's rays to reach the Earth or causes the sun's rays to be 'trapped' or 'bounced' back toward Earth. The third model has simply greenhouse gases in the atmosphere, the heating mechanism is not present in this model. Model 4 consists of greenhouse gases which 'trap' the sun's rays, heating the earth and in model 5, sun's rays are 'bounced' or reflected back and forth between the Earth's surface and greenhouse gases, heating the Earth (Shepardson et al, 2011). Many students only see solar rays from the sun as involved in the greenhouse effect; they may lack the concept of terrestrial radiation (Koulaidis & Christidou, 1999). Students may also fail to understand the Earth's energy balance as a whole. These mental models show that the12-13 year old students lack a clear understanding of the greenhouse effect (Shepardson et al, 2011). In another study by Shepardson et al (2009) with older students (grades 9 to 12), students indicated that they think global warming is caused by greenhouse gases and air pollution in general. This is in line with a study by Jakobsson and colleagues, showing that students aged 14-15, struggle with the difference between the natural and anthropogenic greenhouse effect and have difficulties understanding what a greenhouse gas is (Jakobsson et al, 2009; Andersson&Wallin, 2000). This is also demonstrated in the study of McNeill and Vaughn (2012) in which some students described the barrier of greenhouse gases bouncing back heat from the earth. These students believe climate change is caused by heat produced by industry that is trapped and bounced back toward the surface of the earth (McNeill&Vaughn, 2012). Another indication was, correctly, that carbon dioxide or greenhouse gases form a layer in the atmosphere that traps and reflects the sun's energy (Shepardson et al, 2009). According to the results presented in the previous section, students have difficulties accepting scientific models and explanations, and that common sense ideas tend to dominate students' reasoning even after teaching (Driver, 1983).

It should be noted that some of the results about the knowledge and misconceptions of students emerge as artifacts of the methods used. For example interviews or questionnaires are used, in which students have to respond to questions that appear out of context, and where they might feel an obligation to come up with some kind of answer (Jakobsson et al, 2009).

This study will build on an earlier study by Van Joolingen and Kenbeek (2013) about the understanding of students' drawings of complex systems. Van Joolingen and Kenbeek analyzed drawings of students about the greenhouse effect by scoring objects, processes and annotations. The aim was to find out to what extent students' drawings of such a system can provide insight into students' understanding of the system. By using factor analysis, two factors representing a specific view on the system, were found. These factors are the 'heat vs. light': a view on how energy is transported from the Sun to the Earth, and the 'function of the atmosphere' factor representing the way the role of the atmosphere is conceptualized. Almost all students in this study were able to extract and represent the relevant objects, but less than half of the relevant processes were represented in the students' drawing summaries. The mental model on the nature of the energy that flows from the Sun to the Earth may determine a student's position along the 'heat vs. light dimension'.

The objective of this study is to examine the drawings of a physical system and the mental models of students using different methods. By doing so, we hope to confirm the different viewpoints on the system, found by the factor analysis by Van Joolingen and Kenbeek. In order to explore these subjects, the following research questions are designed:

1. To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere, be interpreted to provide insight into their understanding and mental models of the system?
2. To what extent can the scoring method used by Van Joolingen and Kenbeek (2013) be applied to drawings generated by students in a different context?
3. Can we provide evidence that drawings, scored by this scoring method, provide an actual representation of students' conceptual understanding?

To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

2. Method

In this study three data sets were used. Next to the original dataset from Van Joolingen and Kenbeek (2013) (from now on labeled dataset A), consisting of 66 drawings, we used 35 drawings collected from German students (from now on named dataset B) as well as 72 drawings we collected specifically for this study (from now on named dataset C). From dataset C, ten students were selected to be videorecorded during the drawing task and ten others were selected to participate in a post-hoc semi-structured interview.

2.1 Participants

The three datasets consist of drawing summaries made about a text about the Energy of the Earth. Dataset A consists of 66 drawing summaries by ninth grade VWO students from the Netherlands, all were 14 or 15 years old. Dataset B consists of 35 drawing summaries by ninth grade pre-university students from Tubingen, Germany made in 2014 for a previous study. 77 Dutch Havo/VWO students (35 girls, 42 boys) participated in the current study, leading to dataset C; their age ranged from 14 to 17. No specific sampling method is used, since the number of participants is very high (77). Three classes were involved, 1 ninth grade Havo/VWO (26 students), 1 ninth grade bilingual class (26 students) and 1 tenth grade VWO (25 students). 7 of the students who were filmed were from the Havo/VWO class and 3 from the bilingual class, due to logistical reasons. Ten students from the tenth grade VWO class were randomly selected for the post-hoc interview. The participants in dataset C are from two secondary schools, one comprehensive school nearby Utrecht and one gymnasium nearby The Hague.

2.2 Analysis

All three datasets are analyzed with the coding rubric (*Table 1)* by Van Joolingen and Kenbeek (2013). The datasets are compared and an exploratory factor analysis is performed.
The reliability of the coding rubric is validated by using it on data of 35 ninth grade students from a German high school who made drawings about the same topic. The German and Dutch texts were compared and minor changes have been made to the Dutch text for the collection of dataset C, because the text might have had influence on the scores of the drawings. In the Appendix six texts, used for the three datasets and an English translation can be found: the original Dutch text by van Joolingen and Kenbeek (2013), the German text, used for dataset B, the text in Dutch used for dataset C and an English translation of all last texts.

The rubric consists of categories that correspond to the pieces of information in the assignment text, focusing on objects and processes represented by the students. The rubric is a coding scheme which codes in three categories: objects, processes and annotations. In a complete drawing summary of the topic, all objects and processes would be represented. Drawing summaries will be scored for whether they include representations of each of these objects and processes. Redundancy of information represented in the drawing summaries will be controlled by allowing each code to be used only once in each drawing summary.

Two processes which are present in the rubric are not written in the text: the processes of the Atmosphere radiating heat in the direction of the Earth (PDAE) or in the direction of the universe (PDAU). These processes can be inferred from the text because of the presentation of the Earth as an object with a heat capacity.

This rubric was verified first by analyzing dataset B. When the method had proved to be reliable, dataset C was analyzed. Part of the data (9 drawing summaries) was scored by a second rater; inter-rater reliability was good, Cohen's κ was 0.67. The information from the video recording and structured interviews was transcribed and analyzed for the same categories: objects, processes and annotations as used in the drawing-analysis.

To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

Table 1: Description of the codes used for the analysis. The first column shows the categories of the labels, in the middle column the label codes are displayed and in the third column a description of each label is given. The labels are abbreviations: e.g. PAAS means the process of the Absorption of heat from the Sun by the Atmosphere.

Category	Code	Description
Object	OS	Sun
	OE	Earth
	OA	Atmosphere
Process	PAE	The Earth absorbs (part of) the heat from the Sun (absorption)
	PDE	The Earth radiates heat
	PFE	The Earth reflects (part of) the sunlight (reflection)
	PFAS	The Atmosphere reflects (part of) the sunlight (reflection)
	PFAE	The Atmosphere reflects (part of) the sunlight that was reflected by the Earth
	PAAE	The Atmosphere absorbs (part of) the heat radiated by the Earth (absorption)
	PAAS	The Atmosphere absorbs (part of) the heat from the Sun (absorption)
	PDAU	The Atmosphere radiates heat in the direction of the universe
	PDAE	The Atmosphere radiates heat in the direction of the Earth
Annotations	AN	Naming
	AX	Explanation
	AL	Legend

2.3 Materials and procedure

For all datasets, the material that is used consists of a short science text on the topic of 'Energy of the Earth' (see Appendix I). This text has the same content as Van Joolingen and Kenbeek (2013) used, with minor revisions, after comparing the texts used for dataset A and B. The drawings in dataset B were made using a different text on the same topic (see Appendix I) . The students also received an assignment text instructing them to create a drawing summary to represent what they understood of the science text. This text instructed them to focus on content rather than on aesthetic aspects. Students read the assignment and science text before they started creating their drawing summaries.

Both texts remained available while the students were drawing. The students were explicitly allowed to use clarifying annotations in their drawing summaries. Both the science text and the assignment were printed on a sheet of A4 paper, and the backside of the sheet was reserved for the students' drawing. Students brought their own drawing materials (pens, pencils, etc.). The study was carried out in a regular classroom during a regular lesson. At the start of the lesson, the teacher handed out the materials and the students worked on the task for approximately ten minutes.

10 randomly selected students of dataset C were video recorded individually and asked to think aloud during the drawing task. The transcriptions of the videos were analyzed for the same processes, annotations and objects as the drawings.
Another 10 students were selected for a post-hoc interview. The question asked was: "Can you explain what you drew for me?" The goal was to make sure we analyzed the same out of a drawing as the student had in his or her mind. The answers to this question were transcribed and analyzed for the same categories: objects, processes and annotations as the drawings. A Cohen's κ was calculated to test the difference in scores between video recorded, transcribed and drawn data. Cohen's κ was used to calculate the inter-instrument reliability.

A factor analysis was carried out in order to check for patterns in the labels assigned and to see if they match the factors found by Van Joolingen and Kenbeek (2013). Because Van Joolingen and Kenbeek found three factors we also used three for factor analysis. Principal Axis Factoring (PAF) was used as extraction method (For an explanation of this procedure, see Costello & Osborne, 2005). No rotation method was used on the data.

To construct an idea of how much and what prior knowledge students use in their drawing summaries, 3 drawing summaries were selected and analyzed thoroughly. These drawing summaries were selected based on the scores according to the rubric, with 1 drawing summary of each participant-group from dataset C (drawings, videorecording and interview).

To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

3. Results

3.1 Comparison of dataset A, B and C

First the texts of dataset A and B were compared to change the text for dataset C if necessary. The German and Dutch texts differ in length: 138 words in Dutch versus 196 words in German. Another difference is that 'Greenhouse effect' is mentioned in the German text, while it is not in the Dutch version. The processes PAAS (The Atmosphere absorbs (part of) the heat from the Sun), PDAE (The Atmosphere radiates heat in the direction of the Earth) and PDAU (The Atmosphere radiates heat in the direction of the universe) are not mentioned in German.

	DATASET A	DATASET B	DATASET C	ANOVA	ANOVA: F (2,177)
PAE	72	66	87	p=.019	F=4.042
PAAE	38	66	30	p=.002	F=6.409
PAAS	62	26	44	p=.002	F=6.637
PFE	46	54	56	p=.445	F=0.814
PFAE	16	54	30	p<.001	F=8.732
PFAS	49	74	79	p<.001	F=8.896
PDE	63	80	77	P=.104	F=2.288
PDAE	12	40	0	p<.001	F=22.054
PDAU	29	29	3	p<.001	F=11.831

Table 2: in this table, the relative frequencies per category are given for each dataset. Also the results from the Anova test are shown.

As is to be seen in *Table 1* the results from the Anova test show that almost all categories score significantly different in the datasets. Tukey HSD tests were conducted on all possible pairwise contrasts. The following processes were found to be significantly different (p<.05): PDAU between datasets A and C (p=.03), and PDAE also between datasets A and C (p<.001).
In all of the categories, except for PAAS, the drawings in dataset B scored higher than the drawing summaries in dataset A.

Dataset A

Fifty-eight students (85%) used annotations to 'name' (AN) parts of their drawing summaries; 35 (51%) used annotations to 'explain' (AX) parts of their drawing summaries, and six (9%) used a 'legend' (AL) in their drawing summary. On average, students represented 2.94 of the 3 objects (SD = 0.24), 3.87 of the 9 processes (SD = 1.22), and used 1.46 of the 3 annotation types (SD = 0.76) (Van Joolingen & Kenbeek, 2013).

Dataset B

Twenty-eight students (80 percent) used annotations to 'name' (AN) parts of their drawings, 25 students (71%) used annotations to 'explain'(AX) parts of their drawings, and 2 students (6 percent) used a 'legend' (AL) in their drawings. On average, students represented 2.97 of the 3 objects (SD=0.17), 4.88 of the 9 processes (SD=1.330), and used 1.57 of the 3 annotation types (SD =0.65).

Dataset C

Regarding the objects category, all students (77) represented the earth and the atmosphere, almost all students (75) represented the sun. The process of 'absorption of heat by the earth (PAE) was represented most frequently (67 students) and the process of 'the atmosphere radiating heat to the earth' (PDAE) was represented least frequent, none of the students represented this process.

63 students (82%) used annotations to explain their drawings. 53% (41) used an explanation to explain more about their drawing. And 7 (9%) used a legend (AL) next to their drawing. On average, students represented 2.97 of the 3 objects (SD=0.16), used 1.44 of the 3 annotation types (SD=0.76) and drew 4.06 of 9 processes (SD=1.34).

3.2 Exploratory factor analysis

The results in the previous section provide a general impression on what students on average represented in their drawing summaries. An exploratory factor analysis was performed to get a more in-depth insight of what the drawing summaries contained. The scores on the three objects were excluded, because they were present in almost all drawing summaries. The factor analysis was already done for Dataset A.

For dataset C, the processes PDAE and PDAU were also excluded, because they scored too low in this dataset.

Table 3: Item loads on the three factors resulting from the factor analysis. Results for all datasets are presented. For each item, the highest item load is marked with an asterisk.

Item name	*Dataset A*			*Dataset B*			*Dataset C*		
	Factor			Factor			Factor		
	1	2	3	1	2	3	1	2	3
PFE	-0.689*	0.186		0.805*	0.124		-0.706*	0.307	0.343
PAAE	0.619*			-.275*			0.218	0.693*	0.289
PDE	0.569*	-0.116	0.114	-0.242*	0.131	0.112	-0.708*	0.219	0.192
PFAE	-0.454*		-0.133	0.789*		0.110	-0.651*	0.333	0.421
PDAE	0.412*	0.121				0.999*	-	-	-
PAAS	-0.145	0.904*		-0.264	-0.121	0.288*	0.350	0.275	0.391*
PDAU	0.490	0.526*	-0.209	-0.184	-0.430*		-	-	-
PAE		-0.386*	0.115	-0.154		0.203*	0.555*	0.087	0.092
PFAS		-0.179*	0.117		0.962*		0.353*	0.104	0.275
AX			0.999*	-0.402	0.533*	0.285	0.015	0.595*	-0.588
AN		0.250	0.347*		-0.715	-0.192*	-0.95	0.662*	-0.425

The results of the factor analysis show a different pattern in all three datasets. The first factor shows some resemblance between dataset A and B, for the categories PFE, PDE, PFAE, but not for PDAE. Dataset C shows a different pattern, which might also be caused by the absence of the processes PDAE and PDAU in the drawings and therefor in the scores.

3.3 Results from video and interview

To answer sub-question 3, the question if we can provide evidence that drawings, scored by this scoring method, provide an actual representation of students' conceptual understanding, t-tests and Cohen's kappa were used.

For the group that was filmed during their drawing process, the t-test ($p= 0.05$) revealed that there was no significant difference in the scores between the drawings and the videos.

Comparing the results of the analysis of the explanation of students who joined in an interview with the analysis of their drawings with an independent sample t-test ($p=.05$) revealed that only PDE as a process was significantly different: $t(18)=.334$, $p=.024$ ($SD=.203$).
The process of the Earth radiating heat was scored in all explanations, while it was only scored in eight out of ten drawings. This result might suggest that students reveal more information while explaining their thoughts with speech, but the sample size is too small to conclude this.

Cohen's kappa was used as an inter-instrument tool to compare the results from the drawing summaries and the spoken text of the students. For the 'video group', kappa was .734 and for the 'interview group' .625. Seeing the significance and the value of kappa leads to the conclusion that there is a good inter-instrument reliability.

3.4 Detailed description of three drawing summaries

To obtain an idea of how much and what knowledge students use in their drawing summaries, apart from the information gained from the rubric-analysis, these three drawing summaries will be described in terms of their content, the representation of prior knowledge, and of the representational formalisms that were used in these drawing summaries.

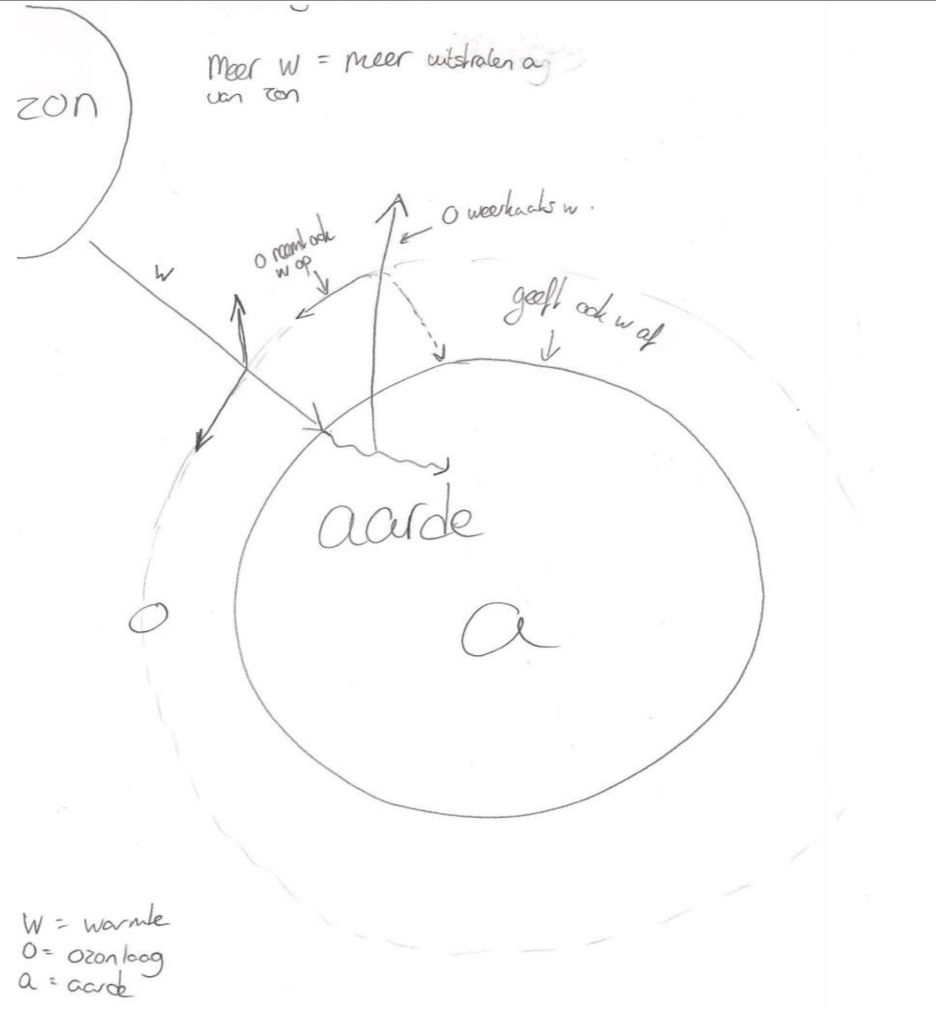


Figure 3. This drawing represents the one with the highest score (13) of the total number of students from dataset C.

To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

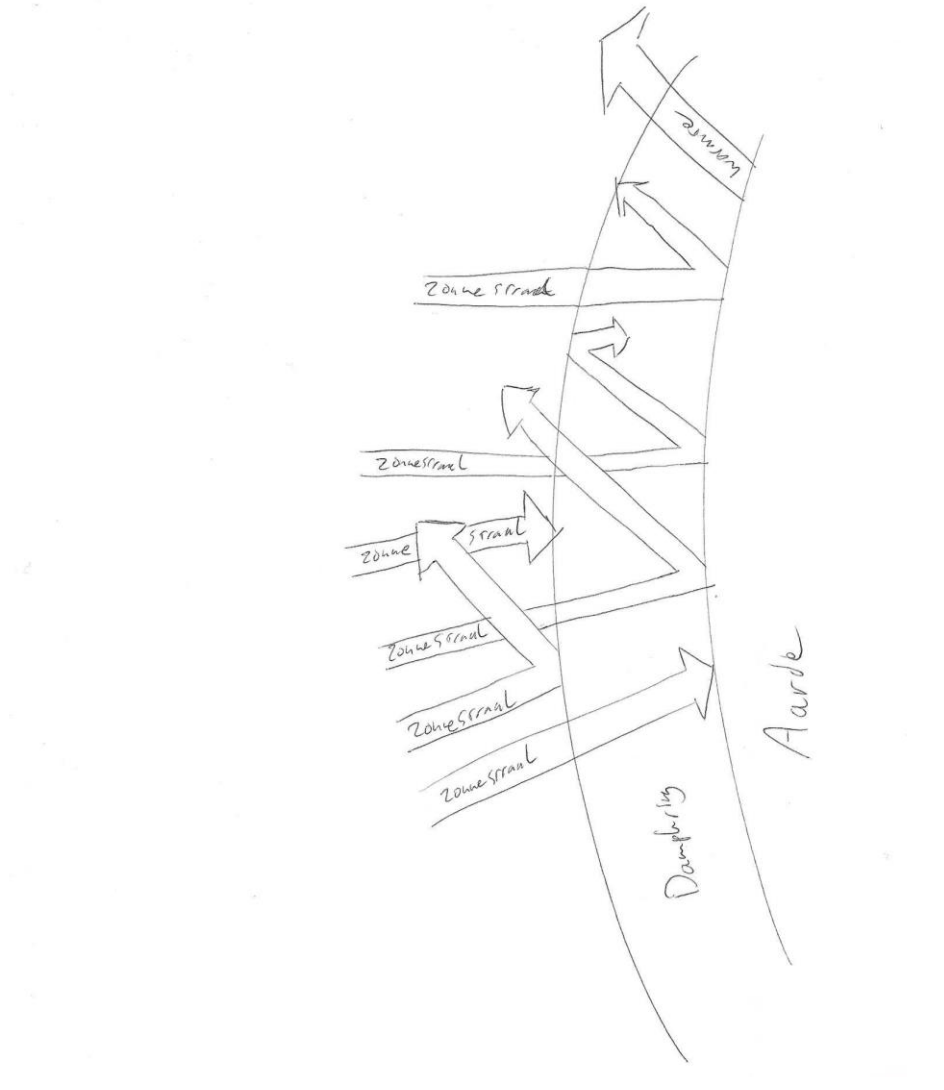


Figure 2. This drawing represents the one with the highest score(8) out of the group of students who was interviewed about the drawing.

Figure 3. This drawing represents the one with the highest score (9) of the students who were videorecorded and asked to think aloud during the drawing process.

In *Figure 1*, the drawing with the highest overall score is displayed. This drawing was one of the few with a legend. There are annotations and an explanation. The sun is represented on the top left corner and the earth in the center of the drawing.

A circle in a lighter color is drawn around the Earth representing the Atmosphere. Small arrows are drawn between the Sun and the Earth, and between the Earth and its atmosphere. These arrows represent some kind of process, presumably the transport of energy, since the 'w' in the drawing represents heat (warmte in Dutch).

In *Figure 2*, the drawing with the highest score of the interviewed students is displayed. When comparing this drawing with the previous one, the absence of the Sun in this drawing is one of the first things to be noticed. But its rays are visible in the form of arrows, which is also written within the arrows ('zonnestraal' in Dutch). The Earth is represented on the right side of the drawing. A semi-circle is drawn around the Earth representing the Atmosphere. The arrows represent some kind of process, presumably the transport of energy.

In *Figure 3*, the drawing with the highest score among videorecorded students is displayed. The sun is represented on the left side of the drawing and the earth in the center. A circle drawn around the Earth represents the Atmosphere. One line is drawn between the Sun and the Earth and two shorter ones from respectively the Earth and the Atmosphere, representing reflection (as is written in the drawing: 'weerkaatsing'). This line presumably represents some kind of process, transport of energy is most likely.

Noticeable in all these drawings is the difference in the representation of sun rays. Two of the drawings show arrows, one shows wavelike shapes. In the first drawing straight arrows and wavelike arrows are combined. This difference in depiction might indicate different mental models about convection or radiation of heat.

4. Discussion

This study started with the question to what extent the rubric used by Van Joolingen and Kenbeek (2013) could be applied to drawings generated by students in a different context. The observation in this study was that the scoring method was useful for scoring of drawing summaries, because it was possible to score the results from dataset B using the rubric. There were some differences in the occurrences of the drawn processes between dataset A and B. PAAS (The Atmosphere absorbs (part of) the heat from the Sun) was the only process that was represented more in the drawings in dataset A than in dataset B. This can be explained by comparing the two texts. PAAS is the only process that is not present in the German text (used for dataset B) and is in the Dutch text (used for dataset A). PAAE (The Atmosphere absorbs (part of) the heat radiated by the Earth) and PFAE (The Atmosphere reflects (part of) the sunlight that was reflected by the Earth) are categories that were not mentioned in the Dutch text but were in the German text. PAAE was mentioned more explicit in the text, which lead to a slightly higher frequency. The variation in frequency of the occurrence of the processes PFAS (The Atmosphere reflects (part of) the sunlight) and PDAE (The Atmosphere radiates heat in the direction of the Earth) in the drawing summaries can not be explained by differences in the text: PFAS is mentioned in both texts, while PDAE is not. These results indicate that mentioning processes in the text can have influence on the occurrence of the processes in the drawing summaries. Because of these results, the text used for dataset C was adjusted. The processes PAAE and PFAE were added to the text.

The aim of this study was to find out if it is possible to provide evidence that drawings, scored by this scoring method, provide an actual representation of students’ conceptual understanding. To study the conceptual understanding of students, we asked ten of them to think aloud during the drawing process. The results of the video recording and the scoring of the drawing summaries lead to a good inter-instrument reliability of .734 (Cohen’s Kappa). It can be concluded that scoring the spoken explanation and scoring drawings produce comparable results. On the other hand, both of these methods might not represent the real conceptual understanding, since most students read the text out loud instead of saying what they thought. Video recording to try to capture the students’ mental models turned out to be very difficult. This quote illustrates that literary: “*ik weet niet zo goed waar ik moet beginnen, dat is echt moeilijk”* [I don’t know where to start, this is really difficult.] And when the students do start speaking out loud, they often read the text out loud. As is the case in this quote: “*Nou, hier staat, hij gaat licht uitstralen en een deel wordt weerkaatst door de dampkring”* [Well, it says here, it radiates light and a part is being reflected by the Atmosphere.] In some videorecordings, parts of prior knowledge are found. For example this quote: *"En ook zoals de wolken en die weerkaatst het weer terug naar de aarde en zo blijft het meer opwarmen"* [And also like the clouds and they reflect it back to the Earth and this makes it warm up more.]

The interviews held more information about the conceptual understanding of students than their drawings, which is shown by a Cohen's kappa of .625. Which is still a good inter-instrument agreement but does show a minor difference. Comparing the results of the analysis of the explanation students who joined in an interview with the analysis of their drawings with an independent sample t-test with alpha 0.05 revealed that only PDE as a process was significantly scored different (p=.024, SD=.203) . It was scored in all oral explanations while it was only scored in 8 out of 10 drawing summaries. This results suggest that students reveal more information while explaining their thoughts with speech, but the sample size is too small to conclude this.

The qualitative results from the interview furthermore reveal that there is not more information to be gained from interviews, when comparing them to drawings. An example of this is a quote from one of the students: "*De zon ehm straalt warmte af en dat gaat door de dampkring heen naar de aarde en dat wordt ook een stukje hier afgebroken maar dat heb ik niet getekend zeg maar"* [The Sun eh, radiates heat and that goes through the Atmosphere to the Earth and that is for a small part discontinued but I did not draw that, so to say.] Another quote showed the presence of prior knowledge, the end of a sentence: "*ja, wat we dan nog geleerd hebben is dat deze laag steeds dikker wordt waardoor steeds meer weerkaatst wordt dus de aarde steeds warmer wordt."* [and yeah, what we have learned is that this layer becomes thicker which leads to more reflection, making the earth heat up more.] This quote reveals information about the ideas of this student about the atmosphere.

The detailed description of the three drawings did show some new information about prior knowledge of the students. The drawings differ on their score, the most characteristic difference is the amount of explanation. This is another demonstration of the fact that more information can be obtained when drawings do not consist of pictures only but also contain text.
Another observation was the different ways to depict solar rays. This might indicate mental models about heat transmission: if the students sees this as radiation or convection. The absence of the sun in one of the drawings shows the focus of the students on the earth.

The main objective for this study was to find out to what extent students' drawings of a physical system, the Energy of the Earth and its atmosphere can be interpreted to provide insight into their understanding and mental models of the system. The method used was the comparison dataset C with the results of Van Joolingen and Kenbeek (2013) (dataset A). Comparing the frequencies of the data, showed that most of the categories show the same pattern. Both of the datasets pointed out that PAE (The Earth absorbs (part of) the heat from the Sun) is drawn most frequently, and PDAE (The Atmosphere radiates heat in the direction of the Earth) least. The result that almost all students are able to represent the three objects (Sun, Earth, Atmosphere) in their drawing is in line with the observation that students are well able to translate text into meaningful drawing summaries, which was described by Ainsworth and Iacovides (2005). On the other hand, representing all the relevant processes in the drawing summaries was more difficult for students, most students represented less than half of them. Van Joolingen & Kenbeek added PDAE and PDAU when students drew this process without it being present in the text. In dataset C, almost none of the students did. To validate the findings by Van Joolingen and Kenbeek (2013), an exploratory factor analysis was performed to gain a deeper insight in the drawn representations. Unfortunately, the data from the current study did not show the same factors as the study by Van Joolingen and Kenbeek (2013). This was also the case when comparing the factors in dataset B and A. It should be noted that the exploratory factor analysis was performed without the processes PDAU and PDAE, because these were absent in most drawings. Next to that, the difference in sample size (dataset A contained 68 drawing summaries, dataset B contained 35 drawing summaries and dataset C contained 76 drawing summaries) might have influenced the scores as well.

The detailed description of three drawings did show interesting information about possible mental models of students on the concept of heat convection or radiation. For this research it lead to far to study this concept more in depth, but for future research this subject might be interesting.

A final remark is that the influence of the text on the results in the drawing summaries makes us question the real representation of mental models in drawings. The text did have more influence on the drawings than we expected. But we can conclude that drawings give a good impression of an explanation of a student, compared to an oral explanation. Video recording is not the most suitable method to capture students' mental models. Students have trouble thinking aloud. Most of what the students say is just reading the text out loud. The interviews produced a little bit of extra information about prior knowledge of the students.

To capture students' mental models, maybe a group assessment is a suitable next step. Because for this the students have to think aloud to explain their mental model to their peers. An example of an assessment like this is showed in the study by Gijlers et al (2013). In this study, students' collaboratively had to create a drawing that was suitable to explain photosynthesis to other fifth grade students that were not familiar with this topic. In one of the conditions, students worked individually on a drawing, after which they could see their partners' drawing. After this the students had to agree on the elements that would have to be in the joint drawing. The open recall test after the assignment showed that better transactivity of students' dialogue lead to higher learning outcomes (Gijlers et al, 2013). Another suggestion for further research, gained from the study by Gijlers et al, is the removal of the text during the drawing phase of the assignment. In this way drawing summaries might contain more of the students' mental models instead of the re-reading of the text.

5. References

Ainsworth, S. (1999). The functions of multiple representations. *Computer & Education, 33* (2-3), 131-152. doi: 10.1016/S0360-1315(99)00029-9

Ainsworth, S. & Iacovides, I. (2005). Learning by constructing self-explanation diagrams. *Presented at the EARLI, Nicosia, Cyprus.*

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in Science. *Science Education, 333,* 1096-1097. doi: 10.1126/science.1204153

Andersson, B. & Wallin, A. (2000). Students’ understanding of the greenhouse effect, the societal consequences of reducing CO_2 emissions and the problem of ozone layer depletion. *Journal of Research in Science Teaching, 37*(10), 1096–1111. doi: 10.1002/1098-2736(200012)37:10<1096::AID-TEA4>3.0.CO;2-8

Brooks, M. (2009). Drawing, Visualisation and Young Children’s Exploration of “Big Ideas”. *International Journal of Science Education, 31*(3), 319-341. doi: 10.1080/09500690802595771

Chi, M. T. H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science, 1,* 73-105.

Driver, R. (1983). The pupil as scientist? Milton Keynes, England: Open University Press, Milton Keynes & Philadelphia, USA. doi: 10.1111/j.1756-8765.2008.01005.x

Gilbert, J. K. (2005). Visualization in Science, Models and Modeling in Science Education, Springer, Dordrecht, The Netherlands.

Gijlers, H., Weinberger, A., Van Dijk, A. M., Bollen, L. & Van Joolingen, W. (2013). Collaborative drawing on a shared digital canvas in elementary science education: The effects of script and task awareness support. *International Society of the Learning Sciences, Inc. and Springer Science+Business Media New York*. doi: 10.1007/s11412-013-9180-5

Gobert, J. D. & Buckley, B. C. (2000). Introduction to model-based teaching and learning in science education. *International Journal of Science Education, 22*(9), 891-894. doi: 10.1080/095006900416839

Greca, I. M. & Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education, 22*(1), 1-11. doi: 10.1080/095006900289976

Jakobsson, A., Makitalo, A., & Saljo, R. (2009). Conceptions of knowledge in research on students' understanding of the greenhouse effect: Methodological positions and their consequences for representations of knowing. *Science Education*, *93*(6), 978–995. doi: 10.1002/sce.20341

Jee, B. D., Gentner, D., Uttal, D. H., Sageman, B., Forbus, K., Manduca, C. A., Ormand, C. J, Shipley, T. F. & Tikoff, B. (2009). Drawing on Experience: How Domain Knowledge Is Reflected in Sketches of Scientific Structures and Processes. *Research on Science Education, 44,* 859-883. doi: 10.1007/s11165-014-9405-2

Koulaidis, V. & Christidou, V. (1999). Models of Students' Thinking Concerning the Greenhouse Effect and Teaching Implications. *John Wiley & Sons, Inc.* doi: 10.1002/(SICI)1098-237X(199909)83:5<559::AID-SCE4>3.0.CO;2-E

Leopold. C. & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction, 22,* 16-26. doi: 10.1016/j.learninstruc.2011.05.005

Mason, L., Lowe, R. & Tornatora, M. C. (2013). Self-generated drawings for supporting comprehension of a complex animation. *Contemporary Educational Psychology 38,* 211-224. doi: 10.1016/j.cedpsych.2013.04.001

McNeill, K. L. & Vaughn, M. H. (2012). Urban High School Students' Critical Science Agency: Conceptual Understandings and Environmental Actions Around Climate Change. *Research in Science Education, 42*(2), 373-399. doi: 10.1007/s11165-010-9202-5

Reis, M. J. and Tunnicliffe, S. D. (2001). Students' Understandings of Human Organs and Organ Systems. *Research in Science Education, 31,* 383-399.

Shepardson, D. P., Choi, S., Niyogi, D. & Charusombat, U. (2011). Seventh grade students' mental models of the greenhouse effect. *Environmental Education Research, 17*(1), 1-17. doi: 10.1080/13504620903564549

Shepardson, D. P., Niyogi, D., Choi, S. & Charusombat, U. (2009). Students' conceptions about the greenhouse effect, global warming, and climate change. *Climatic Change, 104,* 481-507. doi: 10.1007/s10584-009-9786-9

Van Joolingen, W. & Kenbeek, W. (2013). Understanding Students' Science Drawings of Complex Systems.

Van Meter, P. & Garner, J. (2005). The Promise and Practice of Learner-Generated Drawing: Literature Review and Synthesis. *Educational Psychology Review*, *17*(4), 285-325. doi: 10.1007/s10648-005-8136-3

Vosniadou, S. & Brewer, W. F. (1994). Mental Models of the Day/Night Cycle. *Cognitive Science, 18,* 123-183. doi: 10.1016/0364-0213(94)90022-1

To what extent can students' drawings of a physical system, the Energy of the Earth and its atmosphere be interpreted to provide insight into their understanding and mental models of the system?

6. Appendix

Text used for dataset A

Hoe wordt de aarde warm?

De voornaamste bron van energie voor de Aarde is de Zon. Door de straling van de zon warmt de aarde op. De aarde vangt de zonnestralen op, en wordt daardoor warm. Toch wordt de aarde niet alleen maar steeds warmer, dit komt doordat de Aarde zelf ook weer warmte uitstraalt. Hoe warmer de aarde is, hoe meer warmte ze ook uitstraalt. De warmte die de aarde verliest, verdwijnt in het heelal.

In werkelijkheid wordt niet alle warmte van het zonlicht opgevangen door de aarde, maar wordt een *gedeelte* van het licht weerkaatst. Daarnaast speelt de dampkring ook nog een rol. De dampkring is de laag lucht om de aarde heen waarin wij leven. De dampkring heeft twee eigenschappen: hij weerkaatst een gedeelte van het licht en hij neemt er ook een gedeelte van op.

English translation of the text used for dataset A

How does the earth heat up?

The sun is the most important source of energy for the earth. Through the radiation of the sun, the earth heats up. The earth catches the sunrays, and consequently heats up. Yet the earth does not become warmer and warmer, because it also radiates heat itself. The warmer the earth is, the more heat it radiates. The heat the earth loses this way vanishes in the universe.

In fact, not all of the heat from the sunrays is absorbed by the earth, but part of it is being reflected. The atmosphere has a role to play too. The atmosphere is the layer of air around the earth in which we live. The atmosphere has two properties: it reflects part of, and absorbs part of it.

Text used for Dataset B

Der Treibhauseffekt

Der Treibhauseffekt ist die Erwärmung der Temperaturen auf unserer Erde. Dieser wird durch den Ausstoβ und die Produktion verschiedenster Stoffe und Gase verursacht, die zur vermehrten Rückstrahlung der Wärme an der Atmosphäre Richtung Erde führen.

Die Sonne schickt Lichtstrahlung auf die Erde. Ein Teil dieser Strahlung wird schon beim Auftreffen auf die äuβerste Atmosphärenschicht zurückgestrahlt, der Groβteil der Sonnenstrahlung gelangt jedoch bis zur Erdoberfläche. Dort wird die Strahlung der Sonne in Wärme umgewandelt und wiederum abgestrahlt. Ein gewisser Anteil dieser Wärmestrahlung dringt durch die Lufthülle unserer Erde wieder in den Weltraum. Ein anderer Anteil an Wärmestrahlung prallt von Natur aus an der Atmosphäre ab und wird auf die Erde zurückgeworfen. Diese natürliche Reflexion gewährleistet auf der Erde eine Durchschnittstemperatur von etwa 15 Grad Celsius. Ohne diese Reflexion würde dieser Wert -18 Grad betragen.

Nun kommt es in unserer Zeit jedoch zu vermehrtem Ausstoβ von Gasen, wie Kohlenstoffdioxid (CO2) und so genannten "Treibhausgasen". Diese werden zum Beispiel von Autos und Industrie produziert. Diese Gase lagern sich in der Atmosphäre an und bewirken, dass die Wärmestrahlung nicht in den Weltraum dringen kann, sondern zwischen Atmosphäre und Erdoberfläche bleibt. Die logische Auswirking hiervon ist eine weltweite Temperaturerhöhung.

Translation of the text used for dataset B

The greenhouse effect

The greenhouse effect is the rise of temperature on our earth. This is caused by pollution and the production of several substances and gases, which increase the reflection of heat by the atmosphere in the direction of the earth.

The sun sends sunrays to the earth. A part of this rays is being reflected by the atmosphere, the bigger part of the radiation reaches the earth. Here the radiation of the sun is transformed into heat and reflected. A part of the radiation goes through the atmosphere back into the universe. Another part of the radiation stays naturally in the atmosphere and is being radiated back to the earth. This natural reflection is the reason for an average temperature of 15 degrees Celsius on the earth. Without this reflection, the temperature would have been -18.

However today we have a lot of emission of CO2 and 'greenhousegases'. These are, for example, produced by cars and industry. These gases accumulate in the atmosphere and, by doing so, prevent radiation from going into the universe, keeping the heat between atmosphere and earth. The logical consequence of this is a worldwide increase in temperature.

The text used for dataset C

Hoe wordt de aarde warm?

De voornaamste bron van energie voor de Aarde is de Zon. Door de straling van de zon warmt de aarde op. De aarde vangt de zonnestralen op, en wordt daardoor warm. Toch wordt de aarde niet alleen maar steeds warmer, dit komt doordat de Aarde zelf ook weer warmte uitstraalt. Hoe warmer de aarde is, hoe meer warmte ze ook uitstraalt. De warmte die de aarde verliest, verdwijnt in het heelal.

In werkelijkheid wordt niet alle warmte van het zonlicht opgevangen door de aarde, maar wordt een *gedeelte* van het licht weerkaatst. Daarnaast speelt de dampkring ook nog een rol. De dampkring is de laag lucht om de aarde heen waarin wij leven. De dampkring heeft twee eigenschappen: hij weerkaatst een gedeelte van het licht en hij neemt er ook een gedeelte van op. Beiden gelden zowel voor licht direct van de zon als licht dat is weerkaatst door de aarde.

English translation of the text used for dataset C

How does the earth heat up?

The sun is the most important source of energy for the earth. Through the radiation of the sun, the earth heats up. The earth catches the sunrays, and consequently heats up. Yet the earth does not become warmer and warmer, because it also radiates heat itself. The warmer the earth is, the more heat it radiates. The heat the earth loses this way vanishes in the universe.

In fact, not all of the heat from the sunrays is absorbed by the earth, but part of it is being reflected. The atmosphere has a role to play too. The atmosphere is the layer of air around the earth in which we live. The atmosphere has two properties: it reflects part of, and absorbs part of it. Both are the case for light directly from the sun and for light reflected by the earth.