Research Article

Changes in Children's Answers to Open Questions about the Earth and Gravity

Triin Hannust[1] and Eve Kikas[2, 3]

[1] *Institute of Psychology, University of Tartu, Ülikooli 18, 50090 Tartu, Estonia*
[2] *Faculty of Education, University of Tartu, Ülikooli 18, 50090 Tartu, Estonia*
[3] *Institute of Psychology, Tallinn University, Narva mnt 25, 10120 Tallinn, Estonia*

Correspondence should be addressed to Triin Hannust, triin.hannust@ut.ee

Many studies that have been conducted to describe children's knowledge about the Earth and gravity have produced discrepant results. However, as most of these studies have been cross-sectional and they have used different methods for collecting and analyzing data, the question Do children at some point construct internally consistent but incorrect explanations to elementary astronomical phenomena? has not been fully answered. The aim of the study was to further explore this question by examining how children respond to open questions about the Earth and gravity and how these answers change over time. Schoolchildren's ($N = 159$) answers were examined four times with one-year intervals. It was found that directly after learning the topics in school many children gave synthetic responses and some oscillated between correct and incorrect explanations for a time. By the fourth grade more than half of the children were able to give scientifically accurate answers and good knowledge of facts supported children's ability to correctly generalize their existing knowledge. It was also shown that most children do not construct consistent nonscientific models of the Earth and that only thorough understanding of the discussed phenomena will lead to consistent answering.

1. Introduction

The developing understanding of the Earth as an astronomical body and issues related to that topic (e.g., gravity and seasonal changes, see [1–3]) has been of great interest to researchers. Examining learners' responses to questions about these topics provides information about how learners overcome difficulties caused by ambiguous terms (e.g., the term *round* denotes both a three-dimensional sphere as well as a two-dimensional circle), differences in perspective (e.g., the round shape of the Earth versus the seemingly flat surface we walk on), or everyday experiences that seemingly are not consistent with what is taught (e.g., the experience that unsupported things fall versus the fact that people do not fall off the southern hemisphere).

Over the years the topic has been researched quite a lot but the results have been contradictory. While studies have consistently shown that children often fail to fully grasp the adults' explanations, some authors argue that children construct consistent synthetic nonscientific models (see [4–6]), and others believe that before accurate understanding learners' knowledge is mainly fragmented [7–12].

The use of different methods for collecting and analyzing data has been indicated as the possible source of such contradicting findings (see [6, 8, 11, 13, 14]). The aim of the present study was to examine such claims by studying whether directly after learning at elementary school children are able to express scientific knowledge about the Earth and gravity in response to open questions and to see how their answers change over the following three years. We also analyzed whether the occurrence of the patterns of answers denoting consistent synthetic models is the result of chance. For that end the methods used for data collection and the initial search for consistent models were similar to the ones used by the proponents of the mental models approach (e.g., [5]). However, because in the original works only authors'

own opinions and estimations were used to determine the presence of models, the statistical significance of the combinations of answers was verified with the Configural Frequency Analysis.

1.1. The Expressed Knowledge about the Earth and Gravity. Given that conceptual entities cannot be directly observed and therefore have to be inferred from what the respondents say or do, the methods used to assess knowledge and thinking may strongly influence the results [14]. Previous studies have indicated that question format may influence the performance on comprehension tasks [15]. For example, support for the construction of consistent mental models has mostly been gathered when open-ended questions that provide almost no clues about the desired answers are used. It has been claimed that only questions that do not provide children with clues prompt them to make their own generalizations [6]. By using such methods, studies have found indications of several nonscientific models that children use consistently when answering questions about the Earth and the related topics (see [5, 6]). Furthermore, even some adults have been shown to provide synthetic answers when faced with open questions that do not specify the level of abstraction [13, 16].

However, other researchers believe that in order to avoid underestimating children's knowledge, one should use questions that provide clues about the desired response and/or the correct perspective from which respondent is expected to view the objects under discussion [10, 11, 13, 16] or show a model which can be used as an anchor [14]. When such methods have been used, instead of giving responses that would indicate the presence of a synthetic mental models, even young children give answers that are either fragmented or they produce scientifically accurate answers and reach ceiling-level performance at a relatively early age (see [10, 11, 13, 14, 16]). Furthermore, even some of the studies that have used tasks similar to those used by Vosniadou and colleagues have shown that though young children really give more incorrect answers in response to those questions, most of their answers are inconsistent with each other [7, 8]. For example, in their longitudinal study Hannust and Kikas [8] showed that though some children gave answers that were consistent with a synthetic model of the Earth, those models were not very stable in time and some of the patterns denoting models might even have been the result of chance. However, they also showed that though the participants' (who had not officially learnt the topic) tendency to only describe the objects under discussion decreased and their ability to take a more global point of view increased over the years, the majority of children did not express a consistent scientific understanding even during the final assessment. Vygotsky [17] has proposed that learnt scientific information supports the growth of scientific reasoning and serves as a basis for further knowledge development. When viewed from that angle, the result of Hannust and Kikas [8] demonstrating that mainly those children who did not express their knowledge of facts tended to base the rest of their answers on those features of the discussed objects that could be directly observed and

experienced suggests that perhaps the examined children simply had not enough access to content-specific knowledge that would allow them to construct a consistent model of the Earth. It is possible that children start building models (either synthetic or scientific) in response to ambiguous open questions when they are older, have studied the topic in school, and have acquired relevant topic-related facts.

Furthermore, even if children really do not construct synthetic models in response to ambiguous tasks, the question whether people's ability to deal with such misleading tasks improves still remains. When children are explicitly instructed to choose the global perspective, even 6-year olds are able to use it instead of the local perspective (see [11]). Also, whereas most kindergarteners fail to express consistent scientific understanding when Vosniadou and Brewer's methods are used in assessment and produce answers that are either entirely based on the visible features of the world or that inaccurately combine learnt information with experience-based knowledge, about two-thirds of adults and even some children manage to provide correct responses even when faced with such ambiguous questions (see [8, 13, 16]). These studies indicate that something must happen during the intermediate period which allows most of adults to detect trick questions and express their knowledge accurately despite the ambiguity.

Studies examining how children solve problems in physics and mathematics have shown that when children are confronted with unfamiliar tasks for which they lack a previously learnt solution method, they often tend to rely on a single, perceptually salient dimension and use only one rule [18, 19]. Only when children are presented with relevant learning experiences can they adopt ways of solving problems which they would not have been able to implement without instruction [20]. These findings suggest that instruction probably plays a crucial part in people's ability to answer different types of tasks.

Furthermore, it has been shown that while some children progress directly from the least to the most advanced approach of problem solving, others oscillate between correct and incorrect approaches for a period of time, and a few even regress from correct to chance level responding (e.g., [21]). It has also been indicated that flaws embedded in instruction itself can cause learners to become confused and thereby perhaps promote the construction of synthetic models (see [7]).

Therefore, it seems that in order to gain more knowledge about how people's ability to express their knowledge and to resist the incorrect implications of vague questions improves, we need to examine the changes children's answers undergo after learning.

1.2. Rationale of the Study. Studies have shown that even preschoolers are able to answer questions about scientific phenomena correctly when all possible ambiguity is removed form the questions and/or respondents are provided with a model that they can use as a reference point [10–12, 14]. However, when questions do not provide clues about what kind of information should be used as a starting point, most children have been shown to give nonscientific answers whereas many adults are able to interpret the ambiguous

questions correctly (see [5, 6, 8, 13, 16]). These differences suggest that perhaps even if children are able to remember scientific explanations at an early age, their knowledge remains encapsulated for a while and they are not able to access it without specific hints. It also indicates that the ability to solve difficult tasks improves and that once a person has acquired and understood the explanations fully, s/he should be able to go beyond the directly perceptible and to combine one's experiences correctly with the learnt information even in response to vague open-ended questions. One of the aims of the present study was to examine this possibility and to see whether children's skill of dealing with confusing tasks indeed increases over the years.

Studies investigating how children solve unfamiliar tasks in several fields of science have shown that the applied strategies change with experience and that children whose understanding of the phenomena is still somewhat vague might be more vulnerable to implications embedded in the questions than the children who have acquired a scientific understanding (see [14, 20, 21]). For example, Hannust and Kikas [8] showed that those small children who do not express their knowledge of facts also tend to base the rest of their answers on only directly visible features of the discussed objects whereas those who answer factual questions correctly are also more likely to generalize that knowledge. Therefore we hypothesized that the inability to form accurate generalizations in response to open questions is related to children's failure to express their factual knowledge and that children's own correct answers provide them with a frame of reference that allows them to proceed beyond the direct experiences.

Another question that has received attention is whether children who have to answer ambiguous questions about a topic they have not fully understood construct consistent alternative models of the Earth (see [5, 6]) or give answers that are not consistent with each other (see [8]). For example, Hannust and Kikas [8] showed that although in response to ambiguous questions young children gave many inaccurate answers that seemed to combine personal and scientific information, in most cases those answers were not organized into consistent models. Still, because the participants in that study had not been instructed on the topics, and it has been indicated that instruction itself can cause confusion and may promote the construction of synthetic models (see [7]), the present study aimed to examine whether after learning the topic children start constructing consistent synthetic models. We hypothesized that shortly after instruction most children give answers that inaccurately combine the learnt information with everyday experiences. As previous studies have indicated that children often need time to think about the learnt material and sometimes additional information is necessary before understanding is reached (cf. [22]), we decided to examine children's answers during the following three years as well. We expected to find that the number of children who give scientifically accurate answers to all the questions increases with time.

As several studies have indicated that it is important to make sure that the patterns indicating the presence of consistent nonscientific models are not the results of chance [7, 8, 10], a method of analysis (configural frequency analysis) was used that examined whether the combinations of answers which according to Vosniadou and Brewer [5] indicate the presence of synthetic mental models appeared more often than could be expected by chance.

2. Method

2.1. Participants. One hundred and fifty-nine elementary-school children (88 boys and 71 girls) participated in the study. Their knowledge about the Earth and gravity was assessed four times with one-year intervals. During the first assessment they were in the first grade and their mean age was 91.9 months (SD = 5.06). The first assessment was conducted shortly after children had studied the topics of the Earth and gravity at school.

All participants spoke Estonian and attended schools that service middle-SES families of the second largest town of Estonia and the surrounding areas.

Seventeen children with missing data were excluded from the study. Main reasons for attrition were moving to another area and prolonged absence from school caused by sickness or family reasons. Nobody withdrew from the study due to unwillingness to participate further. The obtained scores of children who were not included in the analyses did not significantly differ from those whose data was used.

2.2. Procedure. In order to assess the knowledge of astronomy, interviews with the children of consenting parents were conducted in a separate room in school. Each child was interviewed by an experimenter who wrote down and recorded all of the child's utterances. Interviews were carried out by four experimenters. The scoring of the data was done on the basis of the notes and transcribed data. The process was spread over several months but children were interviewed with roughly one-year intervals. The same questions and tasks were used during all assessments.

2.3. Measures. In keeping with the comments of Vosniadou and her colleagues [6] that it is important to allow children to express alternative ideas about the Earth and to avoid limiting their choices, four open-ended questions demanding verbal answers followed by a drawing task where children were asked to draw a picture according to instructions were used in the assessments. The questions were adapted from Vosniadou and Brewer's study and assessed children's exposure to certain facts (factual questions) and their ability to generate ideas about the phenomena that cannot be directly observed (generative questions) (see also [5]). In accordance with criticisms about the original questions (e.g., not making clear whether the Earth should be discussed from the viewpoint of a person standing on it versus the view from space, see also [11, 16]), the present questions were formulated so that they indicate more clearly that children should take the global perspective when answering. The questions and tasks as well as the scoring key categories are presented in Appendix A and the combinations of answers indicating synthetic mental models that are similar to those proposed by Vosniadou and Brewer [5] are in Appendix B.

Table 1: Children's knowledge about the Earth and gravity during four assessments.

	First grade M (SD)	Second grade M (SD)	Third grade M (SD)	Fourth grade M (SD)	η^2
Number of facts	1.92 (0.80)[a]	2.11 (0.71)[b]	2.43 (0.66)[c]	2.67 (0.56)[d]	.24
Avoid answering	0.16 (0.39)[a]	0.11 (0.31)[a]	0.10 (0.30)[a]	0.08 (0.30)[a]	.01
Descriptions	0.67 (0.79)[a]	0.78 (0.69)[a]	0.16 (0.43)[b]	0.11 (0.33)[b]	.29
incorrect generalizations	1.36 (0.80)[a]	1.03 (0.63)[b]	1.09 (0.77)[b]	0.82 (0.64)[c]	.10
Scientific answers	0.77 (0.81)[a]	1.04 (0.95)[b]	1.66 (0.86)[c]	2.02 (0.70)[d]	.42

The pairs with the same superscript letters do not differ significantly ($P > .05$) based on ANOVA post hoc (Bonferroni corrected) paired comparisons.

Answers to each question were categorized independently from others. Answers to factual questions were coded as either correct or incorrect. When responding to questions that required generalizations a child could use up to four different ways of answering: (1) avoid answering, (2) choose the local perspective and only describe the visible world and adopt the global perspective, (3) provide incorrectly generalized answers (that synthesizes personal experiences with learnt information differently form scientific explanations; e.g., the Earth is round like a disc), and provide (4) scientific answers (that were consistent with current scientific theories). Also, because it is hard to draw three-dimensional objects, to depict the global perspective, and to align people and rain according to the circle, children's drawings with mixed alignment of people and rain were taken to be indications of an understanding about the effects of gravity and therefore were coded as correct, not as synthetic (cf. [23, 24]).

The first author of the paper and an independent coder coded the answers. For each explanation Kendall's Tau was calculated on the basis of 25 percent of the data. Due to different administration times the assessments were distinguishable. The interrater reliability for the no answer scores was 1.00; for descriptions, 0.93; for incorrect generalizations, 0.87 and for scientific answers, 0.97.

3. Results

3.1. Analysis Strategy.
So as to examine changes in children's answers Repeated-Measures ANOVAs were carried out with children's number of factual answers, descriptions, incorrect generalizations, and scientifically accurate answers as dependent variables. Post hoc analyses were done with the Bonferroni correction.

In order to group individuals according to their profiles of answers the CFA module of the SLEIPNER 2.1 program (see [25]) was used. The Bonferroni correction was used to adjust the significance levels in the analyses. CFA performs configural frequency analysis using the exact binomial test. For the purpose of the analysis all theoretically possible value combinations and the frequency for each configuration of variable categories are determined. For each configuration separately, the observed frequency (e.g., number of children exhibiting that pattern) is compared with the expected frequency under the null hypothesis of no relationships in the data to see whether the pattern occurs more often (types)

or less often (antitypes) than could be expected on the basis of chance.

3.2. General Trends in the Development of Knowledge.
In Table 1 the means and standard deviations of children's scores of astronomy knowledge at four assessment times as well as the results of the Repeated-Measures ANOVA are presented. Post hoc analyses used the Bonferroni correction.

The number of factual answers increased steadily, $F(3, 474) = 49.99$, $P < .01$. Similarly, scientific knowledge scores improved, $F(3, 474) = 112.46$, $P < .01$, with each passing year. The changes in the use of descriptions were significant, $F(3, 474) = 63.68$, $P < .01$; during the first two assessments children used descriptions significantly more often than during the last two assessments. The average use of synthetic answers decreased over the years, $F(3, 474) = 132.94$, $P < .01$. In the first grade children gave synthetic explanations more often than during any of the following assessments.

3.3. The Relation between Children's Expressed Factual Knowledge and Generative Answers.
In order to find out whether the answers children give when responding to generative questions depend on their expressed knowledge of facts, CFA was conducted separately for each assessment. The identified types are shown in Table 2.

During the first and second assessments there were a few children ($n = 5$) who did not express any knowledge of facts and in most questions also failed to generalize their knowledge beyond the immediate experiences. Furthermore, the other types identified during the first two assessments showed that most children who failed to answer all the factual questions correctly also tended to use descriptions and synthetic explanations instead of giving scientific answers in response to those questions that demanded generalizations.

The types where children gave mostly scientifically accurate answers to the general questions were first identified in the third grade (e.g., 13 children gave only correct answers and 48 answered correctly all verbal questions but drew a picture that seemed to depict the hollow Earth). In the fourth grade 91 children answered all the verbal questions correctly but 60 of them drew a picture that resembled the hollow Earth.

3.4. Consistency of Answers.
So as to examine whether children's answers across questions within an assessment

TABLE 2: Relationships between children's expressed knowledge of facts and the number of different types of answers given to generative questions.

Number of facts	No answer score	Descriptions score	Synthetic answer score	Scientific answer score	Observed	Expected	P
			First grade				
0	0	2	1	0	2	0.00	.01
2	0	2	1	0	16	0.72	<.001
2	0	1	2	0	16	0.78	<.001
2	0	0	3	0	11	0.40	<.001
2	0	1	1	1	17	2.95	<.001
2	*0*	*0*	*2*	*1*	*15*	*60.05*	*<.001*[a]
			Second grade				
0	1	1	1	0	3	0.03	.01
2	1	2	0	0	3	0.03	.02
2	1	1	1	0	7	0.29	<.001
2	0	2	1	0	13	0.72	<.001
2	0	1	2	0	15	0.78	<.001
2	0	1	1	1	29	2.95	<.001
			Third grade				
2	0	0	3	0	7	0.40	<.001
2	0	0	2	1	16	2.57	<.001
2	0	0	0	3	9	1.07	.01
3	0	0	1	2	48	7.88	<.001
3	0	0	0	3	13	0.92	<.001
			Fourth grade				
3	0	0	1	2	60	7.88	<.001
3	0	0	0	3	31	0.92	<.001

[a] Antitypes denoting patterns that appear less often than expected by chance are in italics.

were consistent with any mental model, two procedures were undertaken. First, a coding system similar to the one proposed by Vosniadou and Brewer [5] (including the condition that one deviant answer is considered acceptable in the set of responses that identify a model, see Appendix B) was used to assess whether any of the previously found consistent nonscientific mental models that children supposedly construct/use when answering questions about the Earth could be identified (see [5]). As the answers to the fourth question (why a ball might fall) were not significant in the identification of consistent models, they were not included in the analyses.

In the first grade 61 children were classified as using a consistent synthetic model (1 disk Earth, 55 hollow spheres, and 5 flattened spheres) and five children gave answers consistent with scientific understanding. In the second grade 82 instances were classified as possible synthetic models (2 disc Earths, 76 hollow spheres, and 4 flattened spheres) and 14 as correct models of the Earth. In the third grade 69 children gave answers that were consistent with a synthetic model (1 disk Earth, 66 hollow spheres, and 2 flattened Earths) and 24 gave scientifically accurate answers. During the last assessment 73 children gave answers that were consistent with the hollow Earth model and 34 children gave answers that were consistent with the scientific model of the Earth.

Second, CFA was conducted to see whether during an assessment any combination of answers (including the patterns that were considered to be indications of consistent mental models) emerged more often than would be expected by chance. Only two statistically significant patterns emerged—in the fourth grade 60 children (29.84 were expected by chance, $P = .002$) answered all verbal questions correctly but drew a picture that resembled a hollow Earth, and 34 children (8.93 expected by chance, $P < .001$) answered all questions correctly and also drew a picture of the spherical Earth.

4. Discussion

The present longitudinal study examined how elementary school children respond to open questions about the Earth and gravity and how their answers change over time. As expected, children's answers underwent several changes during the examined period. More specifically, we found that factual knowledge and the amount of scientific answers increased steadily over time whereas the use of descriptions and the amount of incorrect generalizations decreased. Also, there were indications that in order to progress from the use of descriptions of their surroundings to the use of more global answers children need to learn basic facts about

the topic. When children are confident in their knowledge, they are more likely to respond consistently and give scientifically accurate answers even in response to vague open questions.

4.1. Changes in Factual, Synthetic, and Scientific Answers. One of the aims of the study was to find out whether and how the ability to solve difficult tasks in elementary astronomy changes over time. In the present study all of the participants had learnt new information about the Earth and gravity at school shortly before the first assessment but the results indicated that when they were faced with difficult open questions for the first time, many of them were not able to apply the learnt information. The fact that the amount of synthetic answers was largest in first grade and decreased afterwards suggests that instead of using scientific explanations straight away several children started by constructing answers that synthesized the learnt material with their own experiences (see [6, 26–28]). Prior to the first assessment, the participants had only a few short lessons about elementary astronomy, no time for discussions, and only a little time to reorganize their existing knowledge to fit the new information, which might be the reason why they were not able to implement that knowledge and had to resort to the construction of synthetic answers [29].

There have even been indications that synthetic ideas might exist alongside the scientific knowledge (see [28]) and that those will not be abandoned unless the scientific concepts are understood, seem plausible, and are more useful than old ideas [27]. It has also been shown that when children have not understood the learnt material completely, they may oscillate between correct and incorrect explanations for a time (see [21]) and in fact during the second and third assessments several of the children gave different types of answers in response to different questions, which implies that they had both accurate knowledge and inaccurate ideas, and switched between those.

As the process of problem solving itself has been identified as a factor leading to strategy change (see [30]) and in the present study children had to answer the same questions repeatedly (not to mention that their day-to-day experiences at school involved lots of different situations that required them to solve problems), the result that children's answers continued to change even long after the initial instruction was in accordance with our expectations. We also believed that over time children become more able to answer all types of questions correctly (including the ones that provide no clues about the expected answer) and that expectation was confirmed by the result that during each consecutive assessment children gave more factual and scientific answers than before.

4.2. Relations between Factual, Synthetic, and Scientific Answers. As several studies have shown that it is important to be confident in one's knowledge (see [8, 14, 20, 21]), the second aim of the study was to examine whether children who have not understood the scientific explanations of the discussed phenomena and/or have not enough confidence in the accuracy of their knowledge to express it are more

influenced by vague questions than those who have acquired a scientific understanding. We expected to find that in order to accurately respond to ambiguous open-ended questions one needs to identify the relevant factual information, which then provides the respondent with a frame of reference that allows them to proceed beyond the direct experience and to take the learnt information into account when generating answers that have not been memorized beforehand.

We found that children with some (but not perfect) factual knowledge tended to give descriptive and/or synthetic answers whereas those who answered all factual questions correctly used mostly scientific answers. The results were in accordance with our expectations and with previous findings from other areas of knowledge which have shown that when faced with completely unfamiliar problems children tend to focus on the most prominent perceptually salient dimension of the discussed object whereas children who have some correct knowledge are more likely to pay attention to relevant variables allowing them to change their conceptions so that they reflect the world more accurately and to adopt a more advanced approach of problem solving [18, 19, 31, 32].

The results confirm that a person needs to know basic facts about a topic and s/he also needs to express that knowledge in order to be able to apply it accurately when solving unfamiliar and/or ambiguous tasks (cf. [17, 27, 33]). However, the present results also support the idea that learning about scientific topics does not always lead directly to scientific understanding. It appears that when respondents' knowledge of facts is incomplete and they have to solve difficult or ambiguous tasks, they sometimes construct synthetic ideas as an intermediate step. This might happen when learners have memorized some topic-related information but have not filled it with correct meaning and thus modify the meaning to fit it into their existing conceptual system (cf. [3, 17, 29]).

4.3. The Consistency of Children's Answers. The final aim was to examine whether after learning children's answers are combined into patterns that indicate the presence of consistent nonscientific models (cf. [5, 7, 7, 16]). We believed that if any consistent nonscientific models can be found at all, those probably appear directly after learning.

The initial search for combinations of answers that have previously been considered to indicate consistent mental models (cf. [5, 10, 11]) showed that when one acceptable deviation was allowed to occur, about 38% of the participants during the first assessment, about 51% during the second, about 43% during the third, and about 46% during the final assessment were classified as using a consistent nonscientific model. Most of the identified patterns indicated the presence of the Hollow Earth model.

The percentage of children classified as having a synthetic pattern was indeed higher than the 29% of kindergarteners identified in the study of Hannust and Kikas [8], which seems to confirm that formal instruction indeed increased children's tendency to construct inaccurate generalizations in response to ambiguous open questions (cf. [7]). It also appeared that instead of decreasing after the first assessment, the number of children classified as having a synthetic

TABLE 3: Expected combinations of responses indicating consistent mental models.

Question	Disc/flat Earth	Hollow Earth	Dual Earth	Flattened Earth	Spherical Earth
What is the shape of the Earth?					
	The Earth is angular/round/like a circle.	The Earth is round/like a sphere.	The Earth is angular/round/like a circle/like a sphere.	The Earth is round/oval/like a sphere.	The Earth is round/like a sphere.
If you moved in one direction for many days, where would you end up?					
	If you moved for many days, you would reach the edge of the Earth/end of the Earth.	Not applicable *or* if you moved for many days, you would reach the inner side of the sphere	If you moved for many days, you would reach the edge of the Earth *or* the round Earth in the sky has no edge.	If you moved for many days, you would reach the side of the Earth	If you moved for many days, you would end up back where you started.
Is it possible to fall off the Earth? Where?					
	One can fall off from edge/at the end of the Earth.	Not applicable *or* one cannot fall of the Earth because there is no open edge/we are inside the Earth.	One can fall off from edge/at the end of the Earth.	Not applicable *or* one can fall off at the side of the Earth.	One cannot fall of the Earth because of gravity
Picture of the Earth with people, clouds, and rain					
	A rectangle/circle with people standing everywhere on the surface. People and rain aligned to the edge of the paper. Clouds are above the people/drawn all over.	A circle, people drawn inside, at the bottom of the circle/on a flat surface inside the circle. People and rain aligned to the edge of the paper. Clouds are on top of the circle (inside or outside)/on the upper half of the circle.	Both a circle and a flat line are drawn. People standing on the line, rain and people aligned to the edge of the paper.	An ellipse, people only on the upper part of the ellipse. People and rain, aligned to the edge of the paper	A circle with people everywhere. People and rain at least partially aligned to the circle (towards the centre). Clouds are everywhere on the surface of the circle/around the periphery of the circle.
What does the term *round* mean, when describing the Earth?					
	Disc	Ball	Ball, disc, or both	Ball	Ball

Each child was allowed one acceptable deviation, including a missing or a "do not know" answer (see further [5]). Children, who had at least one unacceptable deviation or more than one acceptable deviation, were classified as inconsistent.

Hollow Earth model increased over the years, which seemed to suggest that as time passed, children became even more confused. However, when the results of the configural frequency analysis, which showed that only two patterns (in the fourth grade 21% of the participants were classified as having a scientifically accurate model of the Earth and 30% gave answers consistent with the Hollow Earth model) appeared above chance levels, are taken into account, the present results start to take on a different meaning.

First, they further confirm that when the appearance rates of the combinations of answers are not analyzed statistically, researchers make the mistake of identifying significant patterns where actually none exist and that most children do not construct alternative models of the Earth (cf. [7–9, 11, 12]).

Second, the results indicated that solutions of tasks that require a person to represent three-dimensional objects (either on paper or as a small-scale model) might depend more on the artistic talent of the respondent and less on the actual knowledge of that person. Nobes and Panagiotaki [13] have previously shown that even some adults find it hard to produce pictures of a spherical Earth in response to rather short and nonspecific instructions. Children experience even more difficulties when they try to draw a three-dimensional object [23, 34] and in fact the most difficult task of the present study really was the one where children had to draw a picture of the Earth—even during the final assessment only 50 children produced an accurate picture. As the majority of the Hollow Earth models were differentiated from the scientific model on the basis of differences in the drawings, it is quite possible that those pictures were actually drawn by children who had understood the concept of a spherical Earth with gravity but failed to depict it on paper. Instead, they ended up drawing nonscientific models because those were easier to represent. This means that although directly after instruction many children indeed fail to express their knowledge in response to ambiguous questions, they become more able to do so after they have had time to organize and understand the acquired material (e.g., in the first grade children knew less than two facts on average and used almost no correct generalizations whereas in fourth grade the majority answered factual questions correctly and provided at least two correct generalizations). However, unless they are also taught how to express that knowledge by using other means besides words they will continue to produce nonscientific depictions of the Earth.

5. Conclusions

In conclusion we find that when drawing tasks and open questions that provide almost no clues about the correct answer are used to examine children's knowledge, children indeed often fail to give scientific answers. They base their initial answers on their own experiences and often proceed by constructing intermediate solutions for each individual task (indicated by high use of synthetic answers and by the lack of a consistent pattern of answers). Still, some children adopt the scientifically accurate explanations straight away and use them even in response to confusing questions (cf.

[19, 21, 32]). The construction of consistent nonscientific models seems to be very rare and in most cases consistency is directly related to complete understanding of the discussed phenomena.

However, although by the fourth grade some children are able to provide scientific explanations even in response to vague open questions, the relatively high appearance rate of incorrect explanations in previous grades as well as the high percentage of those who are not able to express their knowledge fully even in the fourth grade (or even as an adult for that matter) suggests that ambiguous questions are less suited to assess the existence of knowledge and more appropriate when one wants to find out how accessible that knowledge is or how confident respondents are about their knowledge (cf. [10, 11, 13, 16]).

Furthermore, although tasks that require respondents to express their knowledge by using some other medium than words (e.g., drawing or making a model) negate the possibility that the answers have been simply memorized, they may fail to elicit accurate answers for completely other reasons (e.g., lack of skill), which indicates that such tasks should not be the main source of information about the presence and/or nature of knowledge.

Finally, we believe that the best way to achieve scientific knowledge which is easy to access for all children is to make sure that their preliminary knowledge is sufficiently addressed, access to relevant information has been granted, and the basic knowledge necessary for further knowledge construction has been acquired (cf. [7, 17, 22]).

Appendices

A. Questions and Tasks Used in the Assessment of Children's Knowledge and the Categorization of Answers

Children's answers could be classified as avoidance of answering (coded as 0), descriptions from the local perspective (coded as 1), incorrect generalizations manifested as synthetic answers (coded as 2), or scientific answers (coded as 3).

(1) What is the shape of the Earth where all people live? (*factual knowledge*):

   (0) I do not know;
   (1) describing visible features, no reference to shape (brown, straight);
   (2) describing Earth as a flat figure (disc, rectangle);
   (3) sphere or round.

(2) If you started walking or riding a car/train/airplane across the land in one direction and continued going in the same direction for many days, where would you end up? (*generative knowledge*):

   (0) I do not know;
   (1) names specific places (school, at grandma's place, other towns in Estonia or abroad);

(2) referring to the edge or of the Earth (at the edge/end);

(3) back at the same place where you started from.

(3) Is it possible to fall off the Earth? Where would you fall off (if the answer to the first question had been positive)? Why not (if the answer had been negative)? (*generative knowledge*):

(0) I do not know;

(1) describing daily experiences (yes, into a hole; no, it is not allowed);

(2) referring to the edge or the end (yes, over the edge, from the side; no, because there is no open edge);

(3) no, because the Earth holds us back/gravity holds us.

(4) Why does a ball that has been thrown up fall down again? (*factual knowledge*):

(0) I do not know;

(1) referring to the sky (sky is in the way, clouds do not let it), or the features of the ball (a ball cannot fly, the ball is heavy);

(3) referring to gravity (the Earth pulls it back).

(5) Draw the Earth where all people live. Now draw people on Earth; draw some everywhere where they might be. Draw clouds on the picture, draw them where they might be. Draw the rain falling from the clouds (*generative knowledge*):

(0) elements not systematically connected;

(1) a line with people and objects standing on it. people and rain are aligned according to the edge of the paper; clouds are in the upper part of the page;

(2a) a rectangle or circle with people standing everywhere on the surface: people and rain are aligned to the edge of the paper; clouds are positioned above the people or drawn all over;

(2b) a circle with people drawn inside the circle and as standing at the bottom of the circle or on a flat surface inside the circle: people and rain are aligned according to the edge of the paper; clouds are on top of the circle or on the upper half of the circle;

(2c) both a circle and a flat line are drawn. Other elements are drawn; according to either Scheme (1) or (2a);

(2d) an ellipse with people and clouds positioned only on the upper part of the ellipse: people and rain are aligned to the edge of the paper; clouds are above the Ellipse or only in the upper part of the ellipse;

(3) a circle with people standing everywhere: people and rain are at least partially aligned to the circle (towards the centre); clouds are everywhere on the surface of the circle or around the periphery of the circle.

(6) To clarify the meaning of the term round a Ping-Pong ball and a similar-sized disc were shown to the child. The models were shown to the child after the completion of the drawing tasks so that the objects would not influence the answers given to previous questions (cf. [6, 14, 35]). The child was asked the following question: You said the Earth was round. Did you mean round like a ball or round like a disc? (*factual knowledge*):

(0) no choice;

(2) choosing the disc;
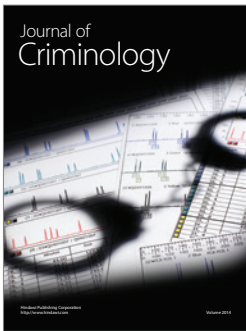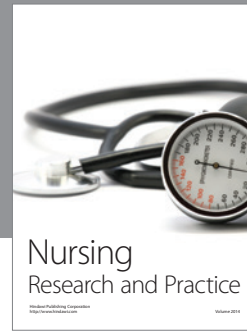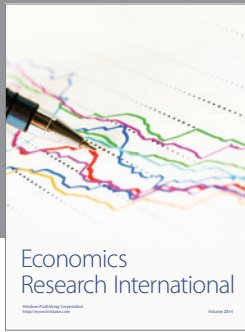
(3) choosing the ball.

B.

See Table 3.

References

[1] E. Kikas, "The impact of teaching on students' definitions and explanations of astronomical phenomena," *Learning and Instruction*, vol. 8, no. 5, pp. 439–454, 1998.

[2] S. Vosniadou, "Capturing and modeling the process of conceptual change," *Learning and Instruction*, vol. 4, no. 1, pp. 45–69, 1994.

[3] S. Vosniadou and W. F. Brewer, "Mental models of the day/night cycle," *Cognitive Science*, vol. 18, no. 1, pp. 123–183, 1994.

[4] W. F. Brewer, "Naïve theories of observational astronomy: review, analysis, and theoretical implications," in *International Handbook of Research on Conceptual Change*, S. Vosniadou, Ed., pp. 155–204, Routledge, New York, NY, USA, 2008.

[5] S. Vosniadou and W. F. Brewer, "Mental models of the earth: a study of conceptual change in childhood," *Cognitive Psychology*, vol. 24, no. 4, pp. 535–585, 1992.

[6] S. Vosniadou, I. Skopeliti, and K. Ikospentaki, "Modes of knowing and ways of reasoning in elementary astronomy," *Cognitive Development*, vol. 19, no. 2, pp. 203–222, 2004.

[7] T. Hannust and E. Kikas, "Children's knowledge of astronomy and its change in the course of learning," *Early Childhood Research Quarterly*, vol. 22, no. 1, pp. 89–104, 2007.

[8] T. Hannust and E. Kikas, "Young children's acquisition of knowledge about the Earth: a longitudinal study," *Journal of Experimental Child Psychology*, vol. 107, no. 2, pp. 164–180, 2010.

[9] G. Nobes, A. E. Martin, and G. Panagiotaki, "The development of scientific knowledge of the Earth," *British Journal of Developmental Psychology*, vol. 23, no. 1, pp. 47–64, 2005.

[10] G. Panagiotaki, G. Nobes, and R. Banerjee, " Is the world round or flat? Children's understanding of the earth," *European Journal of Developmental Psychology*, vol. 3, no. 2, pp. 124–141, 2006.

[11] G. Panagiotaki, G. Nobes, and A. Potton, "Mental models and other misconceptions in children's understanding of

the earth," *Journal of Experimental Child Psychology*, vol. 104, no. 1, pp. 52–67, 2009.

[12] M. Straatemeier, H. L. J. van der Maas, and B. R. J. Jansen, "Children's knowledge of the earth: a new methodological and statistical approach," *Journal of Experimental Child Psychology*, vol. 100, no. 4, pp. 276–296, 2008.

[13] G. Nobes and G. Panagiotaki, "Mental models or methodological artefacts? Adults' 'naïve' responses to a test of children's conceptions of the earth," *British Journal of Psychology*, vol. 100, no. 2, pp. 347–363, 2009.

[14] J. Schoultz, R. Säljö, and J. Wyndhamn, "Heavenly talk: discourse, artifacts, and children's understanding of elementary astronomy," *Human Development*, vol. 44, no. 2-3, pp. 103–118, 2001.

[15] Y. Ozuru, R. Best, C. Bell, A. Witherspoon, and D. S. McNamara, "Influence of question format and text availability on the assessment of expository text comprehension," *Cognition and Instruction*, vol. 25, no. 4, pp. 399–438, 2007.

[16] G. Nobes and G. Panagiotaki, "Adults' representations of the Earth: implications for children's acquisition of scientific concepts," *British Journal of Psychology*, vol. 98, no. 4, pp. 645–665, 2007.

[17] L. Vygotsky, *Thought and Language*, MIT Press, Cambridge, Mass, USA, 1997, Revised edition of 1986, original work published in 1934.

[18] D. J. Messer, K. J. Pine, and C. Butler, "Children's behaviour and cognitions across different balance tasks," *Learning and Instruction*, vol. 18, no. 1, pp. 42–53, 2008.

[19] R. S. Siegler and Z. Chen, "Development of rules and strategies: balancing the old and the new," *Journal of Experimental Child Psychology*, vol. 81, no. 4, pp. 446–457, 2002.

[20] G. S. Halford, G. Andrews, C. Dalton, C. Boag, and T. Zielinski, "Young children's performance on the balance scale: the influence of relational complexity," *Journal of Experimental Child Psychology*, vol. 81, no. 4, pp. 417–445, 2002.

[21] R. S. Siegler and M. Svetina, "What leads children to adopt new strategies? A microgenetic/cross-sectional study of class inclusion," *Child Development*, vol. 77, no. 4, pp. 997–1015, 2006.

[22] I. A. N. Diakidoy and P. Kendeou, "Facilitating conceptual change in astronomy: a comparison of the effectiveness of two instructional approaches," *Learning and Instruction*, vol. 11, no. 1, pp. 1–20, 2001.

[23] M. Blades and C. Spencer, "The development of children's ability to use spatial representations," in *Advances in Child Development*, H. W. Reese, Ed., vol. 25, pp. 157–199, Academic Press, San Diego, Calif, USA, 1994.

[24] A. Toomela, "Drawing development: stages in the representation of a cube and a cylinder," *Child Development*, vol. 70, no. 5, pp. 1141–1150, 1999.

[25] L. R. Bergman, D. Magnusson, and B. M. El Khouri, *Studying Individual Development in and Interindividual Context. A Person-Oriented Approach*, Erlbaum, Mahwah, NJ, USA, 2003.

[26] A. Albanese and M. Vicentini, "Why do we believe that an atom is colourless? Reflections about the teaching of the particle model," *Science and Education*, vol. 6, no. 3, pp. 251–261, 1997.

[27] W. F. Brewer, C. A. Chinn, and A. Samarapungavan, "Explanation in scientists and children," in *Explanation and Cognition*, F. C. Keil and R. A. Wilson, Eds., pp. 279–323, MIT Press, Cambridge, Mass, USA, 2000.

[28] M. Siegal, G. Butterworth, and P. A. Newcombe, "Culture and children's cosmology," *Developmental Science*, vol. 7, no. 3, pp. 308–324, 2004.

[29] M. Chi and S. Ohlsson, "Complex declarative learning," in *Cambridge Handbook of Thinking and Reasoning*, K. J. Holyoak and R. G. Morrison, Eds., pp. 371–399, Cambridge University Press, New York, NY, USA, 2005.

[30] R. S. Siegler, "The rebirth of children's learning," *Child Development*, vol. 71, no. 1, pp. 26–35, 2000.

[31] C. A. Chinn and B. A. Malhotra, "Children's responses to anomalous scientific data: how is conceptual change impeded?" *Journal of Educational Psychology*, vol. 94, no. 2, pp. 327–343, 2002.

[32] R. S. Siegler and Z. Chen, "Developmental differences in rule learning: a microgenetic analysis," *Cognitive Psychology*, vol. 36, no. 3, pp. 273–310, 1998.

[33] S. Vosniadou, "On the nature of naïve physics," in *Reconsidering Conceptual Change: Issues in Theory and Practice*, M. Limon and L. Mason, Eds., pp. 61–76, Kluwer, Dordrecht, The Netherlands, 2002.

[34] A. Karmiloff-Smith, *Beyond Modularity: A Developmental Perspective on Cognitive Science*, MIT Press, Cambridge, Mass, USA, 1996.

[35] S. Vosniadou, I. Skopeliti, and K. Ikospentaki, "Reconsidering the role of artifacts in reasoning: children's understanding of the globe as a model of the Earth," *Learning and Instruction*, vol. 15, no. 4, pp. 333–351, 2005.