



Astronomers' representations of the earth and day / night cycle: Implications for children's acquisition of scientific concepts

Gavin Nobes¹ · Valérie Frède² · Georgia Panagiotaki³

Accepted: 27 December 2021
© The Author(s) 2022

Abstract

Previous researchers (e.g., Vosniadou and Brewer, *Cognitive Psychology* 24:535–585, 1992) have claimed that children have naïve, but coherent, mental models of the earth, such as the flat earth and the hollow sphere. Recent studies have challenged this view, focusing on the original researchers' methods of testing (drawing and open questions) and coding. In this study we tested the construct validity of these methods by asking expert astronomers to complete the same test originally used with 5-year-olds. Many astronomers gave responses that, if given by young children, would have been considered non-scientific and as evidence of naïve mental models. Many gave two or more seemingly contradictory answers to the same questions and, even when only their 'most scientific' responses were considered, fewer than 50% of these expert scientists were classified as having coherent scientific mental models. Comparison with children's responses to a rephrased and disambiguated task indicated that even 6–7-year-olds gave more scientific answers. The astronomers' comments revealed the main reasons why the original task lacks validity: they found many questions confusing and ambiguous. This may well explain many children's responses, too. Since the task incorrectly indicates that experts in the field have fundamental misconceptions, it is likely also to have led to substantial underestimates of children's scientific understanding.

Keywords Astronomy · Scientific expertise · Earth · Children's mental models · Conceptual development

Investigation of scientific misconceptions can shed light on the processes of conceptual development and the sources of information; in particular, concepts might be influenced by individuals' intuitions, observations and experiences, or acquired from cultural sources such as formal education, parent–child conversations, or pictures in books or on screens.

There is evidence that intuitions generate robust misconceptions in, for example, the domain of physics about motion (McCloskey, 1983), sound and heat (Lautrey & Mazens, 2004), and in biology concerning the distinction between living and non-living entities (Inagaki & Hatano, 2002), and prenatal human morphology (Van Schinjdell et al., 2018). Regarding evolution, concepts such as reproduction,

inheritance and biological adaptation appear to be entrenched in a framework of biological essentialism, the belief that evolution is the process by which a species' essence is transformed over time (Shtulman, 2006). Such intuitions can persist and lead to explanations that are coherent and systematic even when contradictory scientific theories have been taught and acquired (Shtulman & Harrington, 2016; Shtulman et al., 2016). Other researchers have highlighted the influence of direct experience in the development of scientific knowledge (e.g., diSessa, 1988, 2017; McDermott, 1984). They argue that intuitive physics (e.g., the concepts of force and motion) consists of many pieces of knowledge – a system of many elements – that are 'fragmented' rather than coherent, and are activated in specific contexts (e.g., diSessa et al., 2004).

In astronomy, some of the most fundamental scientific concepts run counter to both intuition and observation: for instance, the earth appears to be flat and motionless, but is in reality (nearly) spherical and spins on its axis. This area of science therefore lends excellent, perhaps unique, opportunities to investigate the comprehension and acquisition of scientific knowledge by children. If they have flat-earth 'theories', children must have been influenced primarily by their

✉ Gavin Nobes
g.nobes@uea.ac.uk

¹ School of Psychology, University of East Anglia, Norwich, UK

² CLLE, CNRS, UT2J, University of Toulouse, Toulouse, France

³ Norwich Medical School, University of East Anglia, Norwich, UK

own intuitions and observations; after all, it is very unlikely that anyone will have *told* them that it is flat. But if they understand that the earth is spherical, and how day and night are caused, this information must have been acquired from the culture because they could not have directly observed its actual shape and motion.

Many studies have investigated young children's developing knowledge of the earth's shape and the day / night cycle. There are two main theoretical accounts. One is that, consistent with theory theory (e.g., Carey, 1999), children's astronomical concepts are typically coherent (e.g., the earth is flat, and people cannot live 'down under'), and knowledge acquisition is highly constrained by initial and universal presuppositions of flatness and support (Vosniadou, 1994, 2007; Vosniadou & Brewer, 1992; Vosniadou & Skopeliti, 2017). These researchers propose that children form categorical 'mental models' of the earth and of the day / night cycle that are dynamic structures generated from, and constrained by, underlying conceptual structures to solve problems and answer questions.

Following Johnson-Laird (1983), mental models are defined as mental representations that are analogous to the state of affairs they represent. They are theory-like, coherent constructs that, because they are spatially organized, can be manipulated and inspected to enable inference and reasoning (e.g., Gadgil et al., 2012; Goel et al., 1997). This proposal has received support from research on several aspects of reasoning, including deductive reasoning (Knauff et al., 2002), spatial relations (Jahn et al., 2007), and text comprehension (Iani et al., 2017). For example, in Knauff and colleagues' fMRI study, the same right hemispheric cortical areas implicated in spatial working memory, perception and movement control were found to be activated in relational and conditional reasoning. These findings indicate that deductive reasoning is based on spatial representations that are envisaged, transformed and examined to test alternative conclusions, and therefore support the mental model account. In contrast, they are less consistent with accounts of reasoning that emphasize mental logic (the application of formal rules) or visual mental imagery.

According to Vosniadou and colleagues, children's mental models of the earth and day / night cycle are initially based on their intuitions and observations, and only gradually acquire the scientific models as the constraints of flatness and support are overcome. Children's explanations – including those that appear inconsistent with one another – are interpreted as the expressions of their 'naïve', non-scientific mental models of the earth. There are three categories of these models: initial, synthetic and scientific (Vosniadou & Brewer, 1992, 1994). Regarding the earth's shape, a child can form the initial, flat earth model. Later on, when told or shown that the earth is a sphere, the child forms a synthetic model in an attempt to reconcile the apparently irreconcilable intuitive and scientific concepts. For example,

a child can construct a *dual earth* mental model, according to which there are two earths: one of these derives from the child's intuitions and observations, and is a flat plane on which we live; the other reflects scientific information, and is a spherical planet in the sky. In this ingenious way, the child resolves the apparent contradiction of flatness and sphericity. Finally, when all their presuppositions have been relinquished, the child achieves the scientific model in which people live around the spherical earth (Fig. 1).

In contrast, according to the second account, pre-scientific earth concepts lack coherence and are fragmented. Rather than being little scientists who make their own observations and construct their own theory-like mental models, children are considered to be at first theory-free, and then to gradually acquire 'pieces' of knowledge from their culture until – at least in the West – most eventually achieve the coherent scientific model. This view of the acquisition of astronomical concepts resembles diSessa's (1988) regarding

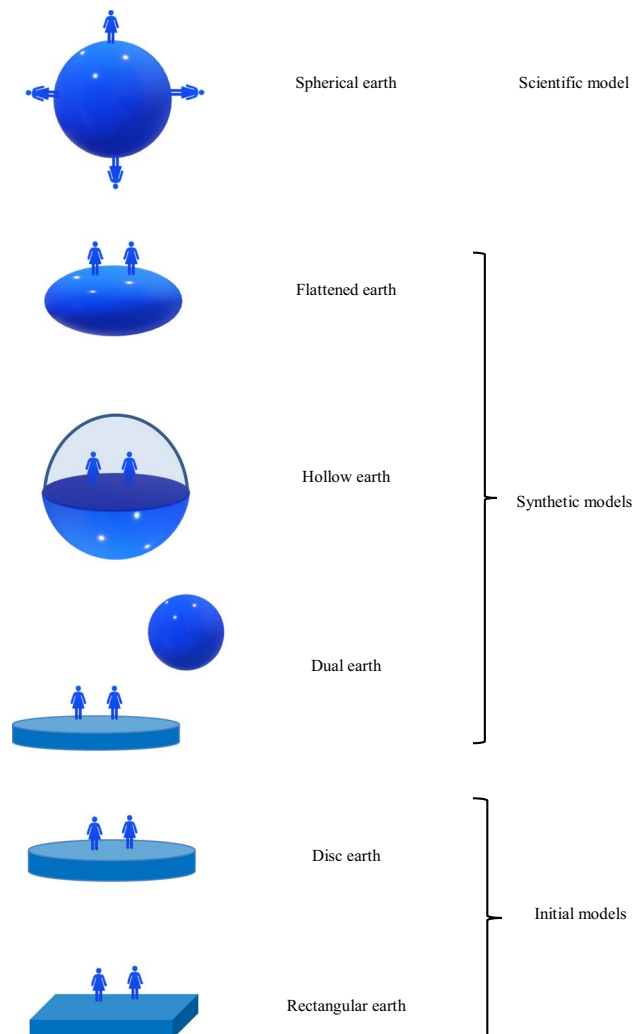


Fig. 1 Mental models of the earth (adapted from Vosniadou & Brewer, 1992)

mechanics, although at least in that domain he argues more for the roles of observations and initial knowledge.

According to this second account, young children simply don't have any views about the shape of the earth or the day / night cycle before they acquire some scientific knowledge. From this perspective, constraining presuppositions are either very weak or non-existent such that, until children acquire the coherent scientific model, they have no initial or synthetic models, but only 'mixed' models, i.e., fragmented or incoherent pieces of knowledge (e.g., Panagiotaki et al., 2006b; Siegal et al., 2011). Consistent with this view, two studies in which children were given 3D models and asked forced-choice questions found little or no evidence of children holding strong intuitions or constructing beliefs of their own (Nobes et al., 2003; Siegal, et al., 2004). Instead, they had either incoherent fragments of knowledge, or the coherent scientific model of the earth.

The replicability and robustness of these two studies have been supported by research using different methods in several countries. For example, Nobes et al. (2005) asked British children, Straatemeier et al. (2008) Dutch children, and Viaopoulou and Papageorgiou (2018) Greek children, to select pictures, and Schoultz et al. (2001) and Ivarsson et al. (2002) used globes or maps when interviewing Swedish children. All reported that even young children showed very little evidence of intuitions of flatness or support, but instead could recognize the earth and understood that it is spherical with people living all around.

Schoultz and colleagues challenged the mental models view from a sociocultural perspective according to which, rather than being located "inside the head", cognition is flexible, situated, and mediated by physical, conceptual and discursive artifacts, of which globes, maps and pictures are examples. They argue that, in the context of such psychological tools (Vygotsky, 1986; Wertsch, 1998), questions make more sense, reasoning is scaffolded, and children's responses are considerably more sophisticated and scientific as a result.

The contrasting findings of researchers who support these two accounts are likely to result from their differing methods. Researchers who support the mental models account (e.g., Vosniadou & Brewer, 1992; Vosniadou & Skopeliti, 2017) use open questions and production tasks (drawing or play-dough modelling) to assess children's understanding in astronomy. In contrast, proponents of the second account (e.g., Nobes et al., 2003; Siegal et al., 2004) initially used selection tasks and asked forced-choice questions.

In an effort to avoid this methodological stalemate, more recently researchers have used Vosniadou and colleagues' own methods. They have focused on two aspects. First, Frède et al. (2011) tested Vosniadou and Brewer's (1992) methods of coding and classifying the purported mental models. They replicated Vosniadou's drawing and open-questions task, and obtained similar findings when they used the same coding scheme.

However, when Frède et al. re-analyzed the same data using statistical methods to test for coherence, no evidence for naïve mental models was found.

Similarly, Hannust and Kikas (2007, 2010, 2012) reported that, even for children as young as two years, combinations of responses that might indicate coherent models were no more likely to occur than by chance. In addition, Panagiotaki et al. (2006a) found that similar methods of questioning to those used by Vosniadou led children to give responses that indicated naïve models, whereas responses to different methods indicated fragmented concepts of the earth.

Second, Nobes and Panagiotaki (2007) investigated possible problems with Vosniadou and Brewer's (1992) original questions and drawing instructions. Instead of testing young children, they asked adults to complete questionnaires with these same questions and instructions. They found that some adults drew flat, dual, or hollow earth pictures, and that many gave non-scientific answers. The adults' comments on the questionnaires and in follow-up interviews indicated that they did not have initial or synthetic models of the earth, but instead found the task confusing and challenging. The authors concluded that non-scientific drawings and answers resulted from semantic (e.g., finding words such as 'earth', 'sky' and 'edge' ambiguous) or pragmatic (e.g., inability to draw the earth, sky and people in one 2D picture) errors rather than conceptual ones. To test this hypothesis, Nobes and Panagiotaki (2009) gave adults a different version of the questionnaire comprising questions and instructions that were re-worded to disambiguate them. Substantially fewer adults now gave non-scientific responses. They concluded that, since even adults found the original questionnaire challenging and confusing, so too would children, and that this was the likely explanation of many children's apparently non-scientific drawings and answers.

Panagiotaki et al. (2009) tested this prediction and explanation by interviewing two groups of 6–7-year-old children. One group was given the original task used by Vosniadou and Brewer (1992), and the second group a new, disambiguated version in which the same questions and instructions were re-worded. For example, whereas in the original version children were asked, "What is the shape of the earth?", in the rephrased questionnaire they were first asked to adopt a global perspective: "Let's pretend you are an astronaut in a big spaceship, travelling in space. You are in space, far away from the earth, but you can still see the earth from your spaceship window. What does the earth look like from your spaceship? What is the shape of the earth?" As predicted, children's responses to the new, rephrased questionnaire were substantially more scientific, and evidence of any naïve mental models was substantially reduced.

However, there remains some uncertainty about the findings with adults. The participants were culturally and educationally diverse, and some found aspects of the task – for

example, the location of the sky in relation to the spherical earth – to be challenging even when the questions were rephrased and simplified. It is possible that the apparently non-scientific responses reported by Nobes and Panagiotaki (2007) reflected the *combination* of ambiguous questions and poor understanding of the earth. And, if some adults' understanding of the earth was poor, then some children, too, would be expected to have non-scientific misconceptions. This would allow for the possibility that at least some children (and adults), have strong presuppositions (e.g., of flatness and support) that lead them to generate naïve mental models.

The present study

In this study a novel approach was taken to investigate this possibility. Whereas in the previous research with lay (i.e., non-expert) adults there was some lack of scientific knowledge, we recruited a sample who we could be certain had excellent scientific understanding of the earth and day / night cycle¹: The participants were all professional astronomers or graduate astronomy students. We reasoned that, if these expert astronomers all performed 'perfectly' on Vosniadou and colleagues' original task (that is, they gave only responses that were coded as scientific following the mental model classification), then this would indicate that, at least for these participants, there was no evidence of ambiguity or over-complexity in the instructions and questions. This would mean that Nobes and Panagiotaki's (2007, 2009) findings of some lay adults giving non-scientific responses must result largely or wholly from their lacking the astronomers' excellent scientific knowledge of the earth, i.e., that the lay adults had misconceptions of the earth. This would be consistent with some adults having presuppositions that are strong enough to influence their thinking in some contexts. And, if some adults have these presuppositions even after many years of exposure to scientific information, we can be sure that many children do, too. A finding of perfect, or near-perfect, performance by expert astronomers on the original task would therefore strongly support its construct validity. It would also indicate that Vosniadou and colleagues' findings of apparently non-scientific responses accurately reflected the lay adults' and children's misconceptions.

On the other hand, any apparently non-scientific responses given by the expert scientists could not possibly reflect conceptual difficulties, and so must necessarily reflect

methodological problems with the task. Since Vosniadou and Brewer's (1992) questionnaire was intended to test children's knowledge of simple and fundamental astronomical concepts, a finding that highly qualified astronomers gave supposedly 'non-scientific' responses, or appeared not to have coherent scientific mental models, would indicate that the task lacked construct validity. This would mean that any claims based on findings from it – such as children having naïve mental models of the earth – would be inadequately substantiated. Moreover, given the repeated failure to replicate these findings using different methods, these claims would likely be incorrect.

The recent studies discussed above (e.g., Frède et al., 2011; Hannust & Kikas, 2010; Nobes et al., 2005; Panagiotaki et al., 2006a; Straatemeier et al., 2008) tested only children's and adults' understanding of the characteristics of the earth, such as its shape and the location of the sky and people. Vosniadou and her colleagues also asked children about the day / night cycle, but the reasons for their non-scientific responses have received less attention and have not previously been investigated with adult participants. To test the validity of this aspect of the original task, in the present study the expert astronomers were asked the original questions both about the earth's characteristics and about the day / night cycle.

Our first prediction concerned the frequency of non-scientific responses (drawings and answers). Our hypothesis was that non-scientific responses arise primarily for methodological reasons, in particular the ambiguity of the instructions and questions. This is the case for all participants, regardless of their knowledge of the earth. We therefore expected that, for pragmatic or semantic rather than conceptual reasons, even some expert astronomers would find some instructions and questions difficult to follow and understand, and so would give 'non-scientific' responses.

The second hypothesis concerned the coherence of concepts. We can be confident that all expert astronomers have coherent scientific earth mental models. However, owing to problems with methods of testing and coding, we expected that some would appear to have naïve (i.e., initial or synthetic) or incoherent (i.e., mixed) mental models of the earth.

We also compared expert astronomers' responses with those given by the two groups of children in Panagiotaki et al. (2009). Though both the experts and the children who were asked the same (or similar) original questions might be influenced by possible semantic and pragmatic problems of the original task, only the children would also be influenced by conceptual problems, i.e., their lack of understanding of the earth. Our third hypothesis was therefore that the astronomers would do considerably better (i.e., give more scientific responses, and appear to have fewer naïve mental models) than the first group of 6–7-year-olds, who responded to the original task.

The fourth hypothesis concerned the experts' performance relative to that of the second group of children in Panagiotaki et al. (2009), who were given rephrased, disambiguated

¹ Chi et al. (1981) described physics undergraduate students as 'novices', and compared their representations of physics problems with those of 'expert' PhD students. For two reasons we consider all the participants in the present study to be 'experts': first, all were graduates researching complex astronomical processes, and most were postdoctoral and professional astronomers; and second, the concepts in question (e.g., the earth's shape and rotation, and the location of the sky) were much simpler than those in Chi et al.'s study, and were so elementary that it is very unlikely that any had true misconceptions concerning these aspects of astronomy.

instructions and questions. This hypothesis was left open. A finding that the astronomers (who had no conceptual problems, but did have the original, possibly ambiguous task) did better than this group of children (who presumably had conceptual problems because they were young, but had less ambiguous questions) would indicate that previous findings of children's non-scientific responses were due primarily to their lack of understanding of the earth, and only partly from any ambiguities of the tasks. In contrast, if the 6–7-year-olds in this second group (who were given the disambiguated version of the task) gave more scientific responses than the expert astronomers, the primary reason for non-scientific responses must be methodological – i.e., problems with the original task – rather than conceptual.

Fifth, because the day / night cycle questions tend to be less ambiguous and so less open to misinterpretation, we predicted that the astronomers' day / night cycle answers would be more scientific than their responses to the earth characteristics questions.

The sixth hypothesis was that even these expert scientists would report finding some of the questions and drawing instructions confusing and ambiguous.

Methods

Sample

The participants were 27 professional astronomers and 17 postgraduate and postdoctoral astronomy students who were conducting research at the Paris Observatory and the Toulouse Observatory. They were aged 21 to 68 years ($M = 35.29$ years, $SD = 11.57$), and 14 were female. Their specialties included planetology, astrometry, celestial mechanics, and geophysics. Professional astronomers and graduate students gave similar percentages of scientific responses, $t = 1.29$, $p = 0.20$, $BF_{01} = 1.68$. This Bayes factor is evidence – albeit only weak or ‘anecdotal’ – in favor of the null hypothesis that there would be no difference between these groups of astronomers.

Measures

The first section of the questionnaire was about the characteristics of the earth (Table 1). It comprised French translations of the instructions and questions used in Nobes and Panagiotaki (2007), all of which were similar or identical to Vosniadou and Brewer's (1992). The second section was about the day / night cycle and consisted of questions that were similar or identical to Vosniadou and Brewer's (1994) and Diakidoy et al.'s (1997), translated into French (Table 2). The multiple-choice answers were the most frequent responses reported and classified as initial, synthetic, or scientific by Vosniadou

et al. (1992, 1994), with the additional option of adding any other response. Finally, participants were invited to report whether, and if so why, they had any difficulties in understanding the instructions and questions (see Supplemental Material for the complete questionnaire).

Procedure

Participants completed the paper questionnaires at work. They were informed that their responses were confidential, that we were conducting a program of research on children's understanding of astronomy, and that, because it was designed for children, they might find some or all the questionnaire very easy.

Analysis

To test the first hypothesis – about the frequency of non-scientific responses – percentages of responses to each question about the earth and the day / night cycle were calculated and are presented in Tables 1 and 2, respectively. In addition, examples of participants' drawings of the earth and the day / night cycle are presented in Figs. 2 and 3. Chi-squared tests were conducted to examine whether the astronomers gave more, or fewer, scientific than non-scientific responses.

The second hypothesis was that some experts would appear to have non-scientific mental models of the earth. Participants' answers about the earth's characteristics were analyzed using Vosniadou and colleagues' (1992, 1994) coding scheme to assign an initial, synthetic, scientific, or mixed mental model of the earth. Following their classification (Appendix 1, Table 4) we selected a priori a pattern of possible responses to fit each model.

A computer program was written in MATLAB² to assign a model to each participant. It calculates a score out of 7 for each mental model (scientific, hollow, dual, or flat) by assessing question by question whether the answer is one of the expected answers for the first model, then the second, and so on. If the answer fits with the expected answer of a model, the program adds one point to the total score for this particular model. When participants score 7 out of 7, they are automatically assigned this model. When their maximum score is 6, the inconsistent answer is checked to see if it can be considered as an ‘acceptable deviation’.³ If so, the model is assigned to the participant. If not, or when their maximum score is less than 6, the participant is assigned a ‘mixed’, or inconsistent, model.

In addition to the MATLAB program, the same analysis was performed manually by an independent coder (the third

² The MATLAB program is available in the Supplemental Material.

³ Following Vosniadou and Brewer (1992, p.557), an acceptable deviation is an answer that is not the expected one for a given model but that can be accepted because it might result from a semantic or pragmatic error. Only one deviation is allowed per model.

Table 1 Astronomers' responses to questions about the earth (N = 44*)

Question	n	%	Scientific vs. non-scientific	
			$\chi^2(1)$	p
Q1-3: Draw the earth; draw where the sky and the clouds go; draw some people to show where they live				
<i>2D drawing with people all around**</i>	18	40.90	26.27	<.001
<i>3D drawing with people all around</i>	3	6.82		
<i>2D drawing with people on the surface</i>	14	31.82		
<i>3D drawing with people on the surface</i>	1	2.27		
<i>3D drawing with people on the top</i>	1	2.27		
<i>2D drawing with only one or two people on the side</i>	2	4.55		
<i>2D drawing of a part of the earth (the top)</i>	3	6.82		
No drawing	2	4.55		
Q4: What is the shape of the earth?				
Round like a disc	0	0.00	1.72	.189
<i>Round like a ball</i>	19	40.42		
Flat like a rectangle	0	0.00		
Oval like a flattened ball	24	51.06		
Hollow ball	0	0.00		
Hemispherical like half a ball	0	0.00		
Other	4	8.52		
Q5: Where is the sky?				
Only on top of the earth	3	6.38	29.13	<.001
Inside the earth	0	0.00		
<i>All around the earth</i>	42	89.36		
Somewhere else (state where)	2	4.26		
Don't know	0	0.00		
Q6: Where do people live?				
Inside the earth	0	0.00	5.67	.017
<i>All around the earth</i>	17	33.33		
Only on the top of the earth	29	56.86		
Somewhere else (state where)	5	9.81		
Don't know	0	0.00		
Q7: If you walked for many days in a straight line, where would you end up?				
At the end/edge of the earth	0	0.00	11.26	<.001
<i>Back where I started</i>	35	74.46		
In space or the sky	2	4.26		
Somewhere else (state where)	10	21.28		
Don't know	0	0.00		
Q8: Is there an end / edge to the earth?				
Yes, there is an end/edge to the earth	21	43.75	0.08	.773
<i>No, there is no end/edge to the earth</i>	25	52.09		
Don't know	2	4.16		
Q9: What is below the earth?				
Ground	17	36.17	4.79	.029
<i>Sky, space</i>	16	34.04		
Water	1	2.13		
Something else (state what)	8	17.02		
No answer	5	10.64		
Total				
<i>Scientific</i>	193	58.31	9.14	.003
Non-scientific	138	41.69		

*N can be more than 44 for some questions as some astronomers gave several answers to the same question.

**Italicized responses are those considered 'scientific' by Vosniadou and her colleagues (Vosniadou & Brewer, 1992; Vosniadou et al., 2004).

author), following the method described in Vosniadou and Brewer (1992). First, a list of expected answers corresponding to each model was formulated a priori. Then, participants whose set of responses all corresponded to a particular mental model's expected answers were

allocated to that model. Next, responses that did not correspond to a model were judged to be either acceptable or unacceptable deviations. Participants with only one acceptable deviation were then assigned to a mental model, and those with more than one, or with one

or more unacceptable deviations, were allocated to the ‘mixed’ category.

Each astronomer’s two earth models (one derived from the program, the other from the manual analysis) were compared. Initial agreement between the program and manual coder was 92.9% ($\kappa = 0.86$, $p < 0.001$) (3 out of 42 drawings were interpreted differently, and two protocols were excluded because the participants did not draw the earth). Agreement was reached following discussion between the second author (who wrote the program) and the manual coder.

Some astronomers gave two answers to the same question (e.g., to the question ‘Where do people live?’, some answered both ‘All around the earth’ and ‘Only on the top of the earth’). We calculated two ‘mental model’ scores for each of these astronomers, one based on their most scientific answers (in this case ‘All around the earth’) and the other on their less scientific answers (in this case ‘Only on the top of the earth’).

Drawings were coded from the analyses of the shape of the earth (round or not, partial, or total view), the location of the sky (on top, all around) and the location of people (at the top of the drawing, all around) following the classification described in Vosniadou and Brewer (1992) (see Appendix 1 Table 4, questions 1–3). For instance, a drawing of a circular earth and the sky and people all around was coded as scientific. For the day / night cycle, drawings were coded depending on the motion of the sun or of the earth (indicated by arrows and the number of earths or suns). For example, if the astronomers drew two suns and a fixed earth, then the drawing was coded as “movement of the sun”. If the drawing showed the earth’s rotation, indicated by arrows around the axis, or by two earths in front of fixed suns, then it was coded as “rotation of the earth”.

To test the third and fourth hypotheses, chi-squared tests were conducted to compare astronomers’ and children’s frequencies of scientific responses to individual questions, and their mental model categories. T-tests were conducted to compare their total numbers of scientific responses.

The fifth hypothesis was that astronomers would give more scientific responses to the day / night questions than to the earth characteristics questions. It was assessed using t-tests of scores on the two sections of the questionnaire.

To test the sixth hypothesis, we examined qualitatively the comments given by the astronomers on the questions and questionnaire. Examples of recurring themes are presented here, and, more fully, in the Supplemental Material.

Results

Frequencies of non-scientific responses (Hypothesis 1): Questions about the earth

Frequencies of answers to the earth characteristics section of the questionnaire are reported in Table 1, and examples

of pictures drawn in response to the instructions (Q1-3) are given in Fig. 2.

Although a majority of the astronomers’ responses were ‘scientific’ (as defined in e.g., Vosniadou & Brewer, 1992, 1994; Vosniadou et al., 2004), over 40% were non-scientific. In response to Q6 (Where do people live?) there were marginally significantly more non-scientific answers – ‘Only on top of the earth’ – than the scientific ‘All around the earth’, $\chi^2(1) = 3.13$, $p = 0.08$. Similarly, although non-significant, more responses to Q9 (What is below the earth?) were non-scientific answers such as ‘Ground’ or ‘Water’ than the scientific ‘Sky or space’, and the shape of the earth (Q4) was described as ‘Oval like a flattened ball’ more frequently than the scientific ‘Round like a ball’. The proportions of scientific and non-scientific answers to Q8, ‘Is there an end and / or edge to the earth?’ were similar. In contrast, many more of the drawings of the earth and their answers to Q5 (Where is the sky?) and Q7 (Where would you end up?) were scientific than non-scientific.

Questions about the day / night cycle

Frequencies of responses are reported in Table 2 and examples of drawings corresponding to the drawing instructions (Q14-15) are given in Fig. 3. More than two-thirds of the astronomers’ responses to these questions were the expected, ‘scientific’ ones. However, in response to Q13 the supposedly non-scientific answer that the sun moves was given much more frequently than that it does not, $\chi^2(1) = 21.43$, $p < 0.001$, and the large majority of astronomers drew the sun moving rather than the earth rotating (Q14-15), $\chi^2(1) = 12.30$, $p < 0.001$. As expected, most answers to Qs 17 and 18 were that the moon is ‘Somewhere in space around the earth’.

Mental models of the earth (Hypothesis 2)

Two of the astronomers did not draw pictures, and so were excluded from this analysis. The remaining 42 sets of responses were analyzed twice because some participants gave more than one answer to some questions. First, when their most scientific answers were considered, 18 of the 42 (42.86%) astronomers’ mental models of the earth were classified as scientific. Of these, 13 (30.95%) included an acceptable deviation. Second, when their least scientific responses were considered, the number of scientific mental models decreased to 11 (26.19%), nine (21.43%) of which included an acceptable deviation.

Although none of the astronomers’ sets of responses were classified as flat, hollow or dual models, three drawings could be interpreted as flat or flattened earths (Fig. 2).

Table 2 Astronomers' answers to questions about the day / night cycle ($N=44^*$)

Question	<i>n</i>	%	Scientific vs. non-scientific	
			$\chi^2(1)$	<i>p</i>
Q10: Where is the sun during the night?				
Behind clouds	0	0.00	14.52	< .001
Gone down	1	1.85		
In space	10	18.51		
<i>On the other side of the earth**</i>	41	75.94		
Other (state where)	2	3.70		
Q11: How does this happen?				
The sun has revolved around the earth	1	1.85	16.67	< .001
The earth has revolved around the sun	6	11.11		
<i>The earth has rotated</i>	42	77.79		
The clouds block the sun	2	3.70		
The sun has gone into space	1	1.85		
The sun has gone under the earth	0	0.00		
Other (state)	2	3.70		
Q12: Does the earth move?				
<i>Yes</i>	43	97.73	40.09	< .001
No	0	0.00		
No answer	1	2.27		
Q13: Does the sun move?				
<i>Yes</i>	36	81.82	23.27	< .001
<i>No</i>	6	13.65		
No answer	2	4.55		
Q14 & 15: Drawing instructions: Here is the earth and an inhabitant. Now, make it so it is day for that person. Now make it so it is night for that person				
Drew the movement of the sun	33	75.00	13.09	< .001
<i>Drew the rotation of the earth</i>	10	22.73		
Other	1	2.27		
Q16: What has happened?				
<i>The earth has rotated</i>	43	86.00	25.92	< .001
The earth has revolved around the sun	4	8.00		
The sun has moved	2	4.00		
The sun became dark	0	0.00		
Other (state)	1	2.00		
Q17: Where is the moon during the day?				
<i>In the sky, where it is night***</i>	0	0.00	29.46	< .001
Exactly opposite the sun	0	0.00		
Behind clouds	0	0.00		
Gone down	0	0.00		
<i>Somewhere in space around the earth</i>	40	90.91		
Other (state)	3	6.82		
No answer	1	2.27		
Q18: Where is the moon during the night?				
<i>In the sky, above our heads***</i>	1	2.17	25.13	< .001
Exactly opposite the sun	0	0.00		
With the stars	2	4.35		
<i>Somewhere in space around the earth</i>	39	84.79		
Other (state)	3	6.52		
No answer	1	2.17		
Total				
<i>Scientific</i>	265	69.74	59.21	< .001
Non-scientific	115	30.26		

*N can be more than 44 for some questions as some astronomers gave several answers to the same question.

**Italicized responses are those considered 'scientific' by Vosniadou and her colleagues (Vosniadou & Brewer, 1994; Vosniadou & Skopeliti, 2017).

***This answer was considered 'scientific' in Vosniadou et al. (2004).

These are pictures 9 and 10, which show only the top of the earth with the sky on the top, and picture 11, which could be interpreted as a disc because the sky and the clouds are inside the earth and above the people and surface.

Two participants gave 6 out of 7 answers that were consistent with the dual model, but both drew scientific earth pictures with people and the sky all around. Their high dual earth scores resulted from their giving scientific answers to some questions, and non-scientific answers to others, but these were not entirely consistent with the dual earth model proposed by Vosniadou and Brewer. These participants were therefore classified as having a mixed mental model.

All of the other participants (22 participants for the ‘most scientific’ view, and 29 for the ‘least scientific’ view) were also classified as having a mixed mental model.

Comparisons between astronomers’ and children’s responses (Hypotheses 3 and 4)

Hypotheses 3 and 4 were tested by comparing the astronomers’ most scientific responses with those given by the two groups of 6–7-year-old children in Panagiotaki et al. (2009) (Table 3). There were four relevant questions about the earth’s characteristics. The third hypothesis – that the astronomers would give more scientific responses than the children who were given the same version of the task – received only weak or anecdotal support, $M_s = 2.48$ vs. 1.93 , respectively, $t(84) = 1.92$, $p = 0.059$, $BF_{10} = 1.11$. Regarding the fourth hypothesis, the astronomers gave significantly fewer scientific responses than the children who had the rephrased task, $M = 3.25$; $t(127) = -3.70$, $p < 0.001$.

When the astronomers’ most scientific answers were considered, the frequencies of scientific mental models did not differ between astronomers (42.86%) and each of the two groups of children (38.09% for the original task, $\chi^2(1) = 0.19$, $p = 0.65$, and 54.12% for the new task, $\chi^2(1) = 1.42$, $p = 0.23$). However, only 26.19% of the astronomers were coded as having scientific mental models when their least scientific answers were analyzed, which is significantly lower than the 54.12% of children who had the rephrased task, $\chi^2(1) = 8.86$, $p < 0.005$ and does not differ from children in the original task, $\chi^2(1) = 1.36$, $p = 0.24$.

Astronomers’ answers regarding the earth’s characteristics and the day / night cycle (Hypothesis 5)

This was tested by comparing the astronomers’ responses to the two sections of the questionnaire. Their answers about the day / night cycle (83.53% correct) were more scientific than the answers about the earth’s characteristics (63.96% correct), $t(43) = 6.92$, $p < 0.001$.

Astronomers’ comments on the questionnaire (Hypothesis 6)

These comments mainly referred to the ambiguity of the questions and the lack of clarity of some of the terms, as well as to the participants’ difficulty in understanding some apparently simple questions. General comments about the questions included⁴:

- Some questions can lead to several answers (because of their wording) (P21)
- I did not understand many of the questions. Clearer wording would have been appreciated (P26)
- *Were there any questions you didn’t understand?* Yes, many. In fact, I understand the point of the question but the imprecise wording will make it very difficult for any reader who knows the right answers but is less sure of himself than a professional astronomer (child, someone without scientific knowledge, etc.)... They show that the authors have a very limited knowledge of astronomy and a slightly defective capacity for logical reasoning. If you use this questionnaire as it is, you will deduce completely erroneous ideas about children’s answers (P6)
- The words “end”, “edge”, “below”, “above” are very imprecise (P13)
- Some wording can induce confusion. For example, question 8 [Is there an end / edge to the earth?], the earth has a finite volume so in that sense it has an end. But in the usual sense, the earth not being flat, it does not have an edge...(P18)

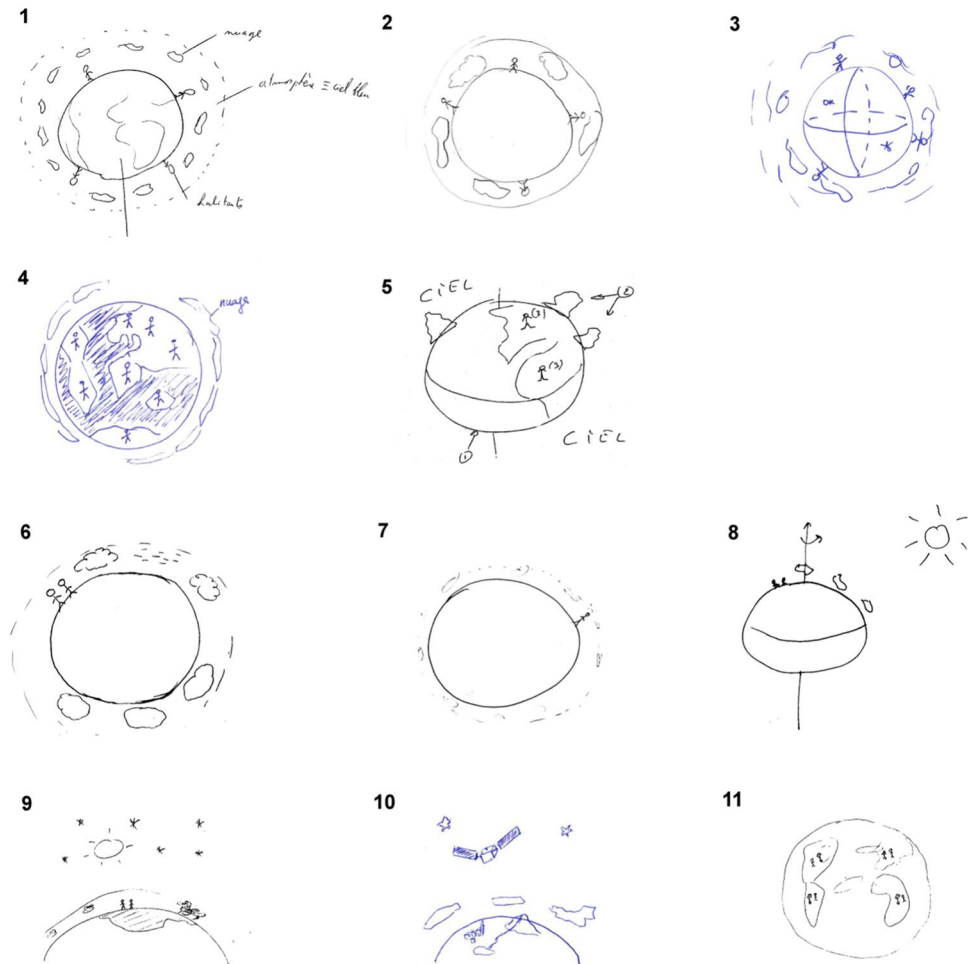
Question 7, about walking in a straight line for many days, caused some confusion:

- The length of the walk is not clearly defined. I don’t know what to answer. We would need to walk a very long time and pass through oceans to come back to the same point. But if we walk for just a few days, we have just moved forward a little, that’s all (P1)
- It is ambiguous; we cannot walk on a straight line remaining on the ground (P8)
- “Being on the surface of the earth, can we walk until we get to an edge of the earth” seems to me a better formulation (P18)

The question that attracted most comments was question 9, ‘What is below the earth?’, with 41% of the astronomers indicating that they found it difficult to answer: 25% said that the question was ambiguous because they were unsure whether it referred to inside or outside the earth.

⁴ See the Supplementary Material for further examples.

Fig. 2 Examples of drawings of the earth. *Scientific representations*: People all around (1–3); People on the surface (4–5); One or two people on the side or top (6–8). *Semi-circular*: People and sky on top of the earth (9–10). *Possible disc representation* (11)



In addition, nine (20%) said that the expression ‘below the earth’ makes no sense. Other comments included:

- The question is maybe deliberately imprecise but again I find it difficult to answer. Below the earth, there is space. Below the earth’s surface there is ground and subsoil. I don’t know what this question means (P1)
- “Below the earth” makes no sense to me: below the surface? (ground?) below the planet? (space?). The interpretation of this expression depends strongly on the image we have of the “earth” (P2)

Some astronomers pointed out the lack of a frame of reference when asking about the relative movements of the earth, sun and moon:

- An object moves relative to other objects. Asking if an object moves without being more precise makes no

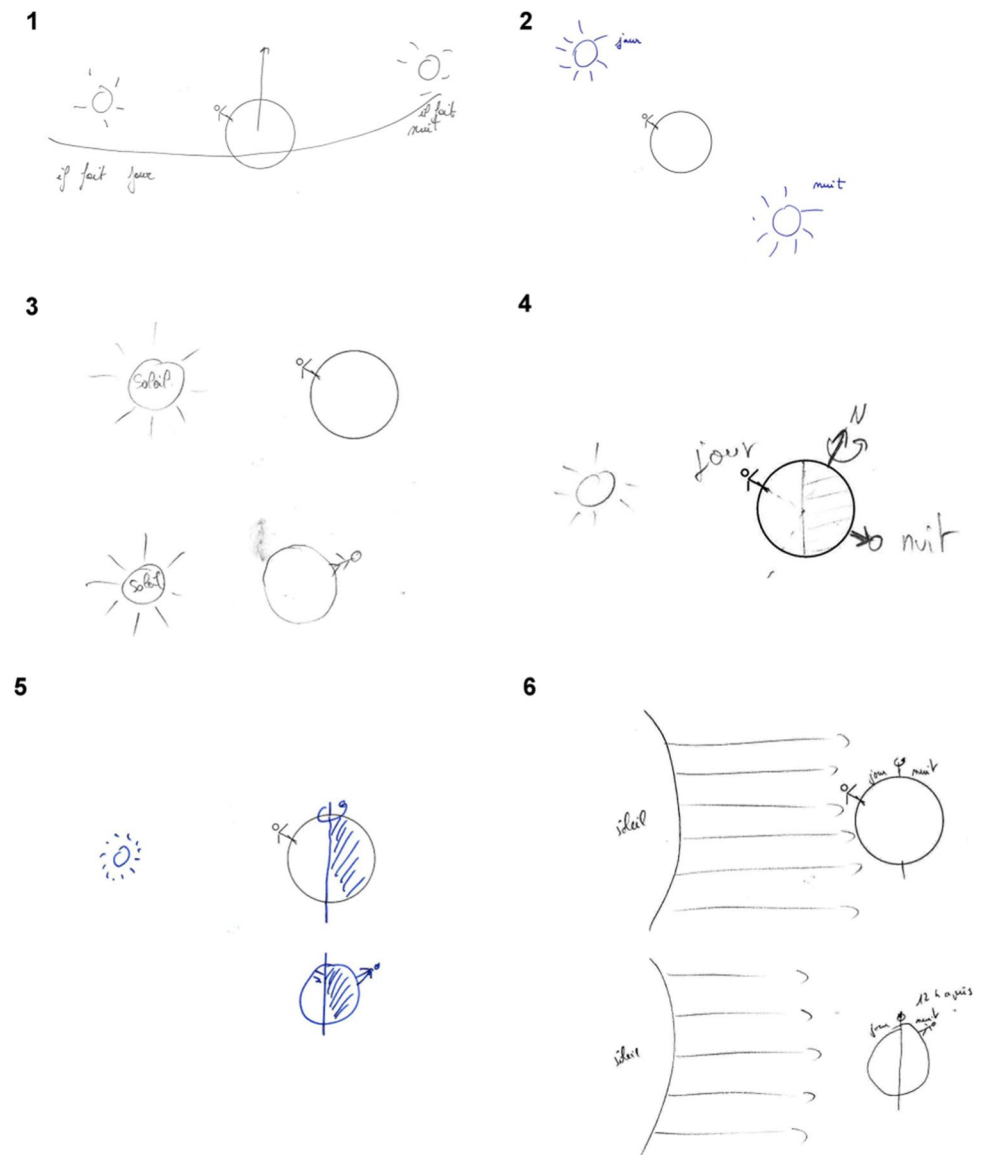
sense... Considered as a celestial body in itself, the earth has no above nor below ...above and below are local concepts (P12)

- Do the sun and earth move? Ok but with regard to what? The sun does not move with regard to the earth but it moves with regard to the other stars (P24)
- The answer depends on the point of view: a [‘The sun has revolved around the earth’] and c [‘The earth has rotated’] answers are both correct [to question 11] (P26)
- Question 13 [Does the sun move?] is ambiguous: the sun moves (also in the galaxy) but it does not influence the day/night cycle, nor the revolution of the earth (P27)

Discussion

In this study, professional and academic astronomers followed drawing instructions and answered questions that were used by Vosniadou and her colleagues (e.g., Diakidoy et al., 1997; Samarapungavan et al., 1996; Vosniadou et al., 2004; Vosniadou

Fig. 3 Examples of drawings of the day / night cycle. Sun moving (1–2). Earth moving (3–6)



& Brewer, 1992, 1994; Vosniadou et al., 2005; Skopeliti & Vosniadou, 2007, 2016) to test young children's understanding of the earth. To our knowledge, this is the first time in this area of science that a test of children's understanding has been investigated by giving the same instructions and questions to expert scientists. Another innovation was that the astronomers were also asked the questions about the day / night cycle: these have not previously been included in studies with adults. In addition, they were asked to give their comments on the instructions and questions.

The first hypothesis, concerning the frequency of responses that the original researchers classified as 'non-scientific', was supported. Many of the astronomers drew pictures and gave answers that would have been considered non-scientific had they been drawn or given by children in Vosniadou and colleagues' studies. Supposedly

non-scientific responses to several questions substantially outnumbered scientific ones. As we can be confident that the expert astronomers had excellent scientific knowledge of the earth's characteristics and the day / night cycle, these responses could only have been given for methodological (e.g., semantic or pragmatic) reasons, not conceptual ones.

The second hypothesis concerned the coherence of astronomers' responses and was also supported. Since many of the astronomers gave multiple responses to questions, it was not possible to assign single mental models to each participant. Instead, we used Vosniadou and Brewer's (1992) method of classification to code each set of responses twice: First, when their 'most scientific' responses were analyzed, fewer than half appeared to have scientific mental models; and second, when their 'least scientific' responses were analyzed, this

decreased to about a quarter. The remaining mental models were classified as mixed, or incoherent, because they included a combination of ‘scientific’ and ‘non-scientific’ responses. Since the astronomers actually had coherent scientific mental models of the earth, the apparent incoherence of their models could only have occurred for methodological reasons.

The third hypothesis was partially supported. Overall, the astronomers’ most scientific responses were only marginally more scientific than the responses given by the 6–7-year-olds in Panagiotaki et al.’s (2009) study who were also given the original task. However, the astronomers’ drawings were more scientific, and more of them answered the question about the long journey’s destination (If you walked for many days in a straight line, where would you end up?) in the expected, ‘scientific’ way. The mental model analysis also indicated that more astronomers than this first group of children (who were given the same, original, version of the task) had scientific representations of the earth.

The fourth hypothesis concerned the astronomers’ responses compared with the second group of 6–7-year-olds in Panagiotaki et al. (2009), who were given a rephrased and disambiguated version of the task. Even when the astronomers’ most scientific responses were considered, the children’s answers about the shape of the earth, what is below the earth, and overall, were significantly more scientific. Together, these tests of Hypotheses 3 and 4 again indicate that the main reason for astronomers’ ‘non-scientific’ responses was methodological rather than conceptual. The astronomers’ conceptual advantage (i.e., their knowing more about the earth) was outweighed by the children’s advantage of responding to rephrased and disambiguated instructions and questions.

The fifth hypothesis – that the astronomers would answer the questions about the day / night cycle with more scientific accuracy than questions about the earth’s characteristics – was supported. This suggests that they found the day / night cycle questions less ambiguous. However, more of the astronomers gave multiple responses to these than to the earth questions, indicating that here, too, Vosniadou and her colleagues’ questions could be answered correctly in several different ways.

Consistent with the sixth hypothesis, the astronomers’ comments indicated that the main reason for their non-scientific responses was that they found the instructions and questions ambiguous and hence difficult to understand. This was supported by their frequently giving more than one answer to the same question, even when these were apparently contradictory (e.g., people live all around the earth, and also only on top of the earth).

Another new point that could only be revealed by investigating experts’ responses was that their views on what should, and what should not, be coded as ‘scientific’ responses sometimes differed from Vosniadou and colleagues’ views. For instance, the large majority of astronomers indicated that the sun *does* move, most described the earth as ‘oval like a flattened ball’, and almost twice as many said that people live ‘only on the top’ as ‘all around the earth’; yet in the original studies these responses were coded as non-scientific. Conversely, the answers ‘when the

moon shines, it is night’ (Vosniadou et al., 2004, p. 212), and that if we can’t see the moon in the day, it is due to the brightness of the sun (Diakidoy et al., 1997, p. 175), were coded as scientifically correct (they are not). The implication of these misclassifications of responses is that many ostensibly ‘non-scientific’ answers given by children are actually more ‘expert-like’ than has generally been recognized in previous research.

The current findings indicate that there are numerous problems with the mental model theorists’ measures of knowledge of the earth and the day / night cycle. Moreover, the comparison of astronomers’ performance on the original task with that of children who were given a rephrased version of the task helps to explain these problems. This comparison also reiterates that children are shown actually to be surprisingly knowledgeable about the earth when the task is phrased in ways that they understand.

However, on their own the current findings do not prove that the same problems apply to, and account for, children’s apparent naïve mental models and high frequencies of non-scientific responses. Instead, they are consistent with, and complement, findings from other studies in which the same questions were rephrased and disambiguated: compared with those who responded to the original task, both lay adults (Nobes & Panagiotaki, 2009) and children (Panagiotaki et al., 2009) gave much higher proportions of scientific responses. This strongly supports the view that the problems with the original task stem principally from the phrasing of the instructions and questions. In addition, when researchers have used different methods that avoid these problematic questions, such as picture and model selection (Nobes et al., 2003, 2005; Siegal et al., 2004; Straatemeier et al., 2008; Viaopoulou & Papageorgiou, 2018), or globes to support children’s understanding (Schoultz et al., 2001), children again show much better understanding of the earth, and little or no evidence of naïve mental models. And finally, when the original or selection methods are used but coherence and consistency are calculated statistically, the initial and synthetic mental models occur no more frequently than would be expected by chance (Frède et al., 2011; Straatemeier et al., 2008; Vaiopoulou & Papageorgiou, 2018).

The findings of this study are therefore consistent with the now extensive body of research that has used a wide variety of methods in several countries, and found little or no evidence of children having naïve mental models of the earth (Frède et al., 2011; Hannust & Kikas, 2007, 2010, 2012; Ivarsson et al., 2002; Nobes et al., 2003, 2005, 2007, 2009; Panagiotaki et al., 2006a, b, 2009; Schoultz et al., 2001; Siegal et al., 2004, 2011; Straatemeier et al., 2008; Vaiopoulou & Papageorgiou, 2018). This body of research indicates that any intuitions or ‘entrenched pre-suppositions’ (e.g., of flatness and support) that are supposed to account for the coherence of initial and synthetic mental models are either very weak or non-existent, because they have little or no effect on children’s thinking. Instead, there is now strong evidence that children’s understanding of even counter-intuitive

Table 3 Comparisons of responses given by astronomers and Panagiotaki et al.'s (2009) two groups of children

Question	Astronomers (N=44)*		Children original task (N=42)		Astronomers vs. children original task		Children new task (N=85)		Astronomers vs. children new task		
	n	%	n	%	$\chi^2(1) / t(84)^\dagger$	p	n	%	$\chi^2(1) / t(127)^\dagger$	p	
Drawing of the earth/people/sky/clouds											
<i>Scientific**</i>	39	88.63	29	69.05	4.98	.026	77	90.59	0.12	.727	
Hollow	0	0.00	5	11.91			2	2.35			
Dual	0	0.00	4	9.52			3	3.53			
Flat/Flattened	3	6.82	4	9.52			2	2.35			
No drawing or don't know	2	4.55	0	0.00			1	1.18			
What is the shape of the earth?											
<i>Round like a ball</i>	19	40.42	13	30.95	1.37	.240	77	90.59	34.22	<.001	
Round like a disc or a circle	0	0.00	26	61.91			4	4.71			
Flat like a rectangle/square	0	0.00	1	2.38			2	2.35			
Oval like a flattened ball	24	51.06	0	0.00			0	0.00			
Other/Don't know	4	8.52	2	4.76			2	2.35			
If you walked for many days in a straight line, where would you end up?											
<i>Back where I started</i>	35	74.46	19	45.24	10.82	.001	57	67.06	2.21	.137	
At the end/edge of the earth	0	0.00	17	40.48			21	24.70			
In space or the sky	2	4.26	0	0.00			0	0.00			
Somewhere else (state where)	10	21.28	3	7.14			4	4.71			
Don't know	0	0.00	3	7.14			3	3.53			
What is below the earth?											
<i>Sky, space</i>	16	34.04	20	47.62	1.12	.290	65	76.48	19.96	<.001	
Ground or water	18	38.30	9	21.43	3.78	.052	10	11.76	14.49	<.001	
Other answers: Something else (state what/ No answer/ Don't know)	13	27.66	13	30.95	0.02	.880	10	11.76	6.25	.012	
Mean score (SD)	2.48 (0.82)		1.93 (1.70)		1.92	.059	3.25 (1.24)		-3.70	<.001	

*N can be more than 44 for some questions as some astronomers gave several answers to the same question.

**Italicized responses are those considered 'scientific' by Vosniadou and her colleagues (Vosniadou & Brewer, 1992; Vosniadou et al., 2004).

†t-tests for comparisons of means.

concepts in this domain of science – such as the earth’s shape and motion – is considerably better than was reported in research that used the original methods of testing, coding and analysis.

The findings of the current study are also consistent with the explanations of children’s and adults’ non-scientific responses and apparent naïve mental models suggested by the recent research, some of which point to problems with the original task’s instructions and questions (in particular, their ambiguity), and others to its methods of coding and analysis (such as ‘finding’ coherent patterns that actually occur at chance levels). Until now, though, expert confirmation of these problems was missing, meaning that some, or even many, adults’ and children’s non-scientific responses reported in previous studies might have occurred not only for methodological reasons, but for conceptual reasons too; that is, the participants might not have known about fundamental aspects of the earth. However, since the expert astronomers in the current study could not possibly have had these conceptual problems, these findings provide perhaps the clearest evidence to date that many non-scientific responses are given for semantic or pragmatic reasons, and that many apparently naïve and mixed mental models (including those reported here) are methodological artifacts rather than true representations of participants’ misconceptions.

The present study therefore contributes to the program of research that supports a view of acquisition of knowledge of the earth in which young children’s concepts show little or no influence of direct observations or intuitions. At first, therefore, children have no views about, for example, the shape of the earth and people’s location on it. They then gradually acquire pieces, or fragments, of knowledge from the culture which, at least in western societies, only becomes coherent when the scientific model is understood and accepted (see also Frède et al., 2011; Siegal et al., 2011).

The implication of the findings of this body of research for science education in this domain is that there is little evidence of strong presuppositions that prevent young children from acquiring the scientific model and that must be overcome before it is understood. Instead, children appear able surprisingly easily and surprisingly early to disregard the apparent evidence of their senses (e.g., that the earth is flat and stationary) in favour of culturally-communicated information. The teacher’s role is therefore more one of providing this information in ways that make sense to and interest young children, rather than of first challenging any supposedly strong non-scientific intuitions in middle or late childhood.

Although it is important to exercise caution in extrapolating from the findings of this body of research to other domains of science, and thus to the acquisition of scientific knowledge in general, at the very least it now provides a clear example of how children can acquire concepts largely or wholly unhindered by strong presuppositions and intuitions, or by the apparent evidence of their senses, such as that the

earth is flat and motionless. Conversely, it highlights the role of cultural transmission in understanding often counterintuitive concepts or unobservable phenomena, and that children acquire these concepts by being exposed to information that they obtain, often as ‘pieces’ of knowledge, from a variety of cultural sources, including formal education, informal conversations with parents, and the media. However, the same is likely also to be the case in other areas of science that include phenomena that, like the sphericity and motion of the earth, cannot be experienced directly because they concern, for example, unobservable entities and concepts such as germs, oxygen, the body’s internal organs, or the cessation of psychological functions after death (Harris et al., 2006; Panagiotaki & Nobes, 2014). Harris and Koenig (2006) argue that, in such areas of science, the testimony of trusted adults plays a fundamental role in the acquisition of knowledge.

The body of recent research to which this study contributes also indicates the importance of considering the contexts of cognition. Ivarsson et al. (2002) and Schoultz et al. (2001) discuss and report how, rather than being solely “inside the head” and context-free, all cognition is situated, flexible, and mediated by cultural artifacts (see also Gottlieb, 1991; Thelen & Smith, 1994; Varela et al., 1991). Our findings are consistent with this account of the context-sensitivity of cognition since they show how even expert scientists appear to have poor scientific understanding in certain contexts, in this case when asked questions that they found ambiguous. The comparison with the previous findings of Panagiotaki et al. (2009) indicates that these experts can appear to know less than young children who were asked disambiguated questions. Similarly, these recent studies indicate that child, lay adult and expert participants responded differently to the original task according to their *interpretation* of the context; some assumed they should take a local perspective when making their responses (i.e., from the earth’s surface, from which the earth seems flat and motionless, with people and the sky only on top, and the sun and moon moving around us), and some a global perspective (as if they were looking at the earth from space). Although it is reasonable to take either or both (as evidenced here by the fact that even some expert scientists did so), the original researchers coded responses from children who took the local perspective as non-scientific, and only those from children who took the global perspective as scientific.

There are similar examples of the influence of context in other areas of science. For instance, Giménez and Harris (2005) explored children’s understanding of the concept of death by telling 7–11-year-olds a story about the death of a grandparent in two different contexts; one secular (where a doctor was present), the other religious (where a priest was present). When they answered questions about the cessation of physical and psychological functions of the dead grandparent, children who heard the secular narrative tended to endorse a biological conception of death, where both physical and psychological

processes cease with death. In contrast, many of those who heard the religious narrative offered a spiritual / religious explanation, according to which many psychological processes continue posthumously. These examples from astronomy and biology show how the contexts in which children and adults are tested, and the participants' interpretations of these contexts, can affect the way they respond to researchers' questions and, consequently, the way researchers interpret their explanations as evidence for or against scientific understanding.

A related point concerns the context in which learning about science takes place, and the sources of information that shape children's knowledge. In particular, the acquisition of scientific knowledge often depends on the cultural context. Frède et al., (2017, 2019) report that, in Burkina Faso, astronomical phenomena are explained in animist oral tradition through the tales and testimonies of professional storytellers, elders and parents. For example, Burkinabe children are told that the earth is flat, "the sun has the will to move around the earth", and that "he had a quarrel with the moon" and so never appears at night. These explanations differ from those taught in these children's schools, where, as in western education, the earth, moon, sun and stars are not considered to be intentional agents, and instead scientific explanations are given. For children growing up in these contexts there is therefore 'interference' between these divergent explanations that can slow acceptance of scientific concepts. In contrast, in western countries, children tend to be given generally convergent information from all sources, and so there is less interference; most children receive the same messages about the earth's sphericity and motion from parents as from teachers. The varied influences of cultural sources in different contexts are also likely to apply to other areas of science and should be tested more widely, especially given that investigation of the acquisition of scientific knowledge has been largely restricted to western countries (Henrich et al., 2010).

While these studies show that conceptions of the earth are strongly influenced by cultural transmission, in other areas of science there could be other sources of information, including – where they are possible – direct observations of phenomena. According to diSessa (1993), children acquire elements, or pieces, of knowledge of physics from their direct experiences, and their conceptual structures consist of a collection of many experiential elements that are independently activated in specific contexts. And, at least in some domains, children can have co-existent, often contradictory scientific *and* non-scientific beliefs at the same time in distinct mental spaces, without necessarily trying to resolve them (Siegal et al., 2004). For example, children – and many adults – often hold contradictory beliefs about death. Some of their coexistent explanations are biologically based, while others reflect spiritual and religious ideas. Panagiotaki et al. (2018) reported that, although 4–11-year-old British children were very good at grasping key biological facts

about death such as its irreversibility or universality, many relied more on supernatural explanations when thinking about the cessation of mental processes following death (e.g., although a dead person cannot come back to life, they are still able to feel things). Similar findings of contradictory but coexistent beliefs have been reported in studies with American (Rosengren et al., 2014) and Mexican children (Gutiérrez et al., 2020).

Limitations

A possible criticism of the use of expert astronomers as participants is that some ambiguities and other problems shown by their answers and comments are not directly relevant to young children, but instead reflect the scientists' expertise. In particular, several commented that the questions lacked a frame of reference, pointing out that movement is always relative; for example, in a heliocentric reference frame, the sun is motionless, but it does move around the centre of the galaxy, which itself is in motion relative to other galaxies. This source of ambiguity for scientists is very unlikely also to explain why children might have found the same questions confusing because, like most lay adults, they do not think in such relativistic terms. Indeed, it is possible that some children gave the same apparently non-scientific responses as scientists for conceptual, rather than semantic reasons. We do not claim, therefore, that all the ambiguities revealed by expert scientists also explain children's supposedly 'non-scientific' responses.

However, relatively few of the astronomers' comments about ambiguities arose from their expertise (see Supplemental Material for further examples). While the astronomers often used more technical language, in fact the large majority of their comments were similar or identical to those made by lay adults (Nobes & Panagiotaki, 2007, 2009) and children (Hannust & Kikas, 2007; Panagiotaki, 2003; Panagiotaki et al., 2009). Considered in the context of this wider literature, it seems likely that the large majority of 'non-scientific' responses given by scientists, lay adults and children alike occurred for similar reasons, in particular the ambiguities of ordinary, everyday language, such as 'below the earth', 'walking for many days in a straight line', 'where do people live', and 'end or edge'. These are probably sources of confusion to all participants, regardless of age or expertise. There is, therefore, a strong case for most – though not all – of the ambiguities revealed by expert astronomers being the same as those facing children, and that they therefore help to explain many of the children's so-called non-scientific responses.

Another possible criticism is that some of the semantic problems reported here might have arisen from the questionnaire having been translated into French. However, several studies have been conducted in other languages (e.g.,

Brewer et al., 1987 – Samoan; Frède et al., 2011 – French; Samarapungavan et al., 1996 – Hindi; Shoultz et al., 2001 – Swedish; Straatemeier et al., 2008 – Dutch; Vosniadou et al., 1996 – Greek) and no issues relating to translation have been reported. Moreover, many of the British adults in Nobes and Panagiotaki (2007) made points concerning the ambiguity of terms used in the questions such as ‘earth’, ‘below’, ‘end’, and ‘edge’ that were very similar to those made by the astronomers in the current study.

Some of the results obtained in this study might have been influenced by the use of a multiple-choice questionnaire. In contrast, in the original task open questions were presented verbally to children. However, apart from being translated into French, the wording of the instructions and questions was similar or identical to that used by Vosniadou and Brewer (1992, 1994). Moreover, the drawing instructions were open (as in the original task), and participants had the options of giving their own answers (‘Other, please state’), or multiple answers, to the questions. These were not, therefore, ‘forced choice’ questions such as those used by Siegal et al. (2004) and Nobes et al. (2003) with children.

Although the percentages of scientific responses given by professional astronomers and graduate students were not significantly different, a limitation of this study is that, because the Bayes factor indicated only weak support for the null hypothesis, we cannot rule out the possibility of a type II error. It is possible that a larger sample would have revealed a clearer difference that might have shed light on the reasons for differences between the astronomers’ responses. However, even the most junior students were postgraduates involved in analyses of complex astronomical processes that contrasted starkly with the simplicity of the concepts tested here. We can therefore be confident that any variation between relatively experienced and inexperienced astronomers’ responses cannot have resulted from differences in their knowledge or understanding.

Similarly, there was only weak, marginally significant support for the third hypothesis that astronomers’ responses would be more scientific than children’s responses to the same, original version of the questionnaire. Again, it is possible that a larger sample of astronomers would have shown that the difference was significant, though failure to reach significance likely resulted more from the relatively high standard deviation of children’s responses (children— $M = 1.93$, $SD = 1.70$ correct responses out of 4; astronomers – $M = 2.48$, $SD = 0.82$). Whether or not larger numbers of astronomers and children would have shown a significant difference, the key point regarding our research questions is that any difference was not substantive, that is, the expert astronomers’ responses to the original task were little or no more scientific than were 6–7-year-old children’s responses.

Future research

The misclassifications of responses revealed here – especially those coded by the original researchers as non-scientific, when actually they are frequently given by expert scientists – show that researchers’ scientific views can lead to misinterpretation of children’s conceptions. Future researchers of children’s scientific knowledge are encouraged to avoid this problem by consulting experts on what is scientifically correct and incorrect. Moreover, just as it has previously been argued that apparently ‘simple’ tasks for young children should be tested on adults (Coley, 2000; Nobes & Panagiotaki, 2007, 2009), this study indicates that these adults should include experts.

The current findings show that the original task does not accurately measure experts’ scientific understanding. To improve methods in this area of research, we therefore recommend use of a more valid instrument. For example, the version used in Panagiotaki et al. (2009) included questions that clarified the perspective (global rather than local) from which they should be answered, and led to children giving substantially more scientific responses. Future researchers are also encouraged to develop tests with repeated measures and mixed methods, such as drawing tasks *and* picture selection, and open *and* multiple-choice questions, to shed further light on the reasons for participants’ responses; in all previous research in this area individual participants have each been tested with only one method.

Recent research from a mental models perspective suggests an alternative approach to assessing understanding. Iani et al. (2017) propose that, as well as propositional representations, mental models can activate motoric representations reflected in gestures. They suggest that ‘anticipatory’ gestures that occur before speech help to organize discourse and prime verbal production, whereas ‘simultaneous’ production of gestures and speech indicates that mental models are already well-articulated. It follows that the relative proportions of anticipatory and simultaneous gestures could therefore provide an index of comprehension. To test this proposal, children and adults were given texts to study and then asked to recall what they read.⁵ Iani et al. report that, consistent with their predictions, participants who showed poor comprehension of a text tended to produce anticipatory gestures, whereas simultaneous gesturing was more typical of participants whose comprehension was better. Future researchers are encouraged to explore the use of this innovative metric of comprehension.

These researchers also report that children and adults who gesture during learning recall more information than those

⁵ For example, one of the texts included the sentence, ‘The force of gravity exerted by the Earth is so strong that even the Moon is attracted towards it, and that is why it revolves around the Earth without ever being able to pull away’. Iani and colleagues describe how a child moved their hand ‘in a circular way around an imaginary central point that starts simultaneously with saying the words “[and that is why] it revolves around the Earth”’ (pp. 1564–1565).

who do not (Cutica & Bucciarelli, 2013; Cutica et al., 2014; Ianì et al., 2017; see also Stevanoni & Salmon, 2005). A related point is the ‘drawing effect’: for example, Wammes et al. (2016) found in seven experiments that, in comparison to writing, drawing while learning resulted in between two and five times as many words being recalled. Since both gestures and drawing involve the coordination of semantic, spatial, perceptual and motor cognitive functions, these findings would appear to be consistent with those of Knauff et al.’s (2002) fMRI study of cortical areas activated in deductive reasoning.

In this study we developed a MATLAB program for classifying mental models. This rule-matching approach ensures objectivity, transparency, and consistency in coding both within and between studies. However, although its use was appropriate here because we aimed to replicate and thereby test Vosniadou and colleagues’ methods, it is recommended that future researchers take a similar statistical approach which allows comparison of participants’ responses with *all* possible patterns of responses, not only the predetermined ones that were interpreted as indicating initial, synthetic and scientific mental models. This type of approach has been taken by Hannust and Kikas (2007, 2010) using configural frequency analysis, Nobes et al. (2005) using cluster analysis, and Straatemeier et al. (2008) and Vaiopoulou and Papageorgiou (2018) using latent class analysis (see also van der Maas & Straatemeier, 2008).

Summary

The investigation of people’s understanding of the earth and the day / night cycle offers researchers perhaps unique insights into the origins, development and structure of

scientific knowledge. Vosniadou and her colleagues pioneered research in this area and have reported intriguing evidence of the role of children’s experiences and intuitions in knowledge acquisition, and of their construction of naïve mental models. This work has inspired much debate and further research, particularly on how to test children’s understanding, and how to interpret children’s responses to these tests. In this study we addressed these issues by investigating the construct validity of Vosniadou and colleagues’ original task. In many ways academic and professional astronomers responded similarly to the young children in the original studies, and the astronomers’ comments and multiple answers to the same questions revealed *why* they gave these responses: in particular, they found it difficult to interpret and answer many of the questions. Moreover, when Vosniadou and Brewer’s (1992) mental model coding scheme was used, fewer than half of the expert scientists were classified as having coherent, scientific earth mental models. When asked rephrased and disambiguated questions, even 6–7-year-old children gave more scientific responses.

These findings support and extend those of other recent studies (e.g., Frède et al., 2011; Hannust & Kikas, 2010, 2012; Nobes & Panagiotaki, 2009; Panagiotaki et al., 2009; Schoultz et al., 2001; Siegal et al., 2004, 2011; Straatemeier et al., 2008). If expert astronomers find the original instructions and questions confusing, young children must find them confusing, too; and if expert astronomers give responses that are classified as ‘non-scientific’ when children make them, then these responses have been misclassified. The current findings therefore contribute to the now substantial body of research that indicates that methodological problems with the original task have led both to underestimates of children’s understanding of the earth and day / night cycle, and to overestimates of the coherence of children’s scientific concepts.

Appendix 1

Table 4

Table 4 Pattern of responses consistent with mental models of the earth

Question	Flat	Hollow	Dual	Spherical
Q1-3. Shape of the earth: drawing (with people and sky)	Square, flat, circle with sky and people on top / Circle with people on the surface and sky on top	Hollow (with people inside the hollow sphere and sky either inside the sphere or on top)	Dual (one flat earth, either circle, square or flat line, and one round earth both in the same drawing. People and sky on top of the flat earth)	Scientific: sphere or circle with people and sky all around / circle or sphere with people on the surface (or people on the side) and sky all around / part of the earth with 3D representation
Q4. Shape of the earth	Round like a disk Flat like a rectangle	Round like a ball a flattened ball Hemispherical like half a ball	Round like a ball Oval like a flattened ball	Round like a ball Oval like a flattened ball
Q5. Location of the sky	Only on the top of the earth	Only on the top of the earth Inside the earth	Only on the top of the earth	All around the earth

Table 4 (continued)

Question	Flat	Hollow	Dual	Spherical
Q6. Location of people	Only on the top of the earth	Inside the earth	Only on the top of the earth	All around the earth
Q7. Long journey's destination	At the end/edge of the earth In space or the sky Somewhere else: water / not far	At the end/edge of the earth Back where I started	At the end/edge of the earth In space or the sky Back where I started Somewhere else: water / not far	Back where I started Somewhere else: water / not far
Q8. End or edge	Yes, there is an end/edge to the earth	Yes, there is an end/edge to the earth or No, there is no end/edge to the earth	Yes, there is an end/edge to the earth	No, there is no end/edge to the earth
Q9. Below the earth	Ground Water, Something else: dirt, rocks, lava, etc	Sky, space Something else: nothing	Ground Water, Sky, space, Something else: nothing, dirt, rocks, lava etc	Sky, space Something else: nothing

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12144-021-02676-6>.

Data Availability All data are available from the authors on reasonable request.

Code availability The Matlab program developed for this research is available from the authors on reasonable request.

Declarations

Conflicts of interest The authors have no conflicts of interest to declare that are relevant to this article.

Ethics approval The study was conducted in accordance with APA and BPS ethical guidelines, and was approved by the School of Psychology Research Ethics Committee, University of East London, UK.

Consent to participate All participants gave their informed consent to participate in the study.

Consent for publication All participants gave their informed consent for their anonymous data to be published in a journal article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Brewer, W. F. (2008). Naive theories of observational astronomy: Review, analysis, and theoretical implications. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 155–204). Routledge.
- Brewer, W.F., Hendrich, D., & Vosniadou, S. (1987). A cross-cultural study of children's development of cosmological models: Samoan and American data. Paper presented at the Third International Conference on Thinking, Honolulu, HI.
- Carey, S. (1999). Sources of conceptual change. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 293–326). Erlbaum.
- Chi, M., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2
- Coley, J. D. (2000). On the importance of comparative research: The case of folkbiology. *Child Development*, 71, 82–90. <https://doi.org/10.1111/1467-8624.00121>
- Cutica, I., & Bucciarelli, M. (2013). Cognitive change in learning from text: Gesturing enhances the construction of the text mental model. *Journal of Cognitive Psychology*, 25, 201–209. <https://doi.org/10.1080/20445911.2012.743987>
- Cutica, I., Iani, F., & Bucciarelli, M. (2014). Learning from text benefits from enactment. *Memory & Cognition*, 42, 1026–1037. <https://doi.org/10.3758/s13421-014-0417-y>
- Diakidoy, I. A., Vosniadou, S., & Hawks, J. D. (1997). Conceptual change in astronomy: Models of the earth and of the day/night cycle in American-Indian children. *European Journal of Psychology of Education*, 12(2), 159–184. <https://doi.org/10.1007/BF03172808>
- diSessa, A. A. (1988). Knowledge in pieces. In G. Forman & P. B. Pufall (Eds.), *Constructivism in the computer age* (pp. 49–70). Lawrence Erlbaum.
- diSessa, A. A. (1993). Towards an epistemology of physics. *Cognition and Instruction*, 10(1), 105–225. <https://doi.org/10.1080/07370008.1985.9649008>

- diSessa, A. A., Gillespie, N., & Esterly, J. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, 28, 843–900. <https://doi.org/10.1016/j.cogsci.2004.05.003>
- diSessa, A. A. (2017). Une introduction accessible à la “connaissance par morceaux”. Modélisation des types de connaissances et de leur rôle dans l’apprentissage. *Education et Didactique*, 11(2), 215–232. <https://doi.org/10.4000/educationdidactique.2781>
- Frède, V., Frappart, S., & Jmel, S. (2019). Burkinabe children’s acquisition and construction of knowledge in astronomy. *Journal of Cross-Cultural Psychology*, 50(7), 870–895. <https://doi.org/10.1177/0022022119860597>
- Frède, V., Frappart, S., & Troadec, B. (2017). Acquisition de connaissances en astronomie : rôle de l’enseignement et du contexte culturel au sein duquel l’enfant se construit. *Psychologie Française*, 62(4), 313–333. <https://doi.org/10.1016/j.psfr.2016.02.001>
- Frède, V., Nobes, G., Frappart, S., Panagiotaki, G., Troadec, B., & Martin, A. (2011). The acquisition of scientific knowledge: The influence of methods of questioning and analysis on the interpretation of children’s conceptions of the earth. *Infant and Child Development*, 20, 432–448. <https://doi.org/10.1002/icd.730>
- Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction*, 22(1), 47–61. <https://doi.org/10.1016/j.learninstruc.2011.06.002>
- Giménez, M., & Harris, P. (2005). Children’s acceptance of conflicting testimony: The case of death. *Journal of Cognition and Culture*, 5(1–2), 143–164. <https://doi.org/10.1163/1568537054068606>
- Goel, V., Gold, B., Kapur, S., & Houle, S. (1997). The seats of reason? An imaging study of deductive and inductive reasoning. *NeuroReport*, 8(5), 1305–1310.
- Gottlieb, G. (1991). Epigenetic systems view of human development. *Developmental Psychology*, 27, 33–34.
- Gutiérrez, I. T., Menendez, D., Jiang, M. J., Hernandez, I. G., Miller, P., & Rosengren, K. S. (2020). Embracing death: Mexican parent and child perspectives on death. *Child Development*, 91, e491–e511. <https://doi.org/10.1111/cdev.13263>
- Hannust, T., & Kikas, E. (2007). Children’s knowledge of astronomy and its change in the course of learning. *Early Childhood Research Quarterly*, 22(1), 89–104. <https://doi.org/10.1016/j.ecresq.2006.11.001>
- Hannust, T., & Kikas, E. (2010). Young children’s acquisition of knowledge about the Earth: A longitudinal study. *Journal of Experimental Child Psychology*, 107(2), 164–180. <https://doi.org/10.1016/j.jecp.2010.04.002>
- Hannust, T., & Kikas, E. (2012). Changes in children’s answers to open questions about the Earth and gravity. *Child Development Research*. Article ID 613674. <https://doi.org/10.1155/2012/613674>
- Harris, P., & Koenig, M. (2006). Trust in testimony: How children learn about science and religion. *Child Development*, 77(3), 505–524. <https://doi.org/10.1111/j.1467-8624.2006.00886.x>
- Harris, P., Pasquini, E., Duke, S., Asscher, J., & Pons, F. (2006). Germs and angels: The role of testimony in young children’s ontology. *Developmental Science*, 9(1), 76–96. <https://doi.org/10.1111/j.1467-7687.2005.00465.x>
- Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–135. <https://doi.org/10.1017/S0140525X0999152X>
- Iani, F., Cutica, I., & Bucciarelli, M. (2017). Timing of gestures: Gestures anticipating or simultaneous with speech as indexes of text comprehension in children and adults. *Cognitive Science*, 41, 1549–1566. <https://doi.org/10.1111/cogs.12381>
- Inagaki, K., & Hatano, G. (2002). *Young children’s naive thinking about the biological world*. Psychology Press. <https://doi.org/10.4324/9780203759844>
- Ivarsson, J., Schoultz, J., & Säljö, R. (2002). Map reading versus mind reading: Revisiting children’s understanding of the shape of the Earth. In M. Limón & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 77–99). Kluwer.
- Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory & Cognition*, 35(8), 2075–2087. <https://doi.org/10.3758/BF03192939>
- Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press
- Knauff, M., Mulack, T., Kassubek, J., Salih, H. R., & Greenlee, M. W. (2002). Spatial imagery in deductive reasoning: A functional MRI study. *Cognitive Brain Research*, 13(2), 203–212. [https://doi.org/10.1016/S0926-6410\(01\)00116-1](https://doi.org/10.1016/S0926-6410(01)00116-1)
- Lautrey, J., & Mazens, K. (2004). Is children’s naive knowledge consistent? A comparison of concepts of sound and heat. *Learning and Instruction*, 14(4), 399–423. <https://doi.org/10.1016/j.learninstruc.2004.06.011>
- McCloskey, M. (1983). Intuitive physics. *Scientific American*, 248(4), 122–130.
- McDermott, L. (1984). Research on conceptual understanding in mechanics. *Physics Today*, 37(7), 24. <https://doi.org/10.1063/1.2916318>
- Nobes, G., Martin, A. E., & Panagiotaki, G. (2005). The development of scientific knowledge of the Earth. *British Journal of Developmental Psychology*, 23(1), 47–64. <https://doi.org/10.1348/026151004x20649>
- Nobes, G., Moore, D. G., Martin, A. E., Clifford, B. R., Butterworth, G., Panagiotaki, G., & Siegal, M. (2003). Children’s understanding of the earth in a multicultural community: Mental models or fragments of knowledge? *Developmental Science*, 6, 72–85. <https://doi.org/10.1111/1467-7687.00257>
- Nobes, G., & Panagiotaki, G. (2007). Adults’ representations of the Earth: Implications for children’s acquisition of scientific concepts. *British Journal of Psychology*, 98, 645–665. <https://doi.org/10.1348/000712607X178119>
- Nobes, G., & Panagiotaki, G. (2009). Mental models or methodological artefacts? Adults’ “naïve” responses to a test of children’s conceptions of the earth. *British Journal of Psychology*, 100(2), 347–363. <https://doi.org/10.1348/000712608X332909>
- Panagiotaki, G. (2003). *Is the Earth flat or round? Unpublished doctoral dissertation*. London, UK: Knowledge acquisition in the domain of astronomy University of Sussex.
- Panagiotaki, G., Hopkins, M., Nobes, G., & Griffiths, D. (2018). Children’s and adults’ understanding of death: Cognitive, parental, and experiential influences. *Journal of Experimental Child Psychology*, 166, 96–115. <https://doi.org/10.1016/j.jecp.2017.07.014>
- Panagiotaki, G., & Nobes, G. (2014). Cultural influences on children’s understanding of the human body and the concept of life. *British Journal of Developmental Psychology*, 32, 276–290. <https://doi.org/10.1111/bjdp.12039>
- Panagiotaki, G., Nobes, G., & Banerjee, R. (2006a). Children’s representations of the earth: A methodological comparison. *British Journal of Developmental Psychology*, 24(2), 353–372. <https://doi.org/10.1348/026151005X39116>
- Panagiotaki, G., Nobes, G., & Banerjee, R. (2006b). Is the world round or flat? Children’s understanding of the earth. *European Journal of Developmental Psychology*, 3(2), 124–141. <https://doi.org/10.1080/17405620500397732>
- Panagiotaki, G., Nobes, G., & Potton, A. (2009). Mental models and other misconceptions in children’s understanding of the earth. *Journal of Experimental Child Psychology*, 104(1), 52–67. <https://doi.org/10.1016/j.jecp.2008.10.003>
- Rosengren, K., Gutiérrez, I., & Schein, S. (2014). Cognitive models of death. *Monographs of the Society for Research in Child Development*, 79, 83–96. <https://www.jstor.org/stable/43772829>

- Samarapungavan, A., Vosniadou, S., & Brewer, W. F. (1996). Mental models of the Earth. Sun, and Moon: Indian children's cosmologies. *Cognitive Development, 11*, 491–521. [https://doi.org/10.1016/S0885-2014\(96\)90015-5](https://doi.org/10.1016/S0885-2014(96)90015-5)
- Schoultz, J., Säljö, R., & Wyndhamn, J. (2001). Heavenly talk: Discourse, artifacts and children's understanding of elementary astronomy. *Human Development, 44*, 103–118. <https://doi.org/10.1159/000057050>
- Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology, 52*, 170–194. <https://doi.org/10.1016/j.cogpsych.2005.10.001>
- Shtulman, A., & Harrington, K. (2016). Tensions between science and intuitions across the life span. *Topics in Cognitive Science, 8*, 118–137. <https://doi.org/10.1111/tops.12174>
- Shtulman, A., Neal, C., & Lindquist, G. (2016). Children's ability to learn evolutionary explanations for biological adaptation. *Early Education and Development, 27*(8), 1222–1236. <https://doi.org/10.1080/10409289.2016.1154418>
- Siegal, M., Butterworth, G., & Newcombe, P. A. (2004). Culture and children's cosmology. *Developmental Science, 7*, 308–324. <https://doi.org/10.1111/j.1467-7687.2004.00350.x>
- Siegal, M., Nobes, G., & Panagiotaki, G. (2011). Children's knowledge of the Earth. *Nature Geoscience, 4*, 134–136. <https://doi.org/10.1038/ngeo1094>
- Skopeliti, I. & Vosniadou, S. (2007). Reasoning with external representations in elementary astronomy. S. Vosniadou, D. Kayser, and A. Protopapas (eds.) *Proceedings of EuroCogSci07, the European Cognitive Science Conference*, 244–249. Delphi, Greece.
- Skopeliti, I., & Vosniadou, S. (2016). The role of categorical information in refutation texts. *Journal of Cognitive Science, 17*(3), 441–468. <https://doi.org/10.17791/jcs.2016.17.3.441>
- Stevanoni, E., & Salmon, K. (2005). Giving memory a hand: Instructing children to gesture enhances their event recall. *Journal of Nonverbal Behavior, 29*, 217–233. <https://doi.org/10.1007/s10919-005-7721-y>
- Straatemeier, M., van der Maas, H. L. J., & Jansen, B. R. J. (2008). Children's knowledge of the earth: A new methodological and statistical approach. *Journal of Experimental Child Psychology, 100*, 276–296. <https://doi.org/10.1016/j.jecp.2008.03.004>
- Thelen, E., & Smith, L. B. (1994). *A dynamical systems approach to the development of perception and action*. MIT Press.
- Vaiopoulou, J., & Papageorgiou, G. (2018). Primary students' conceptions of the Earth: Re-examining a fundamental research hypothesis on mental models. *Preschool and Primary Education, 6*, 23–34. <https://doi.org/10.12681/ppej.14210>
- Van der Maas, H. L. J., & Straatemeier, M. (2008). How to detect cognitive strategies: Commentary on “Differentiation and integration: Guiding principles for analyzing cognitive change.” *Developmental Science, 11*(4), 449–453. <https://doi.org/10.1111/j.1467-7687.2008.00690.x>
- Van Schijndel, T.J.P., van Es, S.E., Franse, R.K., van Bers, B.M.C.W., & Raijmakers, M.E.J. (2018). Children's mental models of prenatal development. *Frontiers in Psychology, 9*. <https://doi.org/10.3389/fpsyg.2018.01835>
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. The MIT Press.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45–69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)
- Vosniadou, S. (2007). The cognitive-situative divide and the problem of conceptual change. *Educational Psychologist, 42*(1), 55–66. <https://doi.org/10.1080/00461520709336918>
- Vosniadou, S., Archodidou, A., Kalogiannidou, A., & Ioannides, C. (1996). How Greek children understand the shape of the earth. *Psychological Issues, 1*, 30–51.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*, 535–585. [https://doi.org/10.1016/0010-0285\(92\)90018-W](https://doi.org/10.1016/0010-0285(92)90018-W)
- Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science, 18*, 123–183. [https://doi.org/10.1016/0364-0213\(94\)90022-1](https://doi.org/10.1016/0364-0213(94)90022-1)
- Vosniadou S., & Skopeliti, I. (2017). Is it the Earth that turns or the Sun that goes behind the mountains? Students' misconceptions about the day/night cycle after reading a science text. *International Journal of Science Education*. <https://doi.org/10.1080/09500693.2017.1361557>
- Vosniadou, S., Skopeliti, I., & Ikospentaki, K. (2004). Modes of knowing and ways of reasoning in elementary astronomy. *Cognitive Development, 19*(2), 203–222. <https://doi.org/10.1016/j.cogdev.2003.12.002>
- Vosniadou, S., Skopeliti, I., & Ikospentaki, K. (2005). Reconsidering the role of artifacts in reasoning: Children's understanding of the globe as a model of the earth. *Learning and Instruction, 15*(4), 333–351. <https://doi.org/10.1016/j.learninstruc.2005.07.004>
- Vygotsky, L. S. (1986). *Thought and language*. MIT Press.
- Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *Quarterly Journal of Experimental Psychology, 69*, 1752–1776. <https://doi.org/10.1080/17470218.2015.1094494>
- Wertsch, J. V. (1998). *Mind as action*. Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.