Exploring Population 2 Students' Ideas about Science

Marit Kjærnsli, Carl Angell, and Svein Lie

Note: A revised version was published in Robitaille D.F. and Beaton A.E. Ed. *Secondary Analysis of the TIMSS Data,* Dordrecht/Boston/London, Kluwer Academic Publisher, 2002, *127-144.*

The aim of this chapter is twofold. One purpose is to illuminate Population 2 students' conceptual understanding of some central phenomena and principles within the different natural sciences. The TIMSS international database is a rich source for various secondary analyses, and in our view, particularly for investigating the conceptual understanding that students around the world have. In addition to examining how much students know and understand within science, we will pay attention to what understanding and conceptions they have.

The second aim of this chapter is to look for consistent similarities and differences between countries concerning the patterns of student responses to the selected items. In a given country, these response patterns depend on curricular antecedents in addition to general cultural and natural factors, such as language, landscape and climate. We will investigate to what extent similar responsepatters are found in "similar" (related to the factors mentioned above) countries. The idea is first to group countries according to cultural and geographical factors, and then to investigate the similarity of the response patterns within each of these groups compared to between the groups. Differences between countries may originate from cultural, geographical and curricular factors. We cannot empirically distinguish between these factors when it comes to explaining differences in response patterns. However, in a few cases we will comment on the possible role of separate factors.

The present analyses were conducted mainly on the basis of free-response items, as these usually allow a deeper understanding of students' own ideas than do multiple choice items. Elsewhere (Angell, Kjærnsli, & Lie, 2000) we have demonstrated that the TIMSS two-digit coding of responses to these items provides a potent tool for secondary analyses of students' conceptual understanding in international achievement tests. But multiple-choice items can also provide valuable information about students' ideas, in particular if the distractors have been carefully constructed for that purpose.

Since the fundamental paper by Driver and Easley (1978) there have been a large number of research studies on students' conceptions within a range of science topics. Overviews (e.g. Wandersee, Mintzes, & Novak, 1993) and bibliographies (e.g. Pfundt & Duit, 1994) have been published. The theoretical paradigm for this large research activity is the so-called constructivist view of learning. The core of this theory is that students learn by constructing their own knowledge. When outer stimuli are treated in the mind together with earlier knowledge and experience of the issue at hand, new insights may be formed. Obviously, within such a framework, it is of crucial importance for science educators and teachers to be aware of students' conceptions within a topic prior to instruction in order to accomplish successful learning.

Selection of science topics for analysis

As said above, this chapter aims at exploring 13-year-old students' conceptions concerning some central topics in classical science. The science topics have been carefully chosen by three criteria. Firstly, there should be an adequate and released Population 2 item that addresses the key point of the phenomena or principle at hand. Secondly, the topic should be regarded as one of the cornerstones of classical science, without being too advanced for 13-year olds. Thirdly, taken together the topics should represent some breadth of the science framework (Robitaille et al. 1993) and allow some general statements about students' understanding of basic scientific concepts. The selected topics and items are listed below. (Item labels in parentheses are the original TIMSS Population 2 labels.):

- the constancy of mass in physical processes. When melting, ice is converted into an amount of water of exactly the same mass. (Q18)
- constancy of temperature during a process of change of state, i.e. melting (Y2)
- the dependency of the process of photosynthesis on light in maintaining an ecosystem (X2b)
- gravity as a force acting on all bodies regardless of their movement and position (K17)
- the crucial role of water balance and temperature regulation for the human body (O16)
- the Earth as a planet with a rotation around its own axis that causes day and night (Q11)
- the ever-changing surface of the Earth, due to building and breaking forces (J1)

Coding rubrics for free-response items

A necessary tool for making diagnostic analyses of responses to free-response items is a coding system that encompasses both the correctness dimension and the diagnostic aspect. In TIMSS this was provided by a two-digit system, originally proposed by the Norwegian TIMSS team. The process of development and the scope and principles of the coding rubrics have been described by Lie, et al. (1996). The idea behind the two-digit system is to apply one two-digit variable to take into account correctness as the first digit (2, 1 or 0 points) and codes for method/error/type of explanation as second digit. The second digit taken alone has no separate meaning across items.

An important feature of the codes is their empirical basis, as all codes are based on authentic responses from many countries in the field trial. The set of actual codes for a particular item therefore reflects the types of responses actually given. Furthermore, score points are allocated from a judgment of quality based on what can be expected from students at age 13, not on requirements of "correctness" from an *a priori* subject matter point of view.

Clustering of countries

We wanted to cluster countries into groups that shared common characteristics. These common characteristics should be well defined from a cultural or geographical perspective, and furthermore, the grouping of countries should also somehow be based on TIMSS achievement data from Population 2. Following Zabulionis (1997), country clusters were constructed based on patterns of differential item functioning in the form of *p*-value residuals (after subtracting over-all *p*-values by items and by country) for all science items. These residuals act as measures of how much more or less difficult an item was, compared to what was "expected" on the basis of the country's general international overall results and the international difficulty of the item.

Table 1. Groups of countries

Group	*Countries*	*Cronbach's alpha*
East Asia	Hong Kong, Japan, Korea, Singapore and Thailand	0.52
East Europe	Bulgaria, Czech Republic, Hungary, Latvia (LSS)*, Lithuania, Romania, Russia, Slovak Republic, Slovenia	0.68
English-speaking	Australia, Canada, England, Ireland, New Zealand, Scotland, United States	0.88
North Europe	Denmark, Iceland, Norway, Sweden, Belgium (Fl)*, Netherlands, Switzerland	0.73
South Europe	Cyprus, Greece, Portugal, Spain	0.44

* The Flemish (Fl.) and French (Fr.) educational systems in Belgium participated separately in TIMSS. Latvia is annotated LSS for Latvian Speaking Schools only.

In Table 1 we have listed the actual country groups. For these groups we have both meaningful cultural or geographical similarities ("validity") and reasonably high achievement similarities ("reliability"). As a measure of this "reliability," Chronbach's alpha is given for each group. The following countries did not fit well into any of the groups and are therefore not included: Austria and Germany (the two were closely linked), France and French Belgium (linked together), Israel, Kuwait, Colombia, the Islamic Republic of Iran, and South Africa. Reference will be made to some of these countries individually. The North European group may seem a bit odd, in particular because of the presence of Switzerland. However, this country as well as Flemish Belgium and the Netherlands, "wanted" to belong to this group according to the empirical clustering of the data. Therefore, we "allowed" them to do so in the present chapter. It is interesting to note that in the SMSO study (Survey of Mathematics and Science Opportunities) linked to TIMSS, Switzerland was often compared to Norway with regard to both teaching style and curricular emphases (Schmidt et al. 1996).

ITEM-SPECIFIC RESULTS

In the following sections, results will be presented and discussed, item by item. The focus will be on the nature of the students' conceptual understanding, as indicated by the occurrence of particular responses within country groups and some individual countries. We have tried to avoid drawing simplistic conclusions about students "having" or "not having" a particular idea, misconception or "conceptual framework," based on responses to stand-alone items. Different aspects of the cognitive networks will be activated depending on the context created by the item.

We do not aim at very precise quantitative statements. Nevertheless, the order of magnitude for percentages of code occurrences is relevant when discussing the occurrence of and background for typical patterns of student thinking. Typically, standard sampling errors for percentages for a country are of the order of a few percent (for the international average less than one percent). Furthermore, the within- and between-country inter-marker reliability for free-response items has been reported by Mullis and Smith (1996). Based on the reported reliabilities, it seems that valid comparisons between frequencies for groups of countries can be made within a few percent. In our presentation of results, frequency distributions will be reported to the nearest percent without any further comments. Results are given for the upper of the two grades included in Population 2. In most countries this means Grade 8 (Beaton et al. 1996). It should also be mentioned that even if response distributions are given for the actual codes for students' responses to free-response items, the descriptions themselves have been simplified. In particular, examples given in the original coding guide (TIMSS, 1995) have been omitted. Furthermore, in a few cases response categories are combined.

For the discussion of the effect of country clustering we need some measure of the proportion of the variability in response patterns that is "explained" by the grouping of countries. For each item we have therefore calculated the between-group variance to total variance ratio (both sum of squares pooled over all response categories).

Constancy of mass (Q18)

One may argue that the history of physics, to a large extent, has been a search for fundamental conservation laws. It has been important to define basic quantities that are conserved under certain conditions. The theme for the first item is a typical example, the concept of mass. Mass is the fundamental ontological concept, a measure of the "amount of matter," a quantity which is conserved during change of state, which is the issue of the item at hand, see Figure 1.

Q18. A glass of water with ice cubes in it has a mass of 300 grams. What will the mass be immediately after the ice has melted? Explain your answer.

Figure 1. Item Q18

For full credit a statement that implies that the mass has not changed is required, and in addition an adequate explanation (see below). Partial credit is given for a correct response with inadequate or no explanation. Table 2 shows the coding guide and the response distributions given in percent.

Table 2. Coding guide and results for item Q18

Code	*Response*	*East Asia*	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
Correct Response						
20	300g with a good explanation	33	26	34	31	25
Partial Response						
10	300g. Explanation is inadequate	6	4	4	6	4
11	300g. No explanation	2	9	1	7	4
Incorrect Response						
70	More than 300g. With explanation	20	10	20	15	15
71	More than 300g. No explanation	1	4	1	2	2
72	Less than 300g. With explanation	16	5	14	12	8
73	Less than 300g. No explanation	1	2	1	2	2
79	Other incorrect	6	4	7	5	8
Blank		15	35	18	20	33

The country clustering can “explain” 42 percent of the variability in response patterns of this item. However, the main between-group differences are largely found among the patterns of wrong responses and non-responses, whereas clusters of countries can account for very little of the variation in correct responses. If we concentrate on the correct aspect of no change of mass (Codes 20, 10, and 11 taken together), discarding for the moment the explanation part, we see that around 40 percent of the students in all country groups responded correctly. Given the clear and direct question in this item it seems to be a straightforward interpretation that about 40 percent of students “believe” in conservation of mass in this case of melting ice.

The rather large non-response rates show large variations between country groups, with particularly high rates in the East and South European groups. The wrong responses are divided between increased mass and decreased mass, with the former being somewhat more frequent. During the marking process in Norway we looked into the explanations given to these responses, which all are necessarily "wrong." Some interesting sources for the misconceptions were evident. Most typically, a common way of reasoning starts from the fact that "ice is lighter than water, since it floats." As is often the case, this daily life expression is in conflict with correct scientific terminology, which states that ice has lower density but is not lighter. Since the process of melting transforms ice into water (which is regarded as "heavier"), the conclusion for these students is that we will end up with more mass.

Explanations of the correct responses could be very different and still receive full credit. The short and easy direct reference to the general principle of constant mass was not given by many. It may well be that some students know this principle, but still do not regard a statement of this "law" as appropriate. After all, referring to a general principle, albeit perfectly correct, really does not explain anything ("Why is it so?"), beyond restating the correct response ("that"). As soon as they try to give an explanation in terms of volume and density, or even microscopic particles, they tend to get confused.

Change of state (Y2)

The next item also focuses on a quantity which, under certain conditions, is constant. Again the topic is the melting process (see Figure 2). The phenomenon focused on here is that during this process the temperature remains constant at the melting point.

Y2. One day when the temperature was just below 0°C, Peter and Ann made snowballs. They put a thermometer into one of the snowballs and it showed 0°C. They tried to make the snowball warmer by holding it in their hands. What do you think the thermometer showed after two minutes?

Figure 2. Item Y2

The concepts of heat and temperature are fundamental to physics, and freezing, melting, boiling, heat transfer, evaporation, and condensation are all phenomena very closely connected to our everyday life. If the notion of everyday physics should have any meaning, the area of heat and temperature must be important. This theme is taught at different levels in school around the world, and a number of common misconceptions have been reported (e.g. Erickson & Tiberghien, 1985; Thomaz, Valente, & Antunes, 1995). Furthermore, a proper discrimination between heat and temperature is very demanding at this age level. Many students have the conception that transferring "heat" to a body always means increasing its temperature.

The response distribution (given in percent) is shown in Table 3. The international mean on this item is only 14 percent for a correct response with an acceptable explanation. Another 30 percent have a partially correct response. This indicates that a substantial part of 13-year-old students do not have adequate knowledge of the fact that snow cannot get warmer than 0 degrees. Another reasonable correct response could have been that snow is a very good insulator and therefore the temperature inside would not change. However, very few students responded in this way.

Table 3. Coding guide and results for item Y2

Code	*Response*	*East Asia*	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
Correct Response						
20	Reports 0 degrees or mentions "the same temperature". Explanation includes: "Snow can not be warmer than 0 degrees" or similar	11	10	13	15	6
29	Other correct	3	3	4	2	3
Partial Response						
10	0 degrees or "the same temperature". No or incorrect explanation	22	16	15	16	17
19	Other partial correct	15	11	15	19	9
Incorrect Response						
70	Above 0 degrees, because the hands are warm	23	12	25	12	20
71	Above 0 degrees, because the snow melts	2	6	5	4	6
72	Above 0 degrees. No explanation	2	4	2	4	2
79	Other incorrect	11	12	13	14	17
Blank		12	26	10	14	24

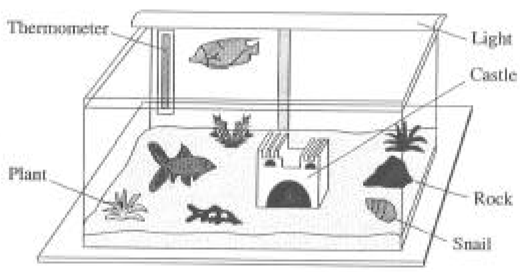
The effect of country clusters is very similar to that of the previous item. The between group-variance is about 40 percent; but, just as for the previous item, most of this is due to large between-group differences in the non-response category. East Europe and South Europe have significantly higher frequency of non-response, and this fact may indicate some cultural differences when it comes to the willingness (motivation) to answer. There are, however, some large differences between countries within the groups. For example in East Asia Hong Kong has 34 percent correct or partially correct while Japan has 74 percent; and, in North Europe, Denmark has 41 percent correct or partially correct while Iceland has 62 percent.

The most common incorrect response is “above 0 degrees because the hands are warm” (Code 70). North and East Europe have fewer responses in this category, a feature that may be explained by more experience with snow and ice in these countries. Also individual countries seem to support that pattern. The two “winter countries” Japan and Korea both have relatively low percentages on Code 70 (7 percent and 12 percent, respectively), whereas Singapore, with its very good overall results, has a percentage for Code 70 as high as 41. If we combine the three codes—70, 71, and 72—we find that, internationally, 24 percent of students seem to believe that the temperature will increase to above 0 degrees.

The role of light in maintaining the ecosystem (X2b)

The process of photosynthesis is fundamental for life on Earth. Much research has focused on students’ understanding of how plants grow and where their “food” comes from (e.g. Driver, et al. 1984; Wandersee 1983). There is also some research showing that light is not regarded as necessary for photosynthesis (Bell, 1985; Stavy, 1987). In TIMSS there are three items that, in different ways, focus on photosynthesis. Here we want to discuss the second part of item X2 (see question b in Figure 3) where the students were asked to explain the role of light in maintaining the ecosystem in the aquarium. The photosynthesis was not mentioned in the stem, and it is not required that they explicitly mention that concept to get the response correct.

X2. In the picture of an aquarium six items are labeled.



Explain why each of the following is important in maintaining the ecosystem in the aquarium.

(a) the plant

(b) the light

Figure 3. Item X2

From the results in Table 4 we can see that there are large differences between the groups; and, for this item, the country clustering can “explain” more than in any of the other items we discuss here: almost 60 percent of the variability of response patterns. The between-group component of the variance is pronounced in every response category except the unimportant Code 11. East Asia stands out as the best group. More than half of the students in East Asia responded correctly as opposed to the English-speaking group where only one-fifth of the students' responses were correct. In spite of the large between-group effect, there are large differences within the East Asia group, correct responses varying from 78 percent in Singapore to 26 percent in Hong Kong.

It is interesting to see which concepts the students use in their responses. From Table 4 we can see that almost all of those who have correct answers in the East Asia countries referred to photosynthesis explicitly: in Singapore as many as 70 percent. Examples of student answers in this category are: "to help the plant making photosynthesis" and "the light provides energy for the plants to make food." In English-speaking countries and in the North Europe group we see the opposite. In New Zealand and Norway, for example, photosynthesis is referred to in only one fourth of the correct responses. Many students in these countries gave answers such as "the light makes the plants grow" (Code 19).

Table 4. Coding guide and results for item X2B

Code	*Response*	*East Asia*	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
Correct Response						
10	Refers to photosynthesis	44	22	9	10	18
11	States that the light provides energy without any further detail	3	3	2	3	2
19	Other correct	6	14	11	18	14
Incorrect Response						
70	Says the fish need the light to see	8	9	22	12	16
71	Says we need light to see	4	1	5	2	5
72	Says light provides warmth for the fish	6	6	20	12	6
79	Other incorrect	18	17	21	27	19
Blank		11	27	12	18	22

Some of the incorrect responses (e.g. "So we can see the fish," Code 71) could arguably have been correct if the question had been why we have light in an aquarium. There are, however, very few responses in this category. Many students state that the light is needed so the fish will have "a good time," "so the fish can see," Code 70, or "the light provides warmth to the fish," Code 72). Both these categories say something about the conditions for the fish, and almost half of the students in English-speaking countries answer in this way. It seems to be a pronounced cultural effect that caring for pets plays a particularly important role in these countries.

In Part A of this item the students are asked to explain why the plant is important in maintaining the ecosystem in the aquarium. The results show that this aspect of photosynthesis is better known. Internationally, 43 percent answered that the plant produces oxygen, and, for example in countries like New Zealand and Norway, where very few students referred to photosynthesis in question b, half of the students answered that plants produce oxygen. It seems that students are much more informed about the substances taking part in photosynthesis than in the role of light as the energy source for the process.

Gravity (K17)

The next item (see Figure 4) is about an apple falling to the ground due to gravity, a concept that is far from easy. Gravity acts on the apple whether it is falling or not, regardless of position and movement. However, when the apple is resting on the ground there is also a force from the ground pointing upwards. The force of gravity and this force from the ground have the same size, and the apple is at rest because the sum of forces is zero. If we neglect the air resistance, only gravity acts on the apple when it is falling.

Students' understanding of mechanics is probably the domain within physics that is most frequently explored by science educators. Many such studies show that students in all age groups all over the world have unsatisfactory understanding of simple but crucial topics of mechanics. And often, the students' ideas and concepts have been shown to be surprisingly similar in many countries (Duit & Treagust, 1995). Looking at Table 5 we see that this item shows remarkably good results. Perhaps all the attention connected to this research has had a positive influence on teaching and learning around the world.

K17. The drawing shows an apple falling to the ground. In which of the three positions does gravity act on the apple?

A. 2 only
B. 1 and 2 only
C. 1 and 3 only
D. 1. 2. and 3

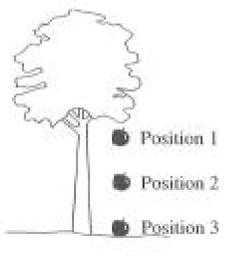


Figure 4. Item K17

The response patterns are quite similar across groups of countries, the between group part of the pooled variance amounting to only about 25 percent. More than 60 percent correct responses in East Asia represents a very satisfactory result, and all the countries within this group score above the international average. However, there are large differences between some countries within groups, East and South Europe in particular. For example in East Europe, the Czech Republic has 81percent correct while Russia has only 42 percent correct and 41 percent for distractor B. Furthermore, in South Europe Spain has 55 percent correct compared to Greece with 30 percent correct and 49 percent for distractor B.

Table 5. Results for item K17

	East Asia	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
A	4	7	6	10	8
B	25	25	30	27	36
C	6	7	8	8	11
D*	64	57	56	53	44
Blank	0	3	1	1	1

The most common incorrect response is B, which indicates the belief that gravity acts only when the apple is falling. Distractor A most likely indicates the same belief. (It is however, not entirely clear if the apple in Position 1 is moving or not). The international mean response frequency is 34 percent for these two alternatives together. The conception that gravity acts only when the effect is observable, is known to be a common misconception. As diSessa (1993) says, according to many students' views only the movement needs to be explained, and gravity provides this explanation in the form of a "moving force." Forces are therefore comprehended as cause for movement and not as cause for changes of the movement (Sjøberg & Lie 1981).

On the other hand, those students who choose C (Figure 4) may think that forces only act when the apple is at rest and no forces act on the falling apple in Position 2. Something similar appears when we are talking about weightlessness. When an astronaut is so-called "weightless," gravity is acting on both the spaceship and the astronaut. There are however, no forces acting between them, and therefore the astronaut has this feeling of weightlessness. It is when we are at rest on the ground we "feel" our weight. This may explain some students' response to this item. In some countries almost 20 percent chose C, e.g. Iceland, Latvia, and South Africa.

Water balance and temperature regulation (O16)

Regulation of the human body's temperature and balance for total body water are vital processes in human physiology. Water is crucial for the internal environment. The two sources of body water are metabolically produced water and ingestion of water. Perspiration is controlled by a mechanism directed toward temperature regulation. The mechanism of thirst is certainly of great importance, since water for evaporation from the skin by the production of sweat, can be made up only by ingestion of water. It is, however, also true that our fluid intake is often influenced more by habit and sociological factors than by the need to regulate the body water. Item O16 (see Figure 5) assesses students' understanding of these processes in human physiology.

O16. Write down the reason why we get thirsty on a hot day and have to drink a lot.

Figure 5. Item O16

The country clustering can “explain” about 45 percent of the variability of response patterns for this item, with group effects in all response categories. From the results in Table 6 we see that there are large differences between East Asia (with above 80 percent correct responses) and the other groups of countries. Correct responses are divided in three categories. To get Code 10 or 12 the students had to refer to perspiration and also its cooling effect. In all groups of countries there are few students that explicitly refer to the cooling effect. The very frequent simple reference to the perspiration of water was coded 11 or 13. The combination of these pairs of categories have been done here due to the difficulty and unimportance of distinguishing between whether replacement of water is implicitly mentioned (Codes 10 or 11) or not (Codes 12 or 13).

Table 6. Coding guide and results for item O16

Code	*Response*	*East Asia*	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
Correct Response						
10+12	Refers to perspiration and its cooling effect	10	12	8	3	4
11+13	Refers to perspiration	68	40	50	62	51
19	Other acceptable explanation	5	3	3	3	2
Incorrect Response						
70	Refers to body temperature (being too hot) but does not answer why we get thirsty	2	4	5	3	5
71	Refers only to drying of the body	4	11	22	11	8
72	Refers to getting more energy by drinking more water	2	4	1	3	6
79	Other incorrect	5	12	6	9	14
Blank		4	16	4	5	12

It seems that the importance of perspiration for the cooling effect is not a central part of teaching in many countries, except in Hungary where 42 percent of the students referred to this effect. In Kuwait and Korea they also had above 20 percent. It was discussed whether Code 10 represented a better answer than the other responses and therefore "deserved" 2 points. From a psychometric point of view this can be supported by an analysis of the Norwegian data that has shown that students in this category have higher overall score than the two other categories. Furthermore, a similar pattern appeared in the results for those in Population 3 (the item was a “link item,” i.e. given in all three populations). The references to the cooling effect dramatically increase in frequency from Population 2 to 3 and from the vocational to the academic branch of upper secondary school (Angell, Kjærnsli, & Lie, 2000). However, a closer look at the item itself reveals that the students are asked to write down *the reason why* we get thirsty, thus implicitly asking for *the one* reason. A response that also refers to the *function* of perspiration, namely temperature regulation of the human body, is definitely a more advanced response, but it cannot reasonably be given a higher score. However, this example also gives another demonstration of the power and flexibility of the coding system. When performing a diagnostic analysis, the codes can be compared and combined according to the main issues under consideration.

From Table 6 we see that most of the incorrect responses have been coded as 71, responses such as "Your throat gets dry." Among individual countries, France stands out with as many as 41 percent in this category, and of the next seven countries, six of them are English-speaking.

Code 72 represents an interesting misconception. These students seem to believe that you have to drink because you get exhausted, and that you get more energy by drinking water. This misconception is more common than seen from the response distribution, simply because a number of responses included sweating in addition *to* this wrong statement and therefore were scored as correct.

Day and night (Q11)

This item (Figure 6) deals with why we have day and night on Earth. The correct answer is of course that the Earth rotates around its own axis (A).

Q11. Which statement explains why daylight and darkness occur on Earth?

A. The Earth rotates on its axis.

B. The Sun rotates on its axis.

C. The Earth's axis is tilted.

D. The Earth revolves around the Sun.

Figure 6. Item Q11

It is common to think of the Sun at rest in the center of the solar system, and all the planets including the Earth orbiting around the Sun. In addition, each planet is rotating around its own axis, so the whole movement is quite complicated. Maybe this complex structure to some extent could explain why as many as 42 percent of the students choose D. There are relatively few students from the various countries who chose the two other distractors (see response distribution in Table 7). Looking at the response alternatives, we see that all of them as isolated statements are true, even if they do not represent correct answers to the question at hand. This fact may well offer a particular challenge to the students.

Table 7. Results for item Q11

	East Asia	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
A*	59	47	36	46	45
B	4	4	5	7	5
C	4	4	8	4	4
D	32	40	49	41	42
Blank	1	6	2	2	5

East Asia clearly stands out as the best group on this item with an average of 59 percent correct responses and with all its countries appearing above the international average (44 percent). The English-speaking countries are at the bottom with 36 percent correct answers. In this group the US is on top with 55 percent correct responses and Scotland at the bottom with 21 percent, which means large differences between countries within this group. Also in North Europe there are large differences between countries (Netherlands 62 percent, Iceland 21 percent). The between-group part of the total variability is 30 percent for this item, but the major part of this is provided by the difference between East Asia and the English speaking countries. The response patterns for the remaining groups of countries are strikingly similar.

Astronomy appears in science curricula in many countries around world, often at the primary level to capitalize on children's interest and enthusiasm in this area and to enrich their experiences of science (Sharp, Bowker, Mooney, Grace, & Jeans, 1999). In spite of this, it is remarkable that so few students know why it is day or night, which definitely is an everyday (and –night) phenomenon! It is worthwhile comparing this finding with the curriculum in the countries participating in TIMSS 2003. A large majority of countries indicated in a questionnaire that the explanation of day and night is taught even at Grade 4. It is easy to underestimate the difficulty for children of understanding the models we have developed for the solar system and which we use to explain day and night, and also the seasons. Most of the models bear no instant relation to any concrete observation the children can make. For example, the two-dimensional figures in many textbooks are difficult to understand if the students are not showed a three-dimensional model as well.

Surface of the Earth (J1)

The last item (see Figure 7) addresses the fundamental aspect of what we may call the Earth's ever-changing surface. In the 19th century there were active debates before the "old Earth" geological paradigm was settled: i.e. that the Earth is at least many millions of years old. Only by accepting this large time scale could the observed slow building and breaking geological processes be argued as providing an explanation for the surface features. The fact that the 20th century has stretched the time span even further to some billions of years, mainly by astronomical evidence, is of minor importance from our point of view here. The crucial point is that what we observe as very insignificant geological processes in daily life or during a human life span, in a geological time scale can (and do so continuously) lift mountains and break them down.

The results are exhibited in Table 8. The four response alternatives may be characterized as the “building” (A), the “breaking” (B), the correct “building and breaking” (C), and the “static” (D) responses. There is a striking similarity between responses from the country groups, and in particular there are about 40 percent correct responses in all of them. The percent of the variability of response patterns that can be accounted for by the grouping of countries is only about 20 percent, the lowest for all items under consideration.

J1. Which BEST describes the surface of the Earth over billions of years?

A. A flat surface is gradually pushed up into higher and higher mountains until the Earth is covered with mountains.

B. High mountains gradually wear down until most of the Earth is at sea level.

C. High mountains gradually wear down as new mountains are continuously being formed, over and over again.

D. High mountains and flat plains stay side by side for billions of years with little change.

Figure 7. Item J1

What mainly concerns us here is the large portion of students responding D. This response strongly indicates that the student really does have the view of a static Earth. The format of this item requires each student to read through all response alternatives and select the “BEST” explanation. Most probably, since D comes last, all three “active Earth” responses have “actively” been rejected.

Table 8. Results for item J1

	East Asia	*East Europe*	*English-speaking*	*North Europe*	*South Europe*
A	18	9	10	10	8
B	17	13	18	14	20
C *	49	41	42	43	38
D	16	33	28	30	31
Blank	1	4	1	3	2

It should be a concern among science educators in many countries that the basic geological paradigm seems to have been understood by so few by the end of compulsory schooling. Even if differences between country groups are small, there are some interesting and very surprising results for individual countries. One might suspect that countries with earthquakes, volcanoes, or even high mountains with glaciers would be more likely to focus more, in the media and in school, on geological processes. However, the results are not better, and in particular there are not fewer "static Earth believers", in countries like Switzerland (31 percent) or Iceland (44 percent). On the other hand, there are particularly few students in this category in Japan (8 percent) and Korea (3 percent). Korea also stands out as having a very high percentage of correct responses (76 percent). We conclude that neither cultural nor geological factors can explain much of the variability between countries in understanding the phenomenon discussed here.

CONCLUSION

The first aim of the item-by-item analyses in this chapter has been to explore students' conceptions of some fundamental ideas in science. The overall results for each selected item show striking similarities in the general response patterns. As a rough guide we may say that something like 25 to 40 percent of the students seem to have demonstrated adequate understanding of the phenomena discussed here. Furthermore, the analyses have revealed some very distinct misconceptions that are common around the world. We can describe these findings by summing up the misconceptions of some international "typical student," representing a considerable percentage of students at age 13 around the world:

- Day and night occur because the Earth orbits the Sun.
- Earth's surface remains the same over billions of years.
- During the process of melting, ice increases its temperature above 0^oC, and the mass is increasing because water is "heavier" than ice.
- Gravity acts only when we can observe the effect, e.g. when an apple is falling and not when it is at rest.
- Even if the need to drink to replace lost water due to perspiration is well known, the cooling effect on the body by perspiration is not recognized.
- The main role of light in an aquarium as an ecosystem is to care for the fish, for visibility, or to maintain the temperature.

The second aim has been to study differences and similarities in response patterns between countries. Country groups were formed according to cultural/geographical regions and similarities of patterns of correct responses for all science items for Population 2. When we consider the total patterns of all response categories for individual items, we find that the grouping of countries can "explain" between 20 percent and 60 percent of the total variability. For some items (particularly J1-Surface of the Earth, and K17-Gravity) the between-group differences are rather small, whereas for other items (X2b-The role of light, and O16-Water balance) there seems to be a pronounced cluster effect. It is difficult to see any clear reason for this different behavior among items, partly because the role and number of response categories varies from item to item. Obviously, there are curricular impacts within some areas, accounting for significant differences between countries that cannot be understood by cultural or geographical factors alone. In order to investigate the differences in student response patterns across countries further, it would be necessary to carry out detailed curricular comparisons between countries for the topics discussed here.

REFERENCES

Angell, C., Kjærnsli, M., & Lie, S. (2000). Exploring students' responses on free-response science items in TIMSS. In D. Shorrocks-Taylor & E. W. Jenkins (Eds). *Learning from others. International comparisons in education* (pp. 159 – 188). Science & Technology Education Library volume 8. Dordrecht: Kluwer Academic Publishers.

Beaton, A., Martin, M.O., Mullis I.V.S., Gonzales, E.J., Smith, T.A., & Kelly, D.A. (1996). *Science achievement in the middle school years. IEA's Third International Mathematics and Science Study (TIMSS).* Chestnut Hill, MA: Boston College.

Bell, B. (1985). Students' ideas about plant nutrition: What are they? *Journal of Biological Education, 19*(3), 213 – 218.

Driver, R., Child, D., Gott, R., Head, J., Johnson, S., Worsley, C. and Wylie, F. (1984). *Science in Schools. Age 15: Report No.2*. *Report on the 1981 APU survey in England, Wales and Nothern Ireland.* London: Assessment of Performance Unit, Department of Education and Science, HMSO.

Driver, R. & Easley, J. (1978). Pupils and paradigms: A review literature related to concept development in adolescent Science students. *Studies in Science Education*, *5,* 61-83.

Duit, R. & Treagust, D. F. (1995). Students' conceptions and constructivist teaching approaches. In B. J. Fraser and H. J. Walberg (Eds.), *Improving Science education*. Chicago, IL:The University of Chicago Press.

Erickson, G. & Tiberghien, A. (1985). Heat and temperature. In R. Driver, E. Guesne & A. Tiberghien (Eds.), *Children's ideas in Science*. UK: Open University Press.

Lie, S., Taylor, A., & Harmon, M. (1996). Scoring techniques and criteria. In M. O. Martin & D. Kelly (Eds). *Third International Mathematics and Science Study technical report. Volume 1: Design and development.* Boston College.

Mullis, I.V.S. & Smith, T.A. (1996). Quality control steps for free-response scoring*.* In M. O. Martin & I.V.S. Mullis (Eds), *Third International Mathematics and Science Study: Quality assurance in data collection*. Chestnut Hill, MA: Boston College.

Pfundt, H. & Duit, R. (1994). *Bibliography: Students' alternative frameworks and Science education* (4th ed.). Germany: IPN at the University of Kiel.

Robitaille, D.F., Schmidt, W.H., Raizen, S., Mc Knight, C., Britton, E., & Nicol, C. (1993). *Curriculum frameworks for Mathematics and Science. TIMSS monograph no. 1.* Vancouver, Canada :Pacific Educational Press.

Schmidt, W.H, Jorde, D., et al. (1996). *Characterizing pedagogical flow. An investigation of Mathematics and Science teaching in six countries*. Dordrecht: Kluwer Academic Publishers.

diSessa, A. A. (1993). Toward an epistemology of Physics*. Cognition and Instruction, 10* (2 & 3), 105-225.

Sharp, J. G., Bowker, R., Mooney, C. M., Grace, M., & Jeans, R. (1999). Teaching and learning astronomy in primary school. *School Science Review*, *80*(292), 75 – 86.

Sjøberg, S. & Lie, S. (1981). *Ideas about force and movement among Norwegian pupils and students. Report 81-11*. Oslo: University of Oslo.

Stavy, R. (1987). How students' aged 13 – 15 understand photosynthesis. *International Journal of Science Education*, *9*(1), 105 – 115.

Third International Mathematics and Science Study (1995). *Coding guide for free-response items-Populations 1 and 2.* TIMSS Doc. Ref.: ICC897/NRC433. Chestnut Hill, MA: Boston College.

Thomaz, M. F., Valente, M. C. & Antunes, M. J. (1995). An attempt to overcome alternative conceptions related to heat and temperature. *Physics Education, 30*(1), 19 – 26.

Wandersee, J. H. (1983). Students' misconceptions about photosynthesis: A cross-age study. In H. Helm & D. Novak (Chairs), *Proceedings of the International Seminar about Misconceptions in Science and Mathematics* (pp. 441-466). Ithaca, NY: Cornell University.

Wandersee, J. H., Mintzes J. J. & Novak, J. D. (1993). Research on alternative conceptions in science. In D. Gabel (Ed). *Handbook of research on science teaching and learning*. New York: Macmillan Publishing.

Zabulionis, A. (1997). *The formal similarity of the mathematics and science achievement of the countries. Vilnius University pre-print No 97-19*. Vilnius: Vilnius University.