

# **PISA and "Real Life Challenges": Mission Impossible?**

*Svein Sjøberg University of Oslo < svein.sjoberg@ils.uio.no >*

*Contribution to Hopman (Ed): PISA according to PISA Revised Version Oct 8 2007*

## ***Introduction***

The PISA project has positive as well as more problematic aspects, and it is important for educators and researchers to engage in critical public debates on this utterly important project, including its uses and misuses.

The PISA project sets the educational agenda internationally as well as within the participating countries. PISA results and advice are often considered as objective and value-free scientific truths, while they are, in fact embedded in the overall political and economic aims and priorities of the OECD. Through media coverage PISA results create the public perception of the quality of a country's overall school system. The lack of critical voices from academics as well as from media gives authority to the images that are presented.

In this article, I will raise critical points from several perspectives. The main point of view is that the PISA ambitions of testing "real-life skills and competencies in authentic contexts" are by definition alone impossible to achieve. A test is never better than the items that constitute the test. Hence, a critique of PISA should not mainly address the official rationale, ambitions and definitions, but should scrutinize the test items and the realities around the data collection. The secrecy over PISA items makes detailed critique difficult, but I will illustrate the quality of the items with two examples from the released texts.

Finally, I will raise serious questions about the credibility of the results, in particular the ranking. Reliable results assume that the respondents in all countries do their best while they are sitting the test. I will assert that young learners in different countries and cultures may vary in the way they behave in the PISA test situation. I claim that in many modern societies, several students are unwilling to give their best performance if they find the PISA items long, unreadable, unrealistic and boring, in particular if bad test results have no negative consequence for them. I will use the concept of "perceived task value" to argue this important point.

## ***The political importance of PISA***

Whether one likes the PISA study or not, one might easily agree about the importance of the project. When OECD has embarked on such a large project, it is certainly not meant as a purely academic research undertaking. PISA is meant to provide results to be used in the shaping of future policies. After 6-7 years of living with PISA, we see that the PISA concepts, ideology, values and not least the results and the rankings, shape international educational policies and also influence national policies in most of the participating countries. Moreover, the PISA results provide media and the public with convincing images and perceptions about the quality of the school system, the quality of their teachers' work and the characteristics of both the school population and future citizen.

Contemporary natural science is often labelled Big Science or Techno-science: The projects are multinational, they involve thousands of researchers, and they require heavy funding. Moreover, the traditional scientific values and ethos of science become different from the traditional ideals of academic science (Ziman, 2000). Prime examples are CERN, The Human

Genome Project, European Space Agency etc. The PISA project has many similarities with such projects, although the scale and the costs are much lower. But the number of people who are involved is large, and the mere organization of the undertaking requires resources, planning and logistics unusual to the social sciences. According to Prais (2007), the total cost of the PISA and TIMSS testing in 2006 was "probably well over 100 million US dollars for all countries together, plus the time of pupils and teachers directly involved."

Why is an organization like the OECD embarking on an ambitious task like this? The OECD is an organization for the promotion of economic growth, cooperation and development in countries that are committed to market economies. Their slogan appears on their website: "For a better world economy."<sup>1</sup>

The OECD and its member countries have not embarked on the PISA project because they have an interest in basic research in education or learning theory. They have decided to invest in PISA because education is crucial for the economy. Governments need information that is supposed to be relevant for their policies and priorities in this economic perspective. Since mass education is expensive, they also most certainly want "value for money" to ensure efficient running of the educational systems. Stating this is not meant as a critique of PISA. It is, however, meant to state the obvious, but still important fact: PISA should be judged in the context of the agenda of the OECD; economic development and competition in a global market economy.

The strong influence that PISA has on national educational policies should imply that all educators ought to be interested in PISA whether they endorse the aims of PISA or not. Educators should be able to discuss and use the results with some insight in the methods, underlying assumptions, strengths and weaknesses, possibilities and limitations of the project. We need to know what we might learn from the study, as well as what we cannot learn. Moreover, we need to raise a critical (not necessarily negative!) voice in the public as well as professional debates over uses and misuses of the results.

### ***The influence of PISA: Norway as an example***

The attention given to PISA results in national media varies between countries, but in most countries it is formidable. In my country, Norway, the results from PISA2000 as well as from PISA2003 provided war-like headings in most national newspapers.

Our then Minister of Education (2001-2005), Kristin Clemet (representing Høyre, the Conservative party), commented on the PISA2000 results, released a few months after she had taken office, following a Labour government: "Norway is a school loser, now it is well documented. It is like coming home from the Winter Olympics without a gold medal" (which, of course, for Norway would have been a most unthinkable disaster!). She even added: "And this time we cannot even claim that the Finnish participants have been doped!" (Aftenposten January 2001).

The headlines in all the newspapers told us again and again "Norway is a loser".

---

<sup>1</sup> These and other quotes in the article are taken from OECDs home site <http://www.oecd.org/>, retrieved Sept 2, 2007.



**Figure 1**  
*PISA results are presented in the media with war-like headings, shaping public perception about the national school system. Here are PISA results presented in the leading Norwegian newspaper Dagbladet with the heading "Norway is a school loser".*

In fact, such headings were misleading. Norway has ended up close to the average among the OECD countries in all test domains in PISA2000 and PISA2003. But for some reason, Norwegians had expected that we should be on top – as we often are on other indicators and in winter sports. When we are not the winners, we regard ourselves as being losers.

The results from PISA (and TIMSS as well) have shaped the public image of the quality of our school system, not only for the aspects that have in fact been studied, but for more or less all other aspects of school. It has now become commonly 'accepted' that Norwegian schools in general have a very low level of quality, and that Norwegian classrooms are among the noisiest in the world. The media present tabloid-like and oversimplified rankings. It seems that the public as well as politicians have accepted these versions as objective scientific truths about our education system. There has been little public debate, and even the researchers behind the PISA study have little to modify and remind the public about the limitations of the study. In sum; PISA (as well as TIMSS) has created a public image of the quality of the Norwegian school that is not justified, and that may be seen to be detrimental. I assume that other countries may have similar experiences.

But PISA does not only shape the public image, it also provides a scientific legitimization of school reforms. Under Kristin Clemet as Minister of Education (2001-2005), a series of educational reforms were introduced in Norway. Most of these reforms were legitimized by

reference to international testing, mainly to PISA. In 2005, we had a change in government, and Kristin Clemet's Secretary of State, Helge Ole Bergesen, published a book shortly afterwards in which he presented the "inside story" on the reforms made while they were in power. The main perspective of the book is the many references to large-scale achievement studies. He confirms that these studies provided the key arguments and rationale for curricular as well as other school reforms. Under the tabloid heading: "**The PISA Shock**", he confirms the key role of PISA:

With the [publication of the] PISA results, the scene was set for a national battle over knowledge in our schools. . [...] For those of us who had just taken over the political power in the Ministry of Education and Research, the PISA results provided a "flying start" (Bergesen 2006: 41-42. Author's translation).

Other countries may have different stories to tell. Figures 2 and 3 provide examples from the public sphere in Germany. In sum: There is no doubt that PISA has provided - and will continue to provide - results, ideologies, concepts, analysis, advice and recommendations that will shape our future educational debates and reforms, nationally as well as internationally.



**Figure 2**

*The political agenda and the public image of the quality of the entire school system is formed by the PISA results. This is an example from the German Newspaper Die Woche after the release of PISA2000 results.*



**Figure 3**

*PISA has become a well-known concept in public debate: A bookshelf in a German airport offering bestselling books with PISA-like tests for self-assessment, very much like IQ-tests.*

### ***PISA: Underlying values and assumptions***

It is important to examine the ideas and values that underpin PISA, because, like most research studies, PISA is not impartial. It builds on several assumptions, and it carries with it several value judgements. Some of these values are explicit; others are implicit and 'hidden', but nevertheless of great importance. Some value commitments are not very controversial, others may be contested.

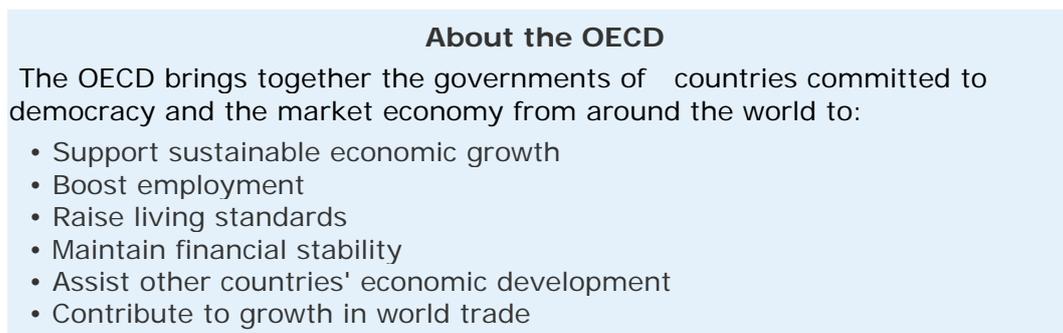
Peter Fensham, a key scholar in international science education thinking and research for many decades, has also been heavily involved in several committees in TIMSS and PISA. He has seen all aspects of the projects from the inside over decades. In a recent book chapter, he provides an insider's overview and critique of the values that underlie these projects. He draws attention to the underlying values and implications:

The design and findings from large-scale international comparisons of science learning do impact on how science education is thought about, is taught and is assessed in the

participating countries. The design and the development of the instruments used and the findings that they produce send implicit messages to the curriculum authorities and, through them, to science teachers. It is thus important that these values, at all levels of existence and operations of the projects, be discussed, lest these messages act counterproductively to other sets of values that individual countries try to achieve with their science curricula. (Fensham 2007: 215,216)

## **Aims and purpose of the OECD**

In the public debate as well as among politicians, advice and reports from OECD experts are often considered to be impartial and objective. The OECD has become an important contributor to the political battle over social, political, economic and other ideas. To a large extent, these persons shape the political landscape, and their reports and advice set the political agenda in the national as well as international debates over priorities and concerns. But the OECD is certainly not a impartial group of independent educational researchers. The OECD is built on a neo-liberal political and economic ideology, and its advice should be seen in this perspective. The seemingly scientific and neutral language of expert advice conceals the fact that there are possibilities for other political choices based on different sets of social, cultural and educational values. Figure 4 shows how the OECD presents itself.



### **Figure 4**

*The basis for and commitments of the OECD as they appear on <http://www.oecd.org/>  
Retrieved Sept 7 2007*

The overall perspective of the OECD is concerned with market economy and growth in free world trade. All policy advice they provide is certainly coloured by such underlying value commitments. Hence, the agenda of the OECD (and PISA) does not necessarily coincide with the concerns of many educators (or other citizens, for that matter). The concerns of PISA are not about 'Bildung' or liberal education, not about solidarity with the poor, not about sustainable development etc. - but about skills and competencies that can promote the economic goals of the OECD. Saying this is, of course, stating the obvious, but such basic facts are often forgotten in the public and political debates over PISA results.

## **Educational and curricular values in PISA**

Quite naturally, values creep into PISA testing in several ways. PISA sets out to shed light on important (and not very controversial) questions like these:

Are students well prepared for future challenges? Can they analyse, reason and communicate effectively? Do they have the capacity to continue learning throughout life? (First words on the PISA home page at <http://www.pisa.oecd.org/>)

These are important concerns for most people, and it is hard to disagree with such aims. However, as is well known, PISA tests just a few areas of the school curriculum: Reading, mathematics and science. These subjects are, consequently, considered more important than other areas of the school curriculum in order to reach the brave goals quoted above. Hence, the OECD implicitly says that our future challenges are *not* highly dependent on subjects like history, geography, social science, ethics, foreign language, practical skills, arts and aesthetics, etc.

PISA provides test results that are closely connected to (certain aspects of) the three subjects that they test. But when test results are communicated to the public, one receives the impression that they have tested the quality of the entire school system and all the competencies that are of key importance for preparing to meet the challenges of the future.

There is one important feature of PISA that is often forgotten in the public debate: PISA (in contrast to TIMSS) does *not* test "school knowledge". Neither the PISA framework nor the test items claim having any connection to national school curricula. This fact is in many ways the strength of the PISA undertaking; they have set out to think independently from the constraints of all the different school curricula. There is a strong contrast with the TIMSS test, as its items are meant to test knowledge that is more or less common in all curricula in the numerous participating countries. This implies, of course, that the "TIMSS curriculum" (Mullis et al 2001) may be characterized as a fossilized and old-fashioned curriculum of a type that most science educators want to eradicate. In fact, nearly all TIMSS test items could have been used 60-70 years ago. The PISA thinking has been freed from the constraints of school curricula and could in principle be more radical and forward-looking in their thinking. (However, as asserted in other parts of this chapter, PISA does not manage to live up to such high expectations.)

PISA stresses that the skills and competencies assessed may not only stem from activities at school but from experiences and influences from family life, contact with friends, etc. In spite of this, both good and bad results are most often considered by both the public and politicians to be attributed to the school only.

### **Values in the PISA-reporting**

The PISA data collection also covers a great variety of dimensions regarding background variables. The intention is, of course, to use these to explain the variance in the test results ("explain" in a statistical sense, i.e. to establish correlations etc.). Many interesting studies have been published on such issues. But the main focus in the public reporting is in the form of simple ranking, often in the form of league tables for the participating countries. Here, the mean scores of the national samples in different countries are published. These league tables are nearly the only results that appear in the mass media. Although the PISA researchers take care to explain that many differences (say, between a mean national score of 567 and 572) are not statistically significant, the placement on the list gets most of the public attention. It is somewhat similar to sporting events: The winner takes it all. If you become no 8, no one asks how far you are from the winner, or how far you are from no 24 at any event. Moving up or down some places in this league table from PISA2000 to PISA2003 is awarded great importance in the public debate, although the differences may be non-significant statistically as well as educationally.

The winners also become models and ideals for other countries. Many want to copy aspects of the school system from the winners. This, among other things, assumes that PISA results can

be explained mainly by school factors – and not by political, historical, economic or cultural factors or by youth culture and the values and concerns of the young learners. Peter Fensham claims:

... the project managers choose to have quite separate expert groups to work on the science learning and the contextual factors—a decision that was later to lead to discrepancies. Both projects have taken a positivist stance to the relationship between contextual constructs and students' achievement scores, although after the first round of TIMSS other voices suggested a more holistic or cultural approach to be more appropriate for such multi-cultural comparisons. (Fensham 2007: 218)

PISA (and even more so TIMSS) is dominated and driven by psychometric concerns, and much less by educational. The data that emerge from these studies provides a fantastic pool of social and educational data, collected under strictly controlled conditions – a playground for psychometricians and their models. In fact, the rather complicated statistical design of the studies decreases the intelligibility of the studies. It is, even for experts, rather difficult to understand the statistical and sampling procedures, the rationale and the models that underlie the emergence of even test scores. In practice, one has to take the results at face value and on trust, given that some of our best statisticians are involved. But the advanced statistics certainly reduce the transparency of the study and hinder publicly informed debate.

## ***PISA items – a critique***

### **The secrecy**

An achievement test is never better than the quality of its items. If the items are miserable, even the best statisticians in the world cannot change this fact. Subject matter educators should have a particular interest, and even a duty, to go into detail on how their subject is treated and 'operationalized' through the PISA test items. One should not just discuss the given definitions of e.g. scientific literacy and the *intentions* of what PISA claims to test. In fact, the framework as well as the intentions and ideologies in the PISA testing may be considered acceptable and even progressive. The important question is: How are these brave intentions translated into actual items?

But it is not easy to address this important issue, as only a very few of the items have been made publicly available. Peter Fensham, himself a member of the PISA (as well as TIMSS) subject matter expert group, deplores the secrecy:

By their decision to maintain most items in a test secret [...] TIMSS and PISA deny to curriculum authorities and to teachers the most immediate feedback the project could make, namely the release in detail of the items, that would indicate better than framework statements, what is meant by 'science learning'. The released items are tantalizing few and can easily be misinterpreted" (Fensham 2007: 217)

The reason for this secrecy is, of course, that the items will be used in the next PISA testing round, and therefore they may not be made public. An informed public debate on this key issue is therefore difficult, to say the least. But we scrutinize the relatively few items that have been made public<sup>2</sup>.

---

<sup>2</sup> All the released items from previous PISA rounds can be retrieved from the PISA website [http://www.oecd.org/document/25/0,3343,en\\_32252351\\_32235731\\_38709529\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/25/0,3343,en_32252351_32235731_38709529_1_1_1_1,00.html)

## Can "real-life challenges" assessed by wordy paper-and-pencil items?

The PISA testing takes place in about 60 countries, which together (according to the PISA homepage) account for 90% of the world economy. PISA has the intention of testing

... knowledge and skills that are essential for full participation in society. [...] not merely in terms of mastery of the school curriculum, but in terms of important knowledge and skills needed in adult life. [...]

The questions are reviewed by the international contractor and by participating countries and are carefully checked for cultural bias. Only those questions that are unanimously approved are used in PISA. (Quotes from [Pisa.oecd.org](http://Pisa.oecd.org), retrieved 5 sept 2007)

In each item unit, the questions are based on what is called an "authentic text". This, one may assume, means that the original text has appeared in print in one of the 60 participating countries, and that it has been translated from this original.

There are many critical comments that can be made to challenge the claim that PISA lives up to the high ambition of testing real-life skills. An obvious limitation is the test format itself: The test contains only paper-and-pencil items, and most items are based on the reading of rather lengthy pieces of text. This is, of course, only a subset of the types of "knowledge and skills that are essential for full participation in society". Coping with life in modern societies requires a range of competencies and skills that cannot possibly be measured by test items of the PISA units' format.

## Identical "real-life challenges" in 60 countries?

But the abovementioned criticism has other and equally important dimensions: The PISA test items are by necessity exactly the same in each country. The quote above assures us that any "cultural bias" has been removed, and items have to be "unanimously approved".

At first glance, this sounds positive. But there are indeed difficulties with such requirements: Real life is different in different countries. Here are, in alphabetical order, the first countries on the list of participating countries: Argentina\*, Australia, Austria, Azerbaijan\*, Belgium, Brazil\*, Bulgaria\*, Canada, Chile\*, Colombia\*, Croatia\*, the Czech Republic, Denmark<sup>3</sup>

We can only imagine the deliberation towards unanimous acceptance of all items among the 60 countries with the demands that there should be no cultural bias and that context of no country should be favoured.

The following consequences seem unavoidable: The items will become decontextualised, or with contrived 'contexts' far removed from the reality of most real life situations in any of the participating countries. While the schools in most countries have a mandate to prepare students to meet the challenges in that particular society (depending on level of development, climate, natural environment, culture, urgent local and national needs and challenges, etc.), the PISA tests only aspects that are shared with *all* other nations. This runs contrary to current curriculum trends in many countries, where the issue of providing local relevance and context have become urgent. In many countries, educators argue for a more contextualized (or 'localized') curriculum, at least in the obligatory basic education for all young learners.

---

<sup>3</sup> The list is from <http://www.pisa.oecd.org/> Countries marked with a \* are not members of OECD, (but are also assumed to unanimously agree on the inclusion of all test units.)

The item construction process also rules out the inclusion of all sorts of controversial issues, be they scientific, cultural, economic or political. It is indeed enough that the authorities in *one* of the participating countries have objections.

To repeat: Schools in many countries have the mandate of preparing their learners to take an active part in social and political life. While many countries encourage the schools to treat controversial socio-scientific issues, such issues are unthinkable in schools in other countries. Moreover, a controversial issue in one country may not be seen as controversial in another. In sum: The demands of the item construction process set serious limitations on the actual items that comprise the PISA test.

Now, all the above considerations are simply deductions from the demands of the processes behind the construction of the PISA instrument. It is, of course, of great importance to check out such conclusions against the test itself. But, as mentioned, this is not an easy task to complete, given the secrecy over the test items. Nonetheless, the items that have been released confirm the above analysis: The PISA items are basically decontextualised and non-controversial. The PISA items are - in spite of an admirable level of ambition - nearly the negation of the skills and competencies that many educators consider important for facing future challenges in modern, democratic societies.

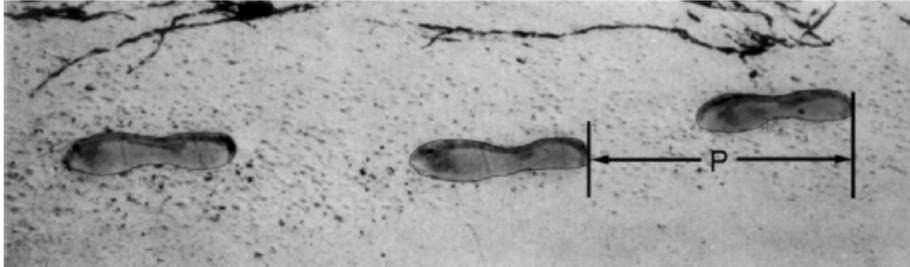
PISA items have also been criticized on other aspects. Many claim that the scientific content is questionable or misleading and that the language is strange, often verbose. In the next paragraph, two examples of PISA units are discussed in some detail, one from Mathematics, the other from Science.

## A PISA mathematics unit: Walking

In Figure 5 below the complete PISA test unit called Cloning is reproduced.

### M124: Walking

---



The picture shows the footprints of a man walking. The pacelength  $P$  is the distance between the rear of two consecutive footprints.

For men, the formula,  $\frac{n}{P} = 140$ , gives an approximate relationship between  $n$  and  $P$  where,

$n$  = number of steps per minute, and

$P$  = pacelength in metres.

#### Question 1: WALKING

M124Q01- 0 1 2 9

If the formula applies to Heiko's walking and Heiko takes 70 steps per minute, what is Heiko's pacelength? Show your work.

#### Question 3: WALKING

M124Q03- 00 11 21 22 23 24 31 99

Bernard knows his pacelength is 0.80 metres. The formula applies to Bernard's walking.

Calculate Bernard's walking speed in metres per minute and in kilometres per hour. Show your working out.

#### Figure 5

*A complete PISA mathematics unit, "Walking", with the text presenting the situation and the questions relating to the situation.*

#### Comments to Walking

Some details first:

Note that Question 2 is missing! (This may be an omission in the published document.) Note also the end of Q1: "Show your work." And for Q2. "Show your working out." There also seems to be several commas too many. Consider the commas in this paragraph: "For men, the formula,  $n/P = 140$ , gives an approximate relationship between  $n$  and  $P$  where, etc..." . In my view, all the 4 commas seem somewhat misplaced. Perhaps these are merely details, but they are not very convincing as the final outcome of serious negotiations between 60 countries!

The main comments to this unit are, however, more on the content of the item. First of all: Is this situation really a "real-life situation"? How real is the situation described above? Is this type of question a real challenge in the future life of young people – in any country?

But even if we accept the situation as a real problem, it seems hard to acknowledge that the given formula is a realistic mathematization of a genuine situation. The formula implies that when you increase the frequency in your walking, your paces simultaneously become longer.. To my knowledge, a person may walk with long paces and low frequency. Moreover, the same person may also walk using short steps at high frequency. In fact, at least from my point of view, the two factors should be inversely proportional rather than proportional, as suggested in the "Walking" item. In any case, a respondent who tries to *think critically* about the formula may get confused, but those who do not think may easily solve the question simply by inserting the formula.

But the problems do not stop here: Take a careful look at the dimensions given in the figure. If the marked footstep is 80 cm (as suggested in Q3 above), then the footprint is 55 cm long! A regular man's foot is actually only about 26 cm long, so the figure is extremely misleading! But even worse: From the figure, we can see (or measure) the next footstep to be 60% longer. Given the formula above, this also implies a more rapid pace, and the man's *acceleration* from the first to the second footstep has to be enormous!

In conclusion: The situation is unrealistic and flawed from several points of view. Students who simply insert numbers in the formula without thinking will get it right. More critical students who start thinking will, however, be confused and get in trouble!

## A PISA science unit: Cloning

In Figure 6 below the complete PISA test unit called Cloning is reproduced

### S128: Cloning

---

Read the newspaper article and answer the questions that follow.

#### A copying machine for living beings?

Without any doubt, if there had been elections for the animal of the year 1997, Dolly would have been the winner! Dolly is a Scottish sheep that you see in the photo. But Dolly is not just a simple sheep. She is a clone of another sheep. A clone means: a copy. Cloning means copying 'from a single master copy'. Scientists succeeded in creating a sheep (Dolly) that is identical to a sheep that functioned as a 'master copy'.

It was the Scottish scientist Ian Wilmut who designed the 'copying machine' for sheep. He took a very small piece from the udder of an adult sheep (sheep 1).

From that small piece he removed the nucleus, then he transferred the nucleus into the egg-cell of another (female) sheep (sheep 2). But first he removed from that egg-cell all the material that would have determined sheep 2 characteristics in a lamb produced from that egg-cell. Ian Wilmut implanted the manipulated egg-cell of sheep 2 into yet another (female) sheep (sheep 3). Sheep 3 became pregnant and had a lamb: Dolly.

Some scientists think that within a few years it will be possible to clone people as well. But many governments have already decided to forbid cloning of people by law.



**Question 1: CLONING**

Which sheep is Dolly identical to?

- A Sheep 1
- B Sheep 2
- C Sheep 3
- D Dolly's father

**Question 2: CLONING**

In line 14 the part of the udder that was used is described as "a very small piece". From the article text you can work out what is meant by "a very small piece".

That "very small piece" is

- A a cell.
- B a gene.
- C a cell nucleus.
- D a chromosome.

**Question 3: CLONING**

S128Q03

In the last sentence of the article it is stated that many governments have already decided to forbid cloning of people by law.

Two possible reasons for this decision are mentioned below.

Are these reasons scientific reasons?

Circle either "Yes" or "No" for each.

Reason:	Scientific?
Cloned people could be more sensitive to certain diseases than normal people.	Yes / No
People should not take over the role of a Creator.	Yes / No

**Figure 6**

A complete PISA science unit, "Cloning", with the text presenting the situation and the three questions relating to the situation.

**Comments to Cloning**

This task requires the understanding of the rather lengthy 30 lines of text. In non-English-speaking countries, this text is translated into the language of instruction. The translation follows rather detailed procedures to ensure high quality. The requirement that the text should be more or less identical results in rather strange prose in many languages. The original has, we assume, been an "authentic text" in some language, but the resulting translations cannot be considered to be "authentic" in the sense that they could appear in any newspaper or journal in that particular country.

PISA adheres to strict rules for the translation process, but this is not the way prose should be translated to become good, natural and readable in other languages. In my own language, Norwegian, the heading "A copying machine for living being" is translated word by word. This does not make sense, and prose like this would never appear in real texts.

The scientific content of the item may also be challenged. The only accepted answer on Question 1 is that Dolly is *identical* to Sheep 1 (alternative A). It may seem strange to claim that two sheep of very different ages are "identical" – but this is the only acceptable answer. The other two questions are also open for criticism. Basically, they test language skills, reading as well as vocabulary. (The word 'udder' was unknown to me.)

In conclusion: Although the intentions behind the PISA test are positive, it becomes next to impossible to produce items that are 'authentic', close to real life challenges – and at the same without cultural bias and equally 'fair' in all countries. Items have to be constructed by international negotiations, and the result will therefore be that all contexts are wiped out – contrary to the ambitions of the PISA framework.

### ***Youth culture: Who cares to concentrate on PISA tests?***

In the PISA testing, students at the age of 15 are supposed to sit for 2 hours and do their best to answer the items. The data gathered in this way forms the basis of all conclusions on achievement and all forms of factor analysis that explain (in a statistical sense) the variation in achievement. The quality of these achievement data determines the quality of the whole PISA exercise. Good data assumes, of course, that the respondents have done their best to answer the questions. For PISA results to be valid, one has to assume that students are motivated and cooperative, and that they are willing to concentrate on the items and give their best performance.

There are good reasons to question such assumptions. My assertion is that students in different countries react very differently to test situations like those of PISA (and TIMSS). This situation is closely linked to the overall cultural environment in the country, and in particular to students' attitudes to schools and education. Let me give an illustration of such cultures with examples from two countries scoring high on tests like PISA and TIMSS.

### **Testing in Taiwan and Singapore**

An observer from Times Educational observed the TIMSS testing at a school in Taiwan, and he noticed that pupils and parents were gathered in the schoolyard before the big event, the TIMSS testing. The director of the school gave an appeal in which he also urged the students to perform their utmost for themselves and their country. Then they marched in while the national hymn was played. Of course, they worked hard; they lived up to the expectations from their parents, school and society.

Similar observations can be made in Singapore, another high achiever on the international test. A professor in mathematics at the National University of Singapore (Helmer Aslaksen) makes the following comment: "In this country, only one thing matters: Be best – teach to the test!"

He has also taken the photograph from the check-out counter in a typical Singaporean shop, (see Figure 7). This is where the last-minute offers are displayed: On the lower shelf one finds pain-killers, while the upper shelf displays a collection of exam papers for the important public exams in mathematics, science and English (i.e. the three PISA subjects). This is what ambitious parents may bring home for their 13-year-old kids at home. Good results from such exams are determinants for the future of the student.



**Figure 7**

*The context of exams and testing: This is the check-out counter in a shop in Singapore. Last-minute offers are displayed: On the lower shelf: Medicinal pain-killers. On the upper shelf: Exam papers for the important public exams in mathematics, science and English (i.e. the three PISA subjects).*

This is definitely *not* the way such testing takes place in my part of the world (Norway) and the other Scandinavian countries. Here, students have a very different attitude to schooling, and even more so to exams and testing. The students know that the performance on the PISA testing has no significance for them: They are told that they will never get the results, the items will never be discussed at school, and they will not get any other form of feedback, let alone school marks for their efforts. Given the educational and cultural milieu in (e.g.) Scandinavia, it is hard to believe that all students will engage seriously in the PISA test.

### **Task value: "Why should I answer this question?"**

Several theoretical concepts and perspectives are used to describe and explain performance on tests. The concept of *self-efficacy* beliefs has become central to this field. By self-efficacy belief, one understands it to be the belief and confidence that students have in their resources and competencies when facing the task (Bandura 1997). Self-efficacy is rather specific to the type of task in question, and should not be confused with more general psychological personality traits like self-confidence or self-esteem. PISA has several constructs that seek to address self-efficacy, and they have noted a rather strong positive relationship between e.g. mathematical self-efficacy beliefs and achievement on the PISA mathematics test on the

individual level (Knain & Turmo 2003). (It is, however, interesting to note that such a positive correlation does not exist when *countries* are the unit of comparison.)

There is, however, a related concept that may be of greater importance when explaining test results and students' behaviour in test situations. This is the concept of *task value beliefs* (Eccles & Wigfield 1992, 1995). While self-efficacy beliefs ask the question, "Am I capable of completing this task?", the task value belief focuses on the question, "Why do I want to do this task?". The task value belief concerns beliefs about the *importance of succeeding* (or even trying to succeed) on a given task.

It has been proposed that the task value belief may be seen to have three different components or dimensions: These are the 1. attainment value, 2. intrinsic value or interest, and 3. utility value. Rhee et al (2007) explains in more detail:

Attainment value refers to the importance or salience that students place on the task. Intrinsic value (i.e. personal interest) relates to general enjoyment of the task or subject matter, which remains more stable over time. Finally, utility value concerns students' perceptions of the usefulness of the task, in terms of their daily life or for future career-related or life goals. (Rhee et al 2007: 87)

I would argue that young learners in different countries perceive the task value of the PISA testing in very different ways, as indicated in this chapter's previous sections.

Based on my knowledge about the school system and youth culture in my own part of the world, in particular Norway and Denmark, I would claim that many students in these countries assign very little value to all the above three dimensions of the task value of the PISA test and its items. Given the nature of the PISA tasks (long, clumsy prose and contrived situations removed from everyday life), many students can hardly find these items to have high "intrinsic value"; the items are simply not interesting and do not provide joy or pleasure. Neither does the PISA test have any "utility value" for these Scandinavian students; the results have no consequence, the items will never be discussed, there is no feed-back, results are secret and do not count, neither for school marks nor in their daily lives. They do not count for students' future career-related or life goals. Given the cultural and school milieu and the values held by young learners in e.g. Scandinavia, it is hard to understand why they should choose to push themselves in a PISA test situation.

If so, we have an additional cause for serious uncertainty about the validity and the reliability of the PISA results.

## **References**

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bergesen, O. H. (2006). *Kampen om kunnskapsskolen* (Eng: The fight for a knowledge-based school) Oslo: Universitetsforlaget.
- Eccles, J. S., & Wigfield, A. (1992). The development of achievement-task values: A theoretical analysis. *Developmental Review*, **12**, 256–273.
- Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: the structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, **21**, 215–225.

- Fensham, Peter (2007). Values in the measurement of students' science achievement in TIMSS and PISA. In Corrigan et al (Eds) (2007). *The Re-Emergence of Values in Science Education*, (p 215-229) Rotterdam: Sense Publishers.
- Knain, E. & A. Turmo (2003). Self-regulated learning Lie S et al (Eds) *Northern Lights on PISA*, University of Oslo, Norway (p 101-112).
- Mullis, I. V. S., Martin, M. O., Smith, T. A., Garden, R. A., Gregory, K. D., Gonzales, E. J., et al. (2001). *TIMSS Assessment Frameworks and Specifications 2003*. Boston: International Study Center, Boston College.
- Prais, S.J. (2007). England: Poor survey response and no sampling of teaching groups *Oxford Review of Education* vol. 33, no. 1.
- Rhee, C. B., T. Kempler, A. Zusho, B. Coppola & P. Pintrich (2005). Student learning in science classrooms: what role does motivation play? In Alsop, S. (Ed). *Beyond Cartesian Dualism. Encountering Affect in the Teaching and Learning of Science* Dordrecht: Springer, Science and Technology Education Library.
- Ziman, J.M. (2000). *Real Science: What it is and what it means*. Cambridge: Cambridge University Press.