Real world Contexts in PISA Science: Implications for context-based science education

Peter J Fensham

School of Mathematics, Science and Technology Education, Queensland University of Technology, Kelvin Grove, QLD, Australia 4059.

Email: p.fensham@qut.edu.au

Abstract: The PISA assessment instruments for students' scientific literacy in 2000, 2003 and 2006 have each consisted of units made up of a real world context involving Science and Technology, about which students are asked a number of cognitive and affective questions. This paper discusses a number of issues from this use of S&T contexts in PISA and the implications they have for the current renewed interest in context-based science education. Suitably chosen contexts can engage both boys and girls. Secondary analyses of the students' responses using the contextual sets of items as the unit of analysis provides new information about the levels of performance in PISA 2006 Science. .Embedding affective items in the achievement test did not lead to gender/context interactions of significance, and context interactions were less than competency ones. A number of implications for context-based science teaching and learning are outlined and the PISA 2006 Science test is suggested as a model for its assessment.

Key words: Context-based teaching, socio-scientific issues, assessment, interest in science,

Introduction

The most recent round of science curriculum reform has seen a renewed interest in a number of countries in a greater use of real world contexts in the teaching of science (Aikenhead, 2005; Bulte, Westbroek, De Jong, & Pilot, 2006; Hofmann & Demuth, 2007; Mikelskis-Seifert & Duit, 2007; Roehrig, Kruse & Kern, 2007; Millar, 2007; Tytler, 2007). The idea of embedding science teaching and learning in contexts that are real and meaningful to learners, was explored extensively in the late 1980s, as first responses to the challenge of Science for All (e.g. USA: National Science Foundation, 1983; Canada: Science Council of Canada, 1984; UK: The Royal Society, 1985). Innovative projects like CHEMCOM in USA, LORST in Canada, SATIS and Salters' Science and Chemistry in England and Wales, and PLON Physics in The Netherlands all involved real world contexts with applications of science and technology. They became associated under the slogan of Science/Technology/Society (STS) and STS science teaching was discussed and encouraged by authors like Solomon and Aikenhead (1994) and Yager (1996). The PLON project, in particular, exemplified the idea of Concepts in Contexts, that is, concepts exist in science because of the relationships they have with other concepts and because these relationships have application and meaning across a variety of natural contexts (Lijnse, Kortland, Eijkelhof, van

Genderen & Hooymayers, 1990). Eijkelhof and Kortland (1988) describe how this idea was used to develop curriculum materials that facilitate the teaching of physics, the engagement of students and their learning of both the concepts and the contexts involved.

This movement towards a greater role for real world contexts in the teaching of science was then largely negated the mid 1990s, when national curriculum projects in most of these countries defined the science to be learnt, from the beginning of schooling to its last years, simply in terms long lists of concepts in science disciplinary strands (for example,. Project 2061 (AAAS), 1993 and the DES/WO (1991). One reason for this reversal from recognising real world contexts in these curricula was the establishment of Technology as a new and separate subject, subsuming the hitherto practical Industrial Arts subjects, but now emphasizng *Design* as a central creative and intellectual feature. The use of the name, Technology, for these subjects about Designing and Making effectively removed the T in the STS movement that was the bridge between the S of Science and the S of Society. Real world applications of science, in the form of technologies, as a consequence, dropped out of the 1990s reformulations of the science curriculum.

By 2003 many of the more industrially developed countries were reporting that students were emerging from school science with low interest in science for its own sake, and for themselves as career possibilities. These affective consequences are disastrous outcomes from school science. The larger society itself is increasingly aware of major personal, social and global problems, that involve science and technology, and which have complex ethical or moral implications.

Inappropriate applications of science as technologies have undoubtedly contributed to some of these problems, but science and its applications as technologies, remain essential components for their solution.

The 1990s curricula for science education are now being critically assessed, both within educational systems and beyond them, by groups concerned about their failure to contribute to the well being of science in contemporary society. An early example of these assessments was Beyond 2000, a report on the failure of the national curriculum for science in England and Wales (Millar and Osborne, 1998). Among its recommendations to reverse these trends was that science education must re-engage with technology as applications of science in society. Aitkenhead (2005) argued for a humanistic perspective in school science and set out how it would be different from traditional school science. Roberts (2007) in a succinct manner has rephrased this difference in two visions of the appropriate basis developing scientific literacy in school science. Vision I looks inwards to science itself –its products of concepts, laws and theories and its process of investigation Vision II looks outward at societal situations in which science has a role.

It is not an easy task for science teachers to switch to teach in a manner that that takes seriously real world contexts involving science and technology. Their own backgrounds are more likely to be in academic science than in any of its areas of application, and more often than not they will not have had direct experience of either scientific research or investigating real world problems. They may also be insecure about some of the science knowledge involved in these applications, and hesitant about their non-science aspects (Millar, 2007; Roehrig, Kruse & Kern, 2007; King, 2007).

Such real world contexts have, however, been a central feature of the OECD's PISA project for the assessment of scientific literacy among young people, 15 year olds. It is thus timely, following the report of the 2006 assessment when science was the major focus of the project, to consider what implications this project's experience may have for the renewal of interest in context-based science education.

Context in PISA Science

In 1998 the OECD launched the PISA project with the charter of measuring how well prepared in Science (along with Reading and Mathematics) 15 year olds are for life in the 21st Century. This prospective measure was a most unusual task for the assessment of science learning. Assessment of school learning is almost always retrospective, that is, measures are sought for how well what has been taught has been learnt and can be recalled. Furthermore, assessment of science learning has, locally, nationally and internationally, been dominated by items that test students' knowledge of what Roberts in his Vision I, above, calls the products science – concepts,. principles and theories and their application in contrived textbook situations. The OECD's charter for PISA required it to take a radically different approach to assess how well 15 year old students' science knowledge from whatever source can be applied to the situations involving science beyond school that increasingly confront citizens. The emphasis was to shift from measuring students' rather passive stores of knowledge to their ability to actively use knowledge in new situations. Laurie and Bybee (2009) have provided an account of how this different approach led to PISA 2006 Science and its implementation in more than 50 countries.

For the first PISA testing in 2000, Science was a minor domain, and the Science Expert Group decided to concentrate on just one important way science impinges on citizens, namely, through reports in the public media. Accordingly, the Science test was constructed around actual media reports about which students were asked questions. It quickly became apparent that these reports very commonly interweave applications of science (as technologies) with their background science. Studies of the public understanding of science have also confirmed the entwined association of science and its applications in everyday life (Fensham, Law, Li & Wei, 2000). From its beginnings PISA has bracketed Science and Technology (S&T) together, meaning by technology the applications of science. In 2003 Science, still a minor domain, repeated this format, and for 2006 when science was to be the major domain, the PISA Science Framework sets actual real world S&T situations or contexts as the starting point for the construction of the Science Achievement Test (OECD, 2006). The test thus consists of a series of units, each of which starts with a presenting context that is followed by a set of items that reflect the three scientific competencies and two scientific attitudes, that were derived from the project's definition of scientific literacy. Laurie and Bybee, 2009)

Choosing real world S&T contexts

The Science Expert Group for each of the three PISA testings has thus been faced with selecting S&T contexts that are not only real world situations but also are likely to be significant - engaging 15 year olds across the many participating countries as interesting and important aspects of modern life.
A survey by content analysis of S&T media reports would be the technical way of establishing a population of these reports and its categories of contexts. The item writers for PISA 2000, and for

the testings in 2003 and 2006 needed only a much more crude version of analysis of media or real world applications of science, knowing that the “significance” of the contexts they suggested would be subjected to three further stages of validation. The first of these stages was by the Science Expert Group itself who could reject contexts and their associated items on a variety of grounds including their cross-cultural robustness, their appeal to 15 year olds, and the obscurity of their science. The second stage was by the PISA managers from the participating countries, who were even more useful for validating the units with respect to the cross-cultural relevance and appeal to 15 year olds. The third validation for the surviving units was by the students in the field trials (conducted in all participating countries), via the statistical analysis of their responses to the items, to a presenting context,. and the unit more holistically. Some units (or some of their items) were rejected at this final stage if there was evidence of unexpected statistical behaviour or comments from national markers.

Finally, thirty six units were used in the PISA 2006 Science. Most of the specific contexts and their items could be related to just one of five broad categories, *Health, Frontiers of Science, Natural Resources, Environment,* and *Hazards,* but three of them had items that related to more than one category. The international report for PISA 2006, released eight of these units into the public domain (OECD, 2007)

Context as a source of scientific competence

For PISA 2006 Science, three scientific competencies were defined as the cognitive variables to be measured. These are *Identifying scientific issues, Explaining phenomena scientifically,* and *Using scientific evidence.* Together they cover a number of important aspects of the Nature of Science and of conceptual science. Thus the items relating to each of these competencies involved varying degrees of the Knowledge <u>of</u> Science and Knowledge <u>about</u> Science. The former is very familiar in traditional science education and its assessment as content knowledge in the science disciplines. Knowledge <u>about</u> Science is usually less emphasised, although acknowledged in the intended curriculum by terms like Working scientifically, Doing science, Scientific inquiry, etc

A great advantage of even the short description of a real world S&T context to introduce each unit’s set of items in the test is that it enables items about the context that related to each of these scientific competencies and attitudes, or to several items relating to the same competence. This led to a small, but coherent set of items, rather than to isolated items addressing simply one competence at a time. It also meant that a later item in the set could pursue thinking about the context that began in an earlier one. Finally, the unit structure of the test and the presenting contexts could accommodate the variety of assessment modes that PISA wished to use.
The released units provide examples of this coherence and item variety. In *Acid Rain,* Item 2, an open constructed response, assessed *explaining phenomena scientifically*. Item 3, a multiple choice response, assessed *using scientific evidence,* and Item 5, an open constructed response, assessed *identifying scientific issues* by pursuing an aspect in Item 3. In *Genetically Modified Crops,* Item 2, a complex multiple choice response, and Item3, a multiple choice response, both assessed *identifying scientific issues.* In *Grand Canyon*, Item
3 and Item 5, both multiple choice responses, assessed *explaining phenomena scientifically*, and Item 7, a complex multiple choice response, assessed identifying scientific issues. The last of these

items is interesting because it illustrated the facility that the complex multiple choice response format offers for comparing aspects of the Nature of Science that are much more difficult with the simple multiple choice format.

Context and the analysis of achievement

Although the PISA Science test was designed in terms of units about particular S&T contexts, the data from students' responses to the test have been analysed for the international and the national reports using the individual items as the unit of statistical analysis. This is consistent with the design that aimed primarily to measure the three scientific competencies that the project chose to reflect its definition of scientific literacy (OECD, 2006). The students' responses as raw data can, however, be re-analysed in terms of each contextual unit making up the test. That is, a student's raw scores from the set of items associated with each contextual unit can become the basis for considering the performance of a country's students (or any other subgroup of them) on that contextual unit in the test.

Combining these raw scores across the student sample leads to a mean raw score for each country in relation to each contextual unit in the test. When these mean raw scores are related to the maximum possible score for each contextual unit's set of items, a mean percentage correctness can be calculated. Correct responses to items are usually scored 1, but sometimes 2. The maximum score for a unit varies between 2 and 5 depending on the number of items. This derivation of a mean percentage correctness is somewhat akin to the way in which Laurie and Roberts (2009) describe proficiency levels for items and students are calculated and reported in the PISA reports as indications of the percentage of each country's students have achieved for this context. The higher the mean percentage correctness for a unit, the easier its items were, and the lower the mean percentage the more difficult they were. Within each country and across the whole PISA student population, a wide range of percentage correctness was manifested across the 36 units. To illustrate the information that can be obtained with this alternative analysis, findings for four countries are given -.Finland as the highest ranking country, Australia and its neighbour New Zealand as countries with well known similarities and differences, and USA for North America for this journal's audience. Table 1 gives the range of percentage correctness excluding one unit for which the correctness was more than 90% in each case.

A mean percentage correctness across all units for these countries is also given in Table 1, which in a close way relates to the rank order of the mean scores for the participating countries that appears in the PISA reports, when the students' scores have been standardised statistically to an OECD mean of 500 (OECD, 2007; Laurie & Bybee, 2009. An advantage of calculating these mean percentage correctness from the raw scores is that they indicate in a manner more familiar to teachers and their parents just what degree of achievement a group has reached on this test. The mean achievement scores in the official reports are very useful in providing relativities in achievement, and in relating differences to the percentage of students at a given proficiency level, or to a difference in another variable like socio-economic and cultural status. In another way they obscure the level of students' achievement on the test. The fact that Finland has a standardised scale score of 563 indicates a considerably higher performance than in the USA with a score of 489. But what do these scores really mean? Is 563 an almost perfect score, and how poor a score is 489?

Early in the development of PISA, the plans to measure students' ability to applying the scientific competencies, spelt out in the first PISA Framework, to unfamiliar, actual S&T media reports (OECD, 1999) was strongly criticised on two grounds by some members of the managing board and some national managers. Firstly, it was suggested that the project would simply show that 15 year old students, in general across the OECD countries, were unable to answer such test items, either because they were 'not mature enough' or because these skills were 'not being taught' in the current science curricula. Secondly, the verbosity of the presenting media reports was perceived to be so great that girls would be greatly advantaged because of their known greater performance in reading in most of the participating countries.

The second fear was allayed, but not explained, when 26 of the thirty two participating countries in 2000 showed no statistical difference between boys and girls, and the other six had small differences, with three favouring girls and three boys (OECD, 2001). The first criticism was also side stepped, when the wide spread of relative performance among the countries was announced, but the question of whether these were all quite low, or quite high remained. If the relative means were all quite high, little would have been learnt about improving the state of current national science education as preparation for life. If they were all quite low, then the task ahead to improve science education would be formidable indeed.

The use of means of percentage correctness helps to clarify this matter. A mean percentage correctness of 50% for a contextual unit indicates that the performance of the group of students on that context is likely to be made up of roughly equal numbers of students getting most of its few items correct and incorrect. The mean percentages for correctness in Table 1 show that in each of these four countries the performance of students on what was an unfamiliar type of science test was by no means disappointingly low. On the other hand, in none of the countries was the performance so good that there is not scope for improvement - a challenge and opportunity that needs to be taken up by governments in all countries.

As would be expected, countries where the students' standardised mean score for scientific literacy was high, tended also to rank highly on the mean percentage correctness for individual units. Consistent with its very high rank of 1.on the PISA scale scores, Finland's correctness ranking, among the 57 participating countries, was first or equal first for 18 of the 36 units, and for only one unit was its rank of 13, outside the first ten. Similar consistency, but with slightly lower correctness rankings, were found for New Zealand (between 1 and 25) and Australia (between 1 and 20) and, matching their respective PISA ranks of 7 and 8. The correctness rankings of the USA ranged from 6 to 38, with two thirds of the rankings between 24 and 36 – relating well to its PISA ranking of 29.
The findings of the international ROSE (Relevance of Science Education) project demonstrate quite significant difference between 15 year old boys and girls on many of its measures, including the science topics they would like to study in science (Schreiner, C. & Sjøberg, 2004). Extrapolating from these findings, context/gender interactions may be present in the responses to the PISA 2006 Science test.

However, as the overall gender neutrality in so many countries suggested, only a small number of contextual units showed a significant gender difference. For Australia, girls outperformed boys on three *Health* units and one *Natural Resources* unit and one *Environment* Unit, while boys more successful on two *Health* units, two *Natural Resources* units, and one *Hazards* unit. Once again the source of these differences may lie, not so much in a gender response to the context itself, but in the gender differences associated with the particular scientific competencies involved in the unit. Only two units with gender difference for Australia were released. Girls out performed boys on Sunscreen, which included two items involving *Identify scientific investigations*, which usually favour girls. Mary Montague (History of Vaccination) involves three *Explain phenomena scientifically* items which usually favour boys, and hence the observed difference in favour of girls for this unit may well be contextual.

When this gender analysis was extended to all 36 test units, six units had no competence emphasis, that is, no competence was associated with 50% or more of the unit's total score Six units emphasised *Identify scientific investigations,* twelve emphasised *Explain phenomena scientifically,* and twelve, *Using scientific evidence.* The contexts emphasising *Identifying scientific investigations (ISI)* and *Explaining scientific phenomena (ESP)* would, from the Australian PISA findings, potentially favour girls and boys respectively. The contexts emphasising *Using scientific evidence (USE)* should favour neither group. None of the six units emphasising *Using Scientific Evidence*, as predicted, showed any gender difference. Three of the six *Identifying scientific investigations* contexts showed a gender difference in favour of girls, and in the twelve *Explaining scientific phenomena* units four favoured boys and two favoured girls. This analysis suggests that gender/context interactions may be present; but that these are probably less than the known differences between gender and item type differences.

Context and Ease or Difficulty

The difficulty or ease that students have in answering a contextual set of items may be due to the context itself (more or less familiar, more or less interesting, more or less relevant) or to the types of items that are asked about it. Across the four countries in Table 1, there was a remarkable, relative consistency. Units that showed high correctness for Finland were also among the more highly correct in the three other countries. Conversely, the contextual unit for which Finland was least correct, was the least correct in the other three countries. Six contexts featured among the twelve more difficult contexts for all four countries.

This suggests that it is the nature of the items and not the context itself that determines difficulty, and this was borne out by the association these 'hard' and 'easy' contexts had with items of high and low level proficiency respectively.

The units for each country for which their correctness ranking are either high or low would, it seems, be respectively useful for identifying target topics and science content where science teaching is already contributing or where there is scope for improved teaching. For example, Australian students were ranked in the top five countries for 2 units in *Health*, 3 units in *Frontiers of Science,* and one in *Natural Resources.* They were ranked in the top 10 countries for four of the five units in *Environment.* However, Australian students also had low rankings in all four broad categories. The same relative success and failure across these categories was evident for the other three countries. The use of these higher and lower rankings to detect strengths and weakness in science education is confused by two aspects of the PISA design. Firstly, the S&T contexts of the 36 units are not equally distributed among the five broad categories, and secondly, the units can include items at very differing proficiency levels. For example, the items in Greenhouse (*Environment/Health*) are at proficiency levels 3, 5 and 6, whereas those for Mary Montague (*Health*) are only at proficiency levels 2 and 3.

The use of these higher and lower rankings to detect strengths and weakness in science education is confused by two aspects of the PISA design. Firstly, the S&T contexts of the 36 units are not equally distributed among the five broad categories, and secondly, the units can include items at very differing proficiency levels. For example, the items in Greenhouse (*Environment/Health*) are at proficiency levels 3, 5 and 6, whereas those for Mary Montague (*Health*) are only at proficiency levels 2 and 3.

The five categories for the S&T contexts are clearly too broad for identifying the weaknesses or strengths in how well students have been prepared for 21st C life. If the specific S&T contexts in PISA do themselves constitute a source of weakness in a country's PISA scores, they must be explored in a more detailed way. Since only eight of the thirty six units have been released from the 2006 test, only an indication of this type of detailed analysis can be given.

Mary Montague (*Health*), with items at proficiency levels 2 and 3, had very high correctness in all four countries and Clothes *(Frontiers of Science*) and Physical Exercise *(Health)* also had high correctness, despite an item at proficiency level 4 in each.

Greenhouse (*Environment/Health*)*, and* Acid Rain *(Environment*) were difficult for all four countries, but both have items at the higher *proficiency levels* of 5 and 6. Genetically Modified Crops (*Frontiers of Science*) was difficult for Finland, Australia and New Zealand, although its items were only at proficiency levels 2 and 3. It was not difficult for USA. Grand Canyon (*Environment*) had items at the lower *proficiency levels* of 2 and 3, but yet was difficult for Finland, and Sun Screens (*Health*)*,* with three *proficiency level* 4 items was difficult for USA, but not for the other three countries.

Students' attitudes

The decision to embed attitudinal questions in the main Achievement Test for PISA Science in 2006 was a radical one. In the earlier PISA testings, and in the four IEA studies of school science achievement since the 1970s, attitudinal questions have always been part of a student questionnaire that is separate from the achievement test (Gonzales et Al, 2004). A primary reason for this radical decision was that PISA, as a very significant international project involving more than 50 countries, would thus convey a strong message that positive affective outcomes should be as expected from the learning of science as are positive cognitive outcomes. The need for such a message was clear by 2003 when PISA 2006 was being planned since many of the OECD countries were reporting a disturbing down turn in the interest of students in science, both for its own sake and for possible career directions. In this climate it would have been remiss of the PISA project not to take up the opportunity to make student affect about science a central feature of its testing in 2006, when Science was to be the major domain.

Another reason for embedding attitudinal items in the S&T contextual units of the test, taken by all students, is that it gives affect about science an expectant social sense. This compares with the personal sense it has when the affective items are in an instrument, that is primarily gathering personal information about a student's self, habits, family and opinions. Embedded items have other advantages over similar attitudinal items in a student questionnaire. Embedded items are able to explore students' attitudes in terms of concrete specific S&T contexts and their associated science, compared with the use in questionnaire items of the words "science" and "technology" which assumes they have a generic meaning across 15 year olds in various countries and cultures. Finally, they have the potential to determine whether students' attitudes vary between S&T contexts and whether the attitudes expressed are related to their cognitive achievement, at the levels of individual items or of a unit's set of items.

The embedded items in PISA 2006 were designed to relate to two attitudinal constructs, *support for scientific enquiry* and *student interest in learning science topics*. The items for the former were Likert type in which the students were asked to express their level of agreement (on a four point scale: *strongly agree to strongly disagree)* with statements that were directly related to the particular S&T context (see examples in Table 2). The items for the latter construct asked students to their degree of interest (on a four point scale: *high interest to no interest)* in learning more about science aspects of the particular S&T context (see examples in Table 2). In the trials of these test units, the students' responses were found to have sufficient coherence for them to be used to constitute a scale for these constructs as variables.

Students' support for scientific investigation

Students in all the participating countries reported strong levels of support for scientific enquiry. However, there were some marked differences in the level of support for particular scientific investigations about the same presenting context. In Mary Montague (*Health*) 94% (on average) supported research to develop new vaccines. In contrast 30% disagreed with the cause of disease requiring scientific research, and 87 favoured investigation of unconventional treatments of diseases. For the eight statements in the three released items in Table 2 only six countries (of 57) had less than 60% positive agreement, and the great majority were over 80%. This is not to say that there was no variation in the students' support. In contrast 30% disagreed with the cause of disease requiring scientific research. ,, For the four exemplary countries, Table 2 gives the mean percentages for agreement (*strongly agree* plus *agree*) about each of these statements in the released units and significant variation is evident.

The flexibility of real world contexts to enable a variety of items is exhibited in Table 2 where several possible scientific investigations were included for each of these S&T contexts. More could have been suggested without exhausting the context albeit so briefly presented in the test. OECD (2007) discusses the released units and their characteristics in detail.

Students' interest in learning science topics

The S&T contexts in the test units also enabled a number of questions to be asked that referred to interest in learning the associated pure or applied science. Across the pooled OECD sample, the correlation of these responses as a scale (reliability index, 0.87) had a very weak negative correlation (r = -0.06) with performance – a sobering finding indeed with respect to the issue in the introduction to this section.

Students' expressed different levels of interest in learning these particular science topics. In Acid Rain (*Environment*), for example, 62% reported positive interest (*high interest* plus *interest*) in human causes of acid rain, but only 49% in the repair of buildings damaged by acid rain. In Genetically Modified Crops (*Frontiers of Science*) only 47% expressed interest in learning more about its topic. Table 3 gives the level of positive interest among the students in learning more about these aspects of science for the four exemplary countries.

Interest in learning these aspects of science is much lower, in general, than support for scientific enquiry. Mean percentages less than 40 were not uncommon, particularly among the OECD countries. Some participating countries commonly showed much higher levels of interest, and often in association with lower mean achievement of scientific literacy, confusing the correlation between them. Again, there was significant variation within countries about the different aspects of the science of a unit, and between countries. Differences between boys and girls were generally small. A more encouraging, albeit still low relation for many countries was the correlation between the measure for *general interest in science learning* (items for this construct were in the student questionnaire) and combined scientific literacy. For Australia, the value of this correlation was $r =$ 0.30.

Implications for context-based teaching

Choice of contexts

Science teachers in the new context-based curriculum will often not be responsible for the choice of contexts because these are prescribed in the curriculum or its accompanying materials (for example, Twenty first Century Science: Millar, 2007). Hopefully, the curriculum writers have been through some objective processes for determining the worth of their contextual choices. A sample of teachers and their students, as validators of these choices should be part of the process. A useful dictum associated with the development of Salters' curricula, based at York University, England is that a science concept should not be introduced until it is important for understanding some familiar context in the lives of their students outside school (Smith, 1088).

When these contextual choices are settled, there is then a substantial task of professional development to be done (King, 2007; Millar, 2007) in familiarising the teachers as a whole group with these contexts, the science and technology involved, and how to use them effectively in science teaching. Teachers' own affective engagement with these S&T contexts is important if they are to engage their students.

Teachers who are faced with choosing contexts for their science teaching have two advantages over those with prescribed contexts. To an extent know their students and hence will have ideas about which contexts may be relevant and engaging. Certainly they can get feedback quite rapidly from the students about their choices. Secondly, engaging with topics of their own choice is psychologically easier than with imposed contexts.

Neglected S&T contexts of popular importance

None of the PISA Science testings have made use of two very different types of S&T context that regularly receive attention in the public domain. They are also largely neglected in school science. Public interest in them is, however, very high, judging from the prominence given by the media. The first of these are related to cosmology including reports about space exploration.

The reporting, for example, of a star that is older than the universe, or of a scientific experiment in Switzerland seeking to reproduce conditions within micro-seconds of the birth of the universe, might seem to be esoteric for school science. Nevertheless, such phenomena regularly command prime space in the mass media. *Bright lights and biggest of bangs* was a headline in newspapers. Five hundreds words followed, plus an artist's conception of the bright gamma ray burst close up that was visible as a faint light in the sky on March 19, 2008. It was the explosion of a dying star, 7.5 billion light years away, as the star collapsed into a black hole, driving bright gas jets into space (Soderberg et al, 2008) It was *"one of the biggest bangs since the big one"*, and "*the most distant astronomical object ever seen by humans"* – claims for attention that are not commonly matched in school science!!

Again, when the Sloan Digital Sky Survey (www.sdss.org) sought expressions of interest from individuals to categorise images of a million galaxies via their home computers, more than 40,000 persons world volunteered, making this probably the largest ever professional/amateur scientific investigation. Already some volunteers have already found examples of new astronomical phenomena .

It is possible that the concern of the PISA project with its OECD charter, namely, "the preparedness of 15 year olds for 21st C life" has led it to focus only on scientific knowledge and scientific competencies that are pragmatically useful. Sjøberg (2007) has argued that this focus has led to the neglect of other roles science plays, for scientists and for the public. One such a role is creating a sense of wonder about the natural world, an appreciation of the amazing interconnectedness of the biophysical world we all inhabit, and the mystery of the universe itself. How to incorporate the importance of this "sense of wonder" into school science curricula remains a challenge that has little to learn from PISA. Part of the current malaise in science education, particularly in the elementary years, may be due to curriculum designers, and hence teachers, moving too rapidly to the 'useful' concepts of science, before the phenomena and contexts for which these concepts were invented have been adequately appreciated.

The second type of prominent, but neglected, S&T contexts are military contexts, for example, those associated, since 2001, with the so-called, *War on Terror*, like weapons of mass destruction. As PISA 2006 Science was being prepared, it was obvious that modern warfare and military weapons must be included under *Hazards* in its matrix of important S&T contexts, because of the immense consequences –Personal, Social and Global - associated with their use in conflict. As an international project involving countries with diverse political stances on the War on Terror it is perhaps not surprising that no units involving these contexts have yet appeared in the Achievement Test. Nevertheless, units have been included in relation to controversial environmental and health issues that are equally divide the participating countries politically.

Political delicacy cannot be used as an excuse to exclude S&T contexts of importance from intra-national science education. Australia joined USA and some other countries in an invasion of Iraq on the basis of the presence of undeclared weapons of mass destruction in that country. It is likely that few citizens and their politicians in Australia were conversant with exactly what these weapons were, although the scientific properties that give them this classification are quite simple to spell out (Weapons of Mass Destruction Commission, 2006). In the event, no such weapons were found, and other reasons for the invasion and its continuance had to be developed. A small study of 13 and 15 year old students in a strong science high school in Australia showed that the their seven or nine years of school science had provided them with only a little more awareness of the scientific character of these weapons than they had constructed from the imprecise discussion of them in the mass media (Fensham and Moulds, 2007).

S&T Contexts as interdisciplinary science

The use in the PISA project of real world contexts made clear that they are rarely mono-disciplinary in a scientific sense, and even less commonly are they purely scientific. This disciplinary complexity has several advantages for teaching science. One of these is the fact that the same context can be used by teachers of a single science subject or by teachers of a more integrated science subject. In the former the single disciplinary aspect of the context can be the focus, but with acknowledgement of how it interacts with other science knowledge. The PLON unit, *Ionising Radiation,* primarily dealt with the physics of radioactivity and its measurement with the *curie* as unit, but also introduced students to the effect of such radiation on human bodies and that phenomenon's unit measure, the *sievert*. The use of MRI or ultrasound in human biology are conversely contexts where the teaching focus could be on the biology, with acknowledgement to the physics of these powerful tools of medical science, or it could be more on the physics with the biology as applications. In more integrated science teaching, the inter-disciplinary character of the science of these contexts enables a more holistic and balanced approach to the underlying science. Rennie (2007) has reported how teams of students can successfully carry out investigations in real world situations as an extension of their school science.

Another advantage of real world contexts is their inclusion of non-scientific aspects –social, aesthetic, economic, and ethical. The teaching of a scientifically investigable question or the nature of evidence in science can be greatly helped by comparing these notions with questions and the use of evidence in these other aspects of living. In some of the PISA units, items requiring such comparative understanding were included. Comparative teaching of this type has not been a feature of science education or of its research, but it is a powerful new pedagogy for both affectively and cognitively engaging students with context-based science. At one point in the development of PISA 2006 Science it was intended that items about the students' sense of responsibility for the environment would also be embedded in the Achievement Test units, but these items were then shifted to the students questionnaire. If they had been embedded, the test would have moved, at least a little, towards meeting the interest that Sadler and Zeidler (2009) have in the moral development of students that real world contexts offer.

Contexts and more balanced science teaching

It has been recognised for more than a decade that there is a need to balance the rather static learning of the conceptual content of science with the more dynamic Nature of Science, but curriculum documents have often listed these as separate strands with a hand waving vagueness of how they should be interrelated with the conceptual teaching. Most science teachers are equipped from their own education in the sciences with some degree of conceptual knowledge. Few of them have had the opportunity to personally experience the Nature of Science as researchers or to have studied it explicitly as a component of their education. They do not therefore find it easy to do justice in their teaching to the interrelationship intended. They are not helped by the fact that assessment instruments, local, national and international, continued throughout the 1990s to emphasise conceptual understanding (Fensham, 1999). Without the reinforcement of assessment that is authentic to the intentions of the curriculum they are bound to fail in practice (Llewellyn, 2005).

Real world S&T contexts have been shown in PISA to be rich sources for items that emphasise both Knowledge <u>of</u> Science and Knowledge <u>about</u> Science. It is as natural to ask of these contexts*, What scientific questions relate to this issue? a*s it is to ask, *What are the underlying science concepts?* Science & Technology contexts therefore facilitate the intentions of science curricula about the Nature of Science and assist teachers to achieve a connectedness between its scientific procedures and the science concepts that are related to them.

Including affective development in Context-based teaching

One of the goals for context-based science teaching is that it will lead to more engagement of students and to their developing more interest in science. As indicated earlier, the choice of appropriate contexts for science teaching is a first step to realising this goal. Real world contexts from the students' lives outside of school have the potential to generate personal intrinsic interest, and their social or global significance can add to this potential an extrinsic quality to this interest. This potential needs then to be realised with engaging pedagogies. In PISA 2000 Science the unexpected gender neutrality seems to have resulted, at least in part, from a number of the presenting contexts being stories that involved people and science. The engagement of boys and girls equally across so many countries supports the suggestions by Malcolm (1996) and Millar and Osborne (1988) that *Science as A Story* need to become a quite central pedagogy in science teaching Despite the limitations of paper and pencil items, the embedding of affective items in the contextual units of PISA Science 2006 was possible. They used the notions of *varying degrees of interest* and of *different opinions* about science issues. If these notions are translated into pedagogical terms for context-based science teaching, they resonate with allowing students to engage with the contexts of the PLON modules in optional ways (Eijkelhof & Koortland, 1988) and with the interest that science students expressed in being able to debate and discuss, at present so lacking in their science classrooms compared with other subjects Lyons(2004).

Conclusion

The PISA project is not an assessment of any present curriculum for school science. It has, to the extent of its success, demonstrated that it is possible to develop assessment tools, centred on real world S&T contexts, that have features that are consonant with the intentions now commonly expressed for the school science curriculum for greater depth in conceptual understanding, a stronger appreciation of the nature of science as a human enterprise and its increasing importance in so many key public issues. A key aspect of implementing these curricular intentions with fidelity is how this context-based teaching and learning of science will be assessed. PISA 2006 Science provides just such an assessment model.

References

Aikenhead, G.(2005). Science for Everyday Life: Evidence-based practice. New York: Teachers College Press.

Bulte, A.M.W., Westbroek, H.B., De Jong, O. & Pilot, A. (2006). A research approach to designing chemistry education using authentic practices as contexts, International Journal of Science Education, 28(10), 1063-1086.

DES/WO (1991). Science in the National Curriculum. London: HMSO

Eijkelhof, H. & Kortland, K. (1998). Broadening the aims of physics education, In P.J. Fensham (Ed.) Development and Dilemmas in Science Education (pp.282-305).London: The Falmer Press.

Fensham, P. J. (1999). International success, but is it science?: Identifying strengths and weaknesses in Australian primary school science from TIMSS and other data, Australian Science Teachers' Journal, 45(2), 39-44.

Fensham, P. J., Law, N., Li, S. & Wei. B. (2000). Public understanding of science as basic literacy, In R.T Cross & P.J.Fensham (Eds.) Science and the Citizen: for educators and the public. Melbourne, Australia: Melbourne Studies in Education.

Fensham, P.J. & Moulds, P. (2007). Students understanding of weapons of mass destruction. Paper presented at annual conference of ASERA, Fremantle, Western Australia, July

Gonzales, P., Guzman, J.C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D. & Williams, T. (2004). Highlights from the Trends in International Mathematics and Science Study: TIMSS 2003. Washington, DC: Institute of Educational Sciences.

Hofmann, D.& Demuth, R. (2007). Chemie in Kontext in der Hauptschule – geht den das? Der Mathematische und Naturwissenschaftliche, Unterricht –MNU, 60(5), 299-303.

Hopman, S. (Ed.) PISA according to PISA. Does PISA keep its promises? Wien: LIT Verlag.

King, D. Teacher beliefs and constraints in implementing a context-based approach in chemistry. Teaching Science: Journal of the Australian Science Teachers Association, 53(1),14-18.

Laurie, R. & Bybee, R. (2009). Overview of Pisa 2006 Science. Journal of Research in Science Teaching, (in press)

Llwellyn, D (2005). Teaching high school science through inquiry. A case study approach. Thousand Oaks, CA: Corwin Press.

Lijnse, P. L., Kortland, K., Eijkelhof, H. M.C., van Genderen, D. & Hooymayer, H.P. (1990). A thematic physics curriculum: a Balance between contradictory curriculum forces, Science Education, 74(1), 95-103.

Lyons, T. (2006). Different countries, Same classes: Students' experiences in their own words, International Journal of Science Education, 28(6), 591-613.

Malcolm, C. (Ed.) (1996). Could We? Should We": Science for year 10. Melbourne, Australia: Curriculum Corporation.

Mikelskis-Seifert, S. & Duit, R. (2007). Physik in Kontext – Innovative Unterrichtsansätze für den Schulalltag. Der Mathematische und Naturwissenschaftliche, Unterricht – MNU, 60(5), 265-274.

Millar, R. & Osborne, J. (1998). Beyond 2000. London: King's College.

Millar, R. (2007). Twenty first Century Science: Implications from the design and implementation of a scientific literacy approach in school science. International Journal of Science Education, 28(13), 1499-1521.

National Science Foundation (1983) Educating Americans for the Twenty First Century, Report of the National Science Board on Pre-College Education in Mathematics, Science and Technology. Washington, DC: National Science Foundation.

OECD (1999). Measuring Student Knowledge and Skills: A new framework for assessment. Paris: OECD.

OECD (2001). Knowledge and Skills for Life: First results from PISA. Paris:OECD.

.OECD (2006). Assessing Scientific, Reading and Mathematical Literacy: A framework for PISA 2006. Paris: OECD.

OECD, (2007). PISA 2006 Science Competencies for Tomorrow's World, Volume 1: Analysis. Paris: OECD.

Project 2061 (AAAS) (1993). Benchmarks for Scientific Literacy. New York: Oxford University Press.

Rennie, L. J. (2007). Learning science outside school. .In S. K. Abell & N.G. Lederman (Eds.) Handbook of Research on Science Education (pp.125-171).

Roberts, D.A. (2007). Scientific literacy/ science literacy. In S. K. Abell & N.G. Lederman (Eds.) Handbook of Research on Science Education (pp.729-780).

Roehrig, G.H., Kruse, R.A. & Kern, a. (2007). Teacher and school characteristiucs and their influence on curriculum implementation, Journal of Research in Science Teaching, 44(7), 883-907.

Sadler, T.D. and Zeidler, D.L (2009) Scientific Literacy, PISA and Socioscientific discourse: Assessment for progressive aims of science education. Journal of Research in Science Teaching, (in press)

Schreiner, C. & Sjøberg, S. (2004). Sowing the seeds of ROSE: Background, rationale,questionnaire development and data collection for ROSE (relevance of science education): A comparative study of students' views of science and science education (Acta Didactica 4/2004). Oslo: Department of Teacher Education and School Development, University of Oslo.

Science Research Council of Canada (1984) Science for Every Student: Educating Canadians for tomorrow's world. Ottawa, Canada: Supply and Service.

Sjøberg, S. (2007) PISA and "Real Life Challenges".: Mission Impossible? In S. Hopman, (Ed.) PISA according to PISA. Does PISA keep its promises? Wien: LIT Verlag.

Sjøberg, S. (2007) PISA and "Real Life Challenges".: Mission Impossible? In S. Hopman, (Ed.) PISA according to PISA. Does PISA keep its promises? Wien: LIT Verlag.

Smith, N. (1988). In support of an application-first chemistry course: Some reflections on the Salter's GCSE scheme. School Science Review, 70 (250), 108-114.

Soderberg, A.M. et al (2008) An extremely luminous X-ray burst at the birth of a supernova, Nature, 454(7201), 246.

Solomom, J. & Aikenhead, G. (Eds.) (1994). STS Education: International perspectives on reform. New York: Teachers College Press.

The Royal Society (1985). Science is for Everybody: Executive Summary from The Public Understanding of Science. London: The Royal Society.

Tytler, R .(2007). Re-imagining Science Education: Engaging students in science for Australia's future. Melbourne Australian Council for Educational Research.

Weapons of Mass Destruction Commission (2006). Final Report: Weapons of Terror- Freeing the world of nuiclear, biological and chemical arms. Stockholm, Sweden: Weapons of Mass Destruction Commission.

Yager, R. E. (Ed.) (1996). Science/Technology/Society as reform in science education. Albany NY: State University of New York Press.