

New Zealand literacy standards in a global context: The uses and abuses of international literacy surveys

WARWICK ELLEY

Emeritus Professor: University Of Canterbury, NZ

ABSTRACT: This article addresses the contentious issue of changes in literacy levels, by outlining the results obtained by New Zealand students in international surveys conducted by IEA over the past three decades. New Zealand students have traditionally achieved at high levels in these surveys, but the 1991 PIRLS survey seemed to suggest that standards were declining, because the country's ranking dropped to 13th place. While the media and some critics saw this as a cause for concern, the evidence shows that New Zealand reading standards have remained remarkably stable over the past 30 years, despite enormous increases in the numbers of immigrants and ESOL children in schools. It is argued that a nation's rank order can be misleading if certain technical assumptions are not met, and that the PIRLS survey had real difficulties in achieving good comparable samples in each country.

Numerous findings in these surveys are quoted to indicate that the New Zealand tradition of teaching reading is effective. Nevertheless, the large gender and ethnic differences in literacy are a continuing cause for concern.

KEY WORDS: Literacy, international surveys, IEA, PIRLS, literacy rankings.

In 2003, a disturbing headline appeared in a New Zealand education journal with a relatively wide circulation. It proclaimed that there had been a "Big drop in New Zealand children's reading performance" (Education Today, 2003). It would appear that yet another crisis in literacy was in the offing. In 2001, New Zealand 10-year-olds had taken part in a cross-national survey of reading-literacy, known as PIRLS – or Progress in International Reading-Literacy Study. This survey was one of a series that had been conducted by the International Association for the Evaluation of Educational Achievement (IEA) over a period of 30 years.

As the media reported, New Zealand students were ranked only 13th out of 35 countries in this survey, and although their mean score was well above the international average, the result gave rise to much soul-searching and "media-bashing". In the article mentioned above, the drop in reading performance was described as "dramatic" and "particularly disturbing to find that New Zealand was almost the lowest of the English speaking countries." One commentator used the result to argue that New Zealand teachers needed to take up the systematic teaching of phonics "that is deliberately separated from reading stories" (Education Today, 2003, p. 30). It appeared that New Zealand had fallen from the grace that earlier surveys had endowed it with. How could this have happened? How well did New Zealand students really read?

In this article, I will review the results of earlier international surveys and try to account for the alleged decline in New Zealand's reading achievement, as suggested by the media in response to the PIRLS survey. Such a review aims to contextualise

the PIRLS survey and give a wider perspective on the reading performance of children in New Zealand classrooms which have been becoming increasingly multicultural and multilingual during the 1990s.

WHAT IS IEA?

IEA is an international research body with headquarters in the Netherlands. It was established to conduct surveys of achievement on large cross-sections of students of comparable ages in its member countries, using the same array of instruments, given under the same conditions. The tests, questionnaires and attitude scales used are developed collaboratively, translated where necessary by rigorous procedures, pilot tested to identify cultural bias, and then administered to carefully selected representative samples of students in each country.

In the past 30 years, IEA surveys have been conducted in reading, mathematics, science, writing, literature, French, pre-school education and computer studies. Some of the early studies attracted only ten or a dozen countries, but recent surveys have seen over 40 countries participate. Officially, the main purpose of the studies is to enable policy makers in each country to investigate the effect on student learning of different traditions and policies which are found in the participating countries. Variations in policies between countries make possible the investigation of issues which are difficult to study within countries. For example, if all schools within a country are required to teach for 200 days, it will not be easy to study the effects of changing the length of the school year.

However, in recent years, a major motive for participation, and a major preoccupation of education commentators, has been to compare mean scores across countries and establish where each country is ranked. While this is an understandable incentive, its validity depends very much on the fulfilment of certain assumptions about equality in age, in sample selection, and test suitability, in particular. These matters are crucial in any interpretation of a country's PIRLS results from 2001.

NEW ZEALAND STUDENT ACHIEVEMENT IN PAST IEA SURVEYS OF READING

In 1970, New Zealand first participated in an IEA study of reading comprehension in 15 countries at two age levels – 14 yrs and 18 yrs (pre-university). To its surprise, New Zealand students produced the highest mean score at both levels, higher than students in England, Scotland, USA, Finland, Sweden, Australia and other countries it likes to compare itself with. New Zealand did not participate at the primary school level in this survey.

While the 18-year-olds' result may have been helped by the fact that New Zealand had a more select section of the age-group still at school (only 13%) than some other countries, the mean score for the 14-year-olds was clearly a positive sign for the quality of its reading programmes. New Zealand had a high mean and a relatively small range (or standard deviation) (Thorndike, 1973). The researchers experimented with tests of vocabulary and speed reading, but these were unable to be translated or

interpreted fairly, so were dropped from consideration. Table 1 shows the results for 15 countries in the 1970 reading comprehension survey (together with Australia, which conducted the survey later, in 1976).

Table 1. Mean scores in reading comprehension in 16 countries (from Thorndike, 1973)

Country	Population II (14yrs)	Population IV (18yrs)
Australia *	25.3	--
Belgium (Fl.)	24.6	25.0
Belgium (Fr.)	27.2	27.6
Chile	14.1	16.0
England	25.3	33.6
Finland	27.1	30.0
Hungary	25.5	23.8
India	5.2	3.5
Iran	7.8	4.4
Israel	22.6	25.2
Italy	27.9	23.9
Netherlands	25.2	31.2
New Zealand	29.3	35.4
Scotland	27.0	34.4
Sweden	25.6	26.8
USA	27.5	21.8

* Australia conducted the survey in 1976, at the 14-year level only.

In the same year, 1970, another survey was conducted on the interpretation of narrative literature, at the same two age levels. Again, New Zealand students had the highest mean, in each sub-test and at both age levels (Purves, 1973). Such results were taken as an indication that the country's teachers were having a positive effect. New Zealand students' performance in 1970 in three other subjects – science, French and civics was less impressive, but generally above the international averages.

In New Zealand, much discussion followed to identify the major factors entailed in these impressive findings in reading and literary interpretation. Some claimed that it validated the kind of literature-based reading programme promoted in the primary schools by the (then) Department of Education – with lots of reading *to*, *with*, and *by* students in a print-saturated environment (Department of Education, 1985). New Zealand has long had a tradition of providing good quality children's literature, along with the popular government-sponsored *School Journals*, *Ready to Read* booklets and other publications. The prevalent use of distinctive reading instructional methods, such as language experience (Ashton Warner, 1963), shared reading (Holdaway, 1979) and guided reading methods designed to exploit the potential of these high-interest reading materials, was thought by some to have made a major contribution to these strong results.

Others preferred to emphasise the fact that New Zealand in the 1970s was an egalitarian society, with full employment, few rich or poor, very few immigrants with English as a second language, and a homogeneous teaching force, well-distributed throughout the nation by virtue of a country service bar. No doubt, each of these, and

other factors, played a part. The analysis of background variables in school and home by the IEA researchers was not able to pin down and rank the critical causal factors. Nevertheless, the results were a clear indication that New Zealand children at the time were good readers, that girls were more proficient than boys in most countries, and that the best readers came from homes with large numbers of books and well-educated parents (Thorndike, 1973).

An important contribution of this study was the demonstration that it was possible to assess students' reading comprehension across countries and languages and obtain a valid result. Evidence for this conclusion came from the fact that the correlations between results for each passage and question across countries were positive and reasonably high. In addition, the tests operated in similar fashion in each country.

THE 1990 SURVEYS OF READING LITERACY

In 1990, New Zealand participated in the next IEA surveys of reading for 9 and 14 year-olds in 32 countries. New tests of narrative prose, expository prose and documents (graphs, maps, tables) were developed collaboratively and pilot tested in each country. These were supplemented by tests of word recognition and student questionnaires about reading habits and attitudes. Teachers and principals also completed questionnaires on their background circumstances and classroom practices. Over 210,000 students and 10,000 teachers took part in this massive survey, which was conducted in 1990 in the Southern Hemisphere, and 1991 in the Northern.

Results from the New Zealand surveys again showed very high levels of performance, with means clustering around 70% in each sub-test at both age levels. However, New Zealand lost its top placing to Finland, which had the highest means at both levels. New Zealand students were ranked fourth at the 14-year level, after Finland, Sweden and France (Elley, 1992). Nevertheless, New Zealand had the most students in the top quartile – that is, proportionately more excellent readers than any other country. At the 9-10 year level, New Zealand was ranked 6th, after the same three countries, plus USA and Italy (Elley, 1992).

There was much more of interest in the analyses of these surveys but a major concern for New Zealand educators was the discovery that the country's spread of scores was surprisingly large, at both age levels. While the country had many excellent readers, it also had a long "tail" of low achievers. Local analyses revealed that this tail was made up predominantly of ESOL children, Maori and Pacific Islands boys, and other pupils in low decile schools. Further study of their profiles showed that they mostly came from homes with few books and spent considerable time watching TV (Wagemaker, 1993).

Another source of concern was gender differences favouring girls, which were also found to be one of the largest among the countries surveyed. While some educators saw this difference as a reflection of the predominance of women teachers in the primary school, a comparative analysis showed that most other countries with smaller gender gaps had even *more* women teachers. The "feminization" explanation began to wilt under the pressure of comparative scrutiny. While the gender gap narrowed considerably by age 14 in most countries, it was interesting to note that countries

where pupils started formal schooling at age five consistently showed larger gender gaps, and these gaps were generally maintained into secondary school (Elley, 1992). A plausible hypothesis emerged, that boys who are introduced to formal instruction at age five before many of them are ready (due to physical and cognitive immaturity) develop negative self-concepts about their ability in reading. Feelings of failure, once established, are difficult for children to shake off. Certainly, New Zealand statistics on children who are struggling in reading consistently show that boys are over-represented (Ministry of Education, 2002).

Amongst the other findings of interest were results that showed that top countries started formal instruction at age seven. New Zealand was the only country in the top 10 that started instruction at age five. Another unexpected finding was related to trends in TV viewing. Generally, good readers watched relatively little TV. But in Scandinavia, the best readers were watching 3-4 hours per night (Elley, 1992). Subsequent investigation revealed that these readers were frequently watching imported children's programmes with sub-titles in their own language. Apparently, the motivation to comprehend what they saw was strong, the cues to understanding were rich, and the ability to read quickly was required to follow the stories. There is now much evidence that captioned TV is a promising support for improving reading comprehension (see, for example, Neuman & Koskinen, 1992).

There was considerable support in these surveys for the kinds of reading programmes common in New Zealand primary school classrooms. When scores were adjusted for socio-economic status, it was clear that *access to books* was a critical factor in reading success – whether in home, school or community. Teachers who read often to children, had large class libraries and provided time for silent reading produced better outcomes, other factors being equal (Elley, 1992).

Furthermore, the pupils who achieved at the highest levels in the top countries surveyed believed that to become a good reader you had to “have many good books around,” “have a lively imagination” and “learn many new words”. By contrast, the good readers in the lowest scoring countries believed that the secret of good reading lay in “learning to sound out the words”, “drilling the hard things”, and “doing lots of homework” (Purves & Elley, 1994). If these students' views indeed reflected the kind of teaching they received, there were some interesting implications in this finding. For instance, it is generally believed that New Zealand teachers conduct most of their phonics teaching within the context of a good story and in response to need, rather than systematically and in isolation from a current text. Such an approach is supported by the contrasts highlighted by the survey data.¹

THE PISA SURVEY OF 2000

In 2000, New Zealand took part in a survey of 15-year-old students in 32 countries, conducted, not by IEA, but by the Organisation for Economic Co-operation and Development (OECD). The assessments were made in reading literacy, with a minor focus on mathematics literacy and scientific literacy. The survey was known as PISA, which is the acronym for the Programme for International Student Assessment. In

¹ Further findings of these IEA surveys of 1990-1991 can be found in Elley (1994).

New Zealand, 3667 students from 153 schools took part. Further surveys are planned by the OECD for 2003 and 2006.

As in previous surveys, new tests were developed collaboratively and pilot tested before use. Reading Literacy was defined as “the ability to understand, use and reflect on written texts, in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate effectively in society”. The sub-skills that were assessed in the survey were described as “retrieving information”, “interpreting texts” and “reflection and evaluation”. A wide range of passages and question types was used to assess students’ skills.

In the overall results, New Zealand students once again achieved at very high levels. They gained a mean score of 529 in reading literacy, where the international mean was 500, and the standard deviation 100. On this occasion New Zealand students were ranked third, behind Finland and Canada, and ahead of UK, USA, Sweden, Japan, plus 25 other countries – and marginally ahead of Australia. Once again, New Zealand produced the highest percentage (19%) of Level 5 students (that is, those in the top 10% internationally). In mathematical literacy, the country’s students were again ranked third out of 31 OECD countries, and in scientific literacy they were ranked sixth (OECD, 2002). On the basis of this survey, New Zealand students were achieving at very high levels in all three of these fundamental skills.

On the debit side, New Zealand again showed a very wide range of scores in reading literacy. It had relatively large gender differences and similarly large differences between mainstream and second language learners. The long “tail” identified in reading literacy in 1990 was still there and generally composed of the same kinds of students (Sturrock & May, 2003).

To illustrate the magnitude of these differences, Table 2 shows the mean scores for each gender and ethnic group in the reading literacy scale of New Zealand students.

Table 2. PISA reading literacy means, by gender and ethnic group (Sturrock & May, 2003)

Ethnic Group	Girls	Boys	Difference
Pakeha	577	532	45
Maori	505	459	45
Pasifika	479	445	34
Asian	541	485	56

Analysis of scores in relation to other background factors showed significant differences between high, middle and low decile groups (with means of 567, 519 and 479 respectively), similar differences by levels of parent occupation (means of 574, 548, 520 and 490), differences by type of school, (means of 576, 539 and 516 for girls’, boys’ and co-educational schools), and differences by size of community (means of 533, 520 and 516 for major urban, secondary urban and rural).²

² Many more findings of literacy achievement in relation to home and school factors are reported in Sturrock and May (2003), or at the New Zealand Ministry of education website at:

PIRLS SURVEY, 2001

In 2001, IEA organised a survey of the reading-literacy levels of 9 to 10-year-olds in 35 countries. This survey was planned as one of a series and was named the Progress in International Reading Literacy Study (PIRLS). This time, two surveys were conducted. The first one repeated the 1990 survey in nine of the countries that took part at that time, using the same tests and questionnaires, and administered to students of the same age as before. This study was designed to assess changes in student achievement over time in these countries. The second survey used a new set of instruments in all 35 countries. As before, the tests and questionnaires for the latter survey were prepared collaboratively and carefully translated and pilot tested before use.

The first survey showed that in 2001, New Zealand 10-year-olds scored at virtually the same level as their counterparts did in 1990, in spite of a 50% increase in the number of ESL speakers in the 2001 sample, and in spite of a number of other social and economic changes which have impacted negatively on schools during the 1990s.

This finding is an interesting one when considered in relation to other studies which show a very stable level of performance over time in New Zealand students' reading. For instance, in 1968 a nation-wide survey was conducted by the New Zealand Council for Educational Research (NZCER) to establish norms for new Progressive Achievement Tests (PAT) in Reading Comprehension and Vocabulary (Elley & Reid, 1969). When the tests were revised in 1990, careful comparisons were made of student performance on the unchanged questions. This analysis showed only negligible differences between 1968 and 1990 on the items common to both tests (Elley, 1993). Tables 3 and 4 show the pattern of results for comparisons made at each grade level.

Table 3. Changes in mean percentage correct in the common items of the PAT reading comprehension tests in 1968 and 1990

Year level	No. of Common Items	Mean Gain/Loss
Year 5	29	Loss 1.2%
Year 6	17	Gain 2.5%
Year 7	27	Gain 3.0%
Year 8	26	Gain 4.1%
Year 9	35	Gain 0.9%

<http://www.minedu.govt.nz/index.cfm?layout=document&documentid=6489&data=1>. The international report is available at the PISA website: www.pisa.oecd.org.

Table 4. Changes in mean percentage correct in the common items of the PAT reading vocabulary tests in 1968 and 1990

Year Level	Test Form	No. of Common Items	Mean Gain/Loss
Year 5	Form A	38	No change
	Form B	37	Loss 0.5%
Year 6	Form A	37	Loss 2.3%
	Form B	36	Gain 0.3%
Year 7	Form A	50	Gain 2.0%
	Form B	46	Gain 2.9%
Year 8	Form A	44	Gain 1.1%
	Form B	51	Loss 1.8%
Year 9	Form A	53	Loss 2.9%
	Form B	48	Loss 3.1%

Since these tables showed that there were eight small gains and six small losses, it was reasonable to conclude that standards had not changed in the intervening 22 years between surveys. There was no improvement and no decline in reading standards from 1968 to 1990. The 2001 PIRLS survey showed, in addition, that standards had not changed at the 9-10 year level since 1990. In other words, the best objective evidence available indicated that New Zealand standards of reading, at primary school level, were the same in 2001 as they were in 1968. *One had to conclude, then, that New Zealand reading standards had been remarkably stable for over 30 years.*

The second PIRLS survey of 2001 showed, however, that New Zealand 10-year-olds were ranked only 13th out of the 35 countries that participated. This result was not expected! While the New Zealand mean score of 529 was well above the international average of 500, it was considerably lower than those of several other countries which New Zealand had surpassed in previous surveys. The highest means were obtained by Sweden (561), Netherlands (554), and England (553). Once again, New Zealand had a large group of very high achievers (17% in the top 10%) but also a very large spread of scores.

In New Zealand, there was considerable media reporting of a supposed decline, based on this fall in its ranking relative to other countries' achievement levels. While the main purpose of IEA surveys is not to compare rankings, the fact that the media and the general public are very interested in such comparisons cannot be ignored. However, the validity of these rankings depends entirely on a number of assumptions which are sometimes difficult to satisfy in these surveys. Indeed, a close analysis of the technical report for PIRLS (Martin et al, 2003) throws some doubt on the validity of these survey assumptions and consequently of the international rankings based on the PIRLS country means.

PROBLEMS IN PIRLS SAMPLES

Several countries in the main PIRLS were found to have changed their rank position markedly in relation to their ranking in previous surveys that they had participated in.

For instance, the Netherlands changed from 21st (in the IEA survey of 1991) to 2nd position, England changed from 17th (IEA, 1996) to 3rd, Latvia jumped from 28th (in PISA, 2000) to 5th, and France dropped from 4th (IEA, 1991) to 18th. New Zealand was not the only country to show a marked change in rank. Were these substantial changes in rank order really unexpected? One way to find out is to compare the rank order of countries in successive surveys.

Table 5 shows the pattern of rank order correlations between countries in the four most recent surveys of reading literacy. Clearly, there is usually a high degree of agreement in the rank order of countries in these surveys. Scandinavian countries, New Zealand and Italy typically achieve at high levels, many central European countries score in the middle ranges and developing countries have lower means. The correlations between national means in successive surveys are normally between 0.5 and 0.8. (In IEA mathematics and science surveys, the correlations are even higher.) However, in the case of PIRLS, the correlations with results from other surveys are close to zero. There is an apparent randomness in the changes in rank order which throws doubt on their validity. Either there have been some extraordinary changes in national policies which have caused these changes, or there were problems in the data collection or analysis.

Table 5. Correlations between national means in recent international reading surveys

	IEA 10yrs.	IEA 14 yrs.	PISA 15 yrs.
IEA 10 yrs	----	0.79	0.62
IEA 14 yrs	0.79	----	0.50
PISA 15 yrs.	0.62	0.50	----
PIRLS 10 yrs	0.18	0.13	-0.03

A close study of the small print in the international reports reveals some likely explanations for the apparent randomness in the rank orders. There were two significant problems in the samples chosen to represent each country. One was the difficulty encountered in persuading schools to participate. The other related to large differences in the mean ages of the samples that actually participated. These problems are elaborated on below.

SELECTING A REPRESENTATIVE SAMPLE

For the first time in these international surveys, there were serious difficulties in selecting an adequate sample of students in several of the top countries. As in previous surveys, sampling plans for each country were developed and approved by a sampling referee. Schools were selected in each country in such a way as to produce sound representative samples of students. However, in several of the high-scoring countries, large numbers of school principals refused to participate. According to the figures provided in the Technical Report (Martin et al., 2003), 47% of schools in the Netherlands, for instance, declined to participate. There were similar levels of refusal

in several other high-scoring countries, for example, England (41%), USA, (39%), Lithuania (57%) and Scotland (25%).

The normal procedure in such circumstances is to replace these schools with others, which are also chosen at random, from the *same strata*. In some cases, this in practice meant from the same *region*, or *school type* or *administrative category*. These three strata are described as “*soft strata*” as they are usually *not* closely correlated with achievement levels. In New Zealand, there were very few refusals (7%), and replacement schools were chosen from the *same* decile levels. Decile levels are “*robust strata*”, because they are closely related to achievement. (Indeed, the final PIRLS mean scores for the top, middle and low decile groupings were 566, 537 and 483 respectively.) Thus, for the New Zealand sample, the replacement schools are likely to be similar in their performance levels to those that declined. However, when many schools decline and are replaced by others from “*soft strata*”, there is an obvious source of bias. The principal of the second (or third) replacement school, that agrees to participate, is more likely to be proud of the achievement levels of his/her students, usually because these levels are higher. There is good evidence from New Zealand surveys, that it is low decile schools that most often decline to participate in primary school reading surveys. Of course, the sampling referee for IEA studies does request that checks be made to compare the performance levels of the replacement schools with others in the same strata, but one cannot infer from such checks how well the declining schools may have performed.

That was not the only sampling problem. Many countries were unable to fill their required quota of schools, even after two rounds of replacement invitations. In England, for instance, the replacement schools were chosen from robust strata, but the researchers could still not persuade enough school principals to participate, so they conducted the survey with a sample that had 19 schools missing. The National Report for England described two checks on the representativeness of their sample. When a check was made against the schools’ achievement levels on the nationwide Key-Stage English tests, given 12 months later to these same schools, it appeared that low achieving schools were under-represented in the PIRLS sample. However, the same report also claimed, on the second check – the eligibility for free school meals criterion – that higher socio-economic schools were also under-represented.

On the face of it, there were important gaps in their sample, gaps that would have reduced the spread of their students’ scores and had unpredictable effects on mean scores. It will be recalled that England rose in the rank order from 17th in the previous IEA survey (Brooks et al., 1996) to 3rd place in PIRLS. There may well have been improvements in England’s reading levels, following a strong emphasis on literacy instruction in recent years, but the reported gaps in the sample raise doubts about the true reasons for the change.

Other high-scoring countries with significant numbers of schools missing were USA (26 schools), Netherlands (16), Canada (28), Scotland (37) and Lithuania (15). These gaps are all large enough to raise the mean scores and to reduce the standard deviation. As usual, New Zealand had an excellent sample, with no missing schools.

There were other sampling problems reported in the Technical Manual (Martin et al, 2003). Canada, for instance, did not include its lower-achieving eastern provinces and

Lithuania excluded its second language students. The inevitable “grey areas” produced when countries decide on which ESL and special-class students to exclude from their samples are bound to have unpredictable effects on national means and standard deviations. Countries which mainstream most of their handicapped students, and which endeavoured to include all students who could possibly take the tests, as New Zealand did, are bound to finish up with a much wider standard deviation than those countries with more restrictive selection policies.

AGE DIFFERENCES IN THE STUDENT SAMPLES

The second major problem that affected the validity of the PIRLS rankings was the substantial differences in the mean ages of the national samples of students who participated. In some countries the average age was 9.8 years, while in others it was 11 years. While there is usually much discussion, during the planning stages, about how to define the student sample in such a way as to allow all countries to select students of similar age, the traditional differences between countries in the dates of the school year, in the structure of school grades, in the promotion policies adopted, and in the age of beginning school – all such differences combine to challenge survey organisers.

Nevertheless, in most of these international surveys, the final samples are reasonably similar in average age. In the 1990 IEA survey, for example, there were only small differences in the top 10 countries in this respect. A careful analysis of the effect of these changes showed that they had negligible effects for all except two countries. *This was not the case in the PIRLS survey.* Five of the top eight countries had an average age over 10.7 yrs, compared with 10.1 yrs. for New Zealand and 9.8 for Italy. The best evidence from the International Report shows that a difference of 12 months is equivalent to a difference of 41 points on the PIRLS scoring system. In other words, if New Zealand students had been as old as the students in the top five countries, they would have produced a mean score of 548 points and been ranked in the top five countries.

Since the main purpose of IEA surveys is not to rank countries in this way, it is not surprising that the IEA researchers refrained from making any such adjustment. Nevertheless, that fact that desirable conditions of equivalent age and adequate sample were not met should alert one to the folly of trying to compare countries without a proper examination of the technical difficulties involved in conducting international studies of this sort.

Another way to show the effect of differences in mean age is to correlate the results in successive surveys, across countries, before and after adjusting for age. When the 1990 IEA national means (9 to 10-year-olds) and the PIRLS national means are compared without adjustment, the correlation between country means is 0.25. When both are adjusted for age differences, the correlation rises to 0.62. A more stable result is obtained when like is compared with like. It seems that, where substantial age differences exist, such an age adjustment is necessary, if country rankings are to have any validity.

Some educators may equally wonder why another adjustment should not be made for differences in the students' number of years at school. Several European school systems delay formal schooling until age 7, while others, including New Zealand, start at age 5. However, successive recent surveys of reading have shown no correlation between age of beginning instruction and achievement levels. It appears that many systems start teaching reading in pre-school, and a seven-year start has not been found to be a disadvantage in reading ability by the time students are 10 years old.

In the light of these sampling problems, it is clear that the rank order of countries in the PIRLS survey should not be accepted at face value. Unlike many of the high-scoring nations, New Zealand had a very good sample, including students from all socio-economic levels in proper proportions. Furthermore, its students were considerably younger than those in most of the countries that were ranked in the top ten. If these two factors had been taken into account, New Zealand students would still have been ranked in the top five countries and possibly higher. On this basis, there is no good reason to suppose that New Zealand students' reading levels have declined.

OTHER FINDINGS FROM PIRLS

The technical difficulties in sampling reported above do not undermine many of the other findings in PIRLS. The large gender differences, for instance, were unaffected by sampling problems. New Zealand boys are not achieving as well as girls at primary school level, in any of the ethnic groups, and the major reasons for this discrepancy need further investigation. Moreover, there are unacceptably large gaps between its ESL and mainstream students, and between various ethnic groups, and these gaps are being addressed vigorously in recent years. As in earlier surveys and those conducted by the National Education Monitoring Project (NEMP, 2002), PIRLS results for New Zealand confirm that Maori and Pacific Islands boys are the most at risk. The mean score for all Pakeha/European students was 550, followed by Asian students (540). By contrast, the mean score for both Maori and Pasifika students was 481, which was below the international average of 500. The mean score for Maori and Pasifika boys was 466 and 465 respectively. These are groups that need strong literacy support.

An innovative move in PIRLS was the use of a questionnaire for parents of students selected for each national sample. Although it was difficult to obtain a full response in most countries, the trends amongst those who replied revealed some interesting findings. In response to a question on book ownership, there was a predictable, moderate relationship between the number of books in the home and the achievement level of children. However, when the question asked for the number of *children's* books in the home, there was a much stronger correlation – almost 0.70, across more than 100,000 students, in over 30 countries. Ten-year-olds who have ready access to many children's books at home become good readers. Although the direction of causality in this relationship is still debatable, it does gain some clarity from the positive evaluations of Alan Duff's "Books in Homes" programme, a New Zealand initiative designed to increase the number of children's books in the homes of children in low decile schools (Elley 1997, Croft & Dunn, 2002). Book access and

book ownership are key variables in raising reading levels, and both are variables capable of being changed by appropriate policies and practices.

Other positive findings noted in the PIRLS international report are positive links between student achievement and such factors as:

- Amount of independent reading
- Parents reading to children at home
- Presence of classroom library
- Frequency of reading for fun
- Number of years preschool education.

Many of these factors continue trends identified in earlier IEA surveys, and it is useful to have them confirmed on larger numbers of students.

Just as helpful are results that show which factors do *not* appear to make a difference. For instance, there appears to be no achievement advantage in starting school at five years of age, or in the number of hours spent on reading instruction, or in the frequency of teacher meetings, or in the use of portfolios in reading or in students answering written questions about their reading. Surprisingly, the survey found that New Zealand teachers more often taught strategies for decoding skills than teachers in most other countries.

CONCLUSION

This article has summarised the findings of a series of international surveys of students' reading performance. New Zealand has participated in ten such surveys between 1970 and 2001. In most of these surveys its students have achieved at very high levels. Although the main purpose of these surveys was not to rank countries, there is no denying that such rankings are of great interest to many. If the main assumptions are met – assumptions of fair, representative samples of students, of similar age, who are tested by suitable tests, under similar conditions of testing, then the rankings *do* indicate relative levels of success in teaching children to read.

In these ten surveys, New Zealand students have been ranked 1st in the first four surveys, 3rd on two subsequent occasions, 4th on one, 6th on two, and thirteenth on one – the PIRLS survey of 2001. On all except the last survey, New Zealand students showed a very high standard when compared with students of similar age in other countries.

In view of the fact that the first of the two PIRLS surveys showed a stable standard of reading from 1990 to 2001, and in view of the problems described above in securing fair samples in some top countries in the second PIRLS study, and of the large differences in mean age of the samples tested, it is reasonable to conclude that New Zealand students have been and are still amongst the best readers in the world.

New Zealand schools consistently produce the most excellent readers, and their attitudes to reading are generally positive. There are still serious problems to address amongst Maori and Pasifika boys in particular, but there is clear evidence from these surveys that, by and large, New Zealand teachers are very good at teaching reading.

REFERENCES

- Ashton-Warner, S. (1963). *Teacher*. New York: Bantam Books.
- Brooks, G, Pugh, A., & Schagen, I. (1996). *Reading performance at nine*. London: National Foundation for Educational Research.
- Croft, A., & Dunn, K. (2002). *An evaluation of the "Books in Homes" programme*. Wellington: New Zealand Council for Educational Research.
- Education Department. (1985). *Reading in the junior classes*. Wellington: Education Department.
- Education Today. (2003). Big drop in New Zealand children's reading performance. *Education Today*, 2, 30.
- Elley, W. (1992). *How in the world do students read?* The Hague: IEA.
- Elley, W. (1993). *How well do New Zealand students achieve, by international standards?* Paper prepared for New Zealand Post Primary Teachers Association, Christchurch: University of Canterbury.
- Elley, W. (Ed.). (1994). *The IEA study of reading-literacy: Achievement and instruction in 32 school systems*. Oxford: Pergamon Press.
- Elley, W. (1997). *An evaluation of Alan Duff's "Books in Homes" Programme: Final report*. Wellington: Ministry of Education.
- Elley, W., & Reid, N. (1969). *Progressive achievement tests of reading: Teachers' manual*. Wellington: New Zealand Council for Educational Research.
- Holdaway, D. (1979). *Foundations of literacy*. Gosford, NSW: Ashton Scholastic.
- Martin, M., Mullis, I., & Kennedy, A. (Eds.). (2003). *The PIRLS technical report*. Chestnut Hill, MA: Boston College.
- Ministry of Education. (2003). *Reading literacy in New Zealand: Final results from PIRLS 2001*. Wellington: Research Division, Ministry of Education.
- Neuman, S., & Koskinen, P. (1992). Captioned television, as comprehensible input. *Reading Research Quarterly*, 27 (1), 94-106.
- NFER. (2003). *Progress in international reading literacy (PIRLS) study: Reading all over the world*. Retrieved 6 April, 2004 from <http://www.standards.dfes.gov.uk/literacy/publications/research/63521/>
- OECD. (2001). *Knowledge and skills for life: First results from PISA 2000: Executive summary*. Paris: OECD.
- Purves, A. (1973). *Literature education in ten countries*. Stockholm: Almquist & Wiksell.
- Purves, A., & Elley, W. (1994). The role of home and student differences in reading performance. In W. Elley (Ed.), *The IEA study of reading literacy: Achievement and instruction in 32 school systems*. Oxford: Pergamon Press.
- Sturrock, F., & May, S. (2002). *PISA 2000: The New Zealand context. Report of results from the Programme for International Student Assessment*. Wellington: Ministry of Education.
- Thorndike, R. (1973). *Reading comprehension achievement in fifteen countries*. Stockholm: Almquist & Wiksell.
- Wagemaker, H. (1993). *Achievement in reading-literacy: New Zealand performance in an international context*. Wellington: Ministry of Education.