Papua New Guinea Journal of Education

ISSN: 0031-1472 Volume 43, 2023

THE EARLY GRADE READING ASSESSMENT (EGRA) SURVEYS: A CLEAR ADVANTAGE SHOWN FOR VERNACULAR-BEFORE-ENGLISH LITERACY IN PNG SCHOOLS

ROBERT PETTERSON

SIL-PNG

www.pngjedu1472.com e-mail: pngjeducat22@gmail.com

THE EARLY GRADE READING ASSESSMENT (EGRA) SURVEYS: A CLEAR ADVANTAGE SHOWN FOR VERNACULAR-BEFORE-ENGLISH LITERACY IN PNG SCHOOLS

ROBERT PETTERSON

SIL-PNG

robbie_petterson@sil.org

ABSTRACT

Choosing an appropriate language for teaching a child to read for the first time is important for educational success. From 2011 to 2013 the Read-PNG project surveyed 5,500 school children in four provinces of Papua New Guinea using the Early Grade Reading Assessment (EGRA) tool, and published the analysed data in four separate reports in 2014; but no inter-provincial comparison was made. This paper carries out that lacking comparison, grouping the data under three literacy methodologies: vernacular literacy, bridging from vernacular to English literacy, and English-only literacy. Doing this reveals compelling evidence that, by grade 4, children in schools where literacy is taught in a child's mother tongue or home language first have far higher *English* literacy skill levels than those in schools where they are taught *only* English literacy (while still learning English as a second language). This finding has important implications for education policy makers in multilingual societies.

Keywords: literacy, PNG, indigenous languages, mother tongue literacy, vernacular literacy, multilingual education, language of instruction, EGRA, Read-PNG, literacy survey

1 INTRODUCTION

This paper compares survey data on the acquisition of literacy in schools in Papua New Guinea under different teaching methodologies, with a view to understanding the actual effectiveness of various educational policies, teaching practices and sociological factors observed.

In Papua New Guinea the language of higher education is English, but many children in the lower grades are only beginning to learn to speak English and are struggling to learn to read it. These struggles were demonstrated in extensive research carried out by the National Department of Education (NDoE) with support from the World Bank in a project called Read-PNG. From 2011 to 2013 Read-PNG conducted detailed surveys of 121 schools spread over four provinces located in diverse parts of the country. These surveys included literacy skill assessments of 5,500 children, and "observation snapshots" of 283 classrooms. Read-PNG published separate survey reports for each of the four provinces in 2014. Although an examination of the data in the reports shows important and significant differences between the provinces that could inform educational policy, Read-PNG did not publish an analysis of those differences. This paper seeks to fill that gap, and to answer the question: how best can children who speak a minority indigenous language at home become literate in a national or world language such as English?

This paper begins with a brief history of language in education in Papua New Guinea, followed by a description of the work of the Read-PNG project and a summary of their survey instruments and survey methodology. The main body of the paper continues with an explanation of the methodology that I used for carrying out an inter-provincial comparison, followed by graphical presentations and discussions of the results of that comparison, especially the significant differences in literacy skills test score means and the numbers of zero-scorers in those tests, which evidently arose from methodological differences that were revealed in teacher and classroom factor analyses. The paper concludes with a summary of the best practices based on the evidence provided by the survey data.

2 LANGUAGE IN EDUCATION IN PNG

As Wroge (2002) and Litteral (2015) explain, starting in the 1870s, missions and churches took the lead in teaching indigenous people to read in some of the local languages or lingua franche, following which limited numbers of students would enter a more formal English or German system. In the early 1950s the colonial government instituted an English-only

education policy developed in Australia. An informal scheme of community-oriented vernacular pre-schools started up in the North Solomons Province in 1979, and an educational advantage was noted for the students affected. These schools proliferated in other areas of the country. In the 1990s the scheme was formalised in a national educational reform, with vernacular literacy to be taught for the first three years before introducing English literacy on entry to primary school in grade 3.

In the early 2000s a teaching methodology called Outcomes Based Education (OBE) was introduced for both elementary and primary schools, but after training and implementation problems (Sinai, 2012), a political decision was made to stop using both OBE (Dept told to act on directives, 2013) and vernacular education in response to widespread public concern that "blamed" poor English standards on "the use of vernacular in schools" (Taita, 2013). This decision was made while the Read-PNG surveys were still being carried out, and, unfortunately, before the survey reports had been published that could have provided wisdom to inform any decision about education policy.

3 THE READ-PNG PROJECT

The Read-PNG project was a 19.2 million USD project funded by the Global Partnership for Education, and carried out by the NDoE with technical support from the World Bank. The objective of the project was "to promote better teaching and learning of reading skills of elementary and primary education" (The World Bank, 2016: vi).

As mentioned above, Read-PNG conducted large-scale literacy surveys of early grade school children in four provinces of Papua New Guinea (PNG). These surveys were carried out as a diagnostic and baseline component for future remedial work planned for a subsequent phase of the project.[1] The data from the surveys were analysed and published as large reports in 2014 (Machuca-Sierra & de Silva, 2014a-d).

The provinces surveyed by Read-PNG were East New Britain (ENB), Madang (MAD), the Western Highlands Province (WHP), and the National Capital District (NCD) (which is centred around the city of Port Moresby). Their locations are shown in the map in Figure 1. Note that these provinces represent four major regions of PNG, namely the islands, the north coast, the highlands, and the south coast.

[1] The original plan was to survey 15 provinces, but this proved over-ambitious (The World Bank, 2016:8). Read-PNG later carried out "Reading Booster" intervention programs in two provinces, Madang in 2013 and Western Highlands in 2014, and distributed reading materials to 7,418 schools in all 22 provinces (The World Bank, 2016:11; Birney, 2016: slide 10).

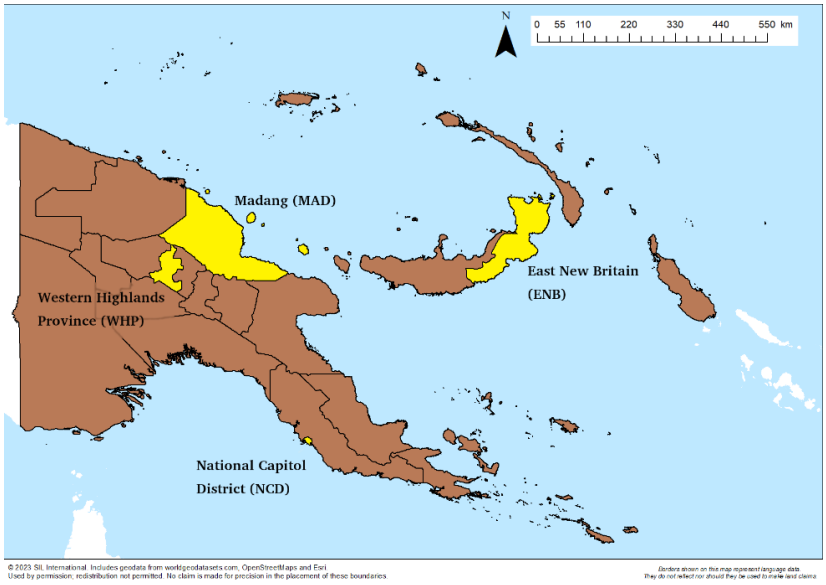

Figure 1. Map of Papua New Guinea, showing the four provinces surveyed by Read-PNG in yellow.

3.1 The Read-PNG survey instruments

There were two parts to the survey. One was quantitative, measuring the literacy skills of each student in a selected class; the other was qualitative, investigating factors in the classroom, or in the teacher's or child's background that might account for how well (or not) those literacy skills had been acquired.

The literacy skills measurements used a survey instrument called Early Grade Reading Assessment (EGRA), which has been used to assess literacy education in many countries around the world. It was designed by RTI International, a non-profit research organisation (see RTI International, 2015). In PNG, the EGRA assessments included a battery of nine tests that covered the first three stages of literacy acquisition, that is, Emergent Literacy, Decoding, and Confirmation and Fluency (see Dubeck & Gove, 2015).

The nine tests assessed the following literacy skills:

(1) Listening Comprehension (students orally answer questions about a story told to them)

(2) Phonemic Awareness (students identify the first sound of some spoken words)

(3) English Letter Names (e.g. *y* is called *wye* [wai])

(4) Letter sounds (e.g. *y* often has a consonantal sound [j] as heard in the word *yoke*).

(5) Invented word reading (e.g. *speng*, *vock*)

(6) Familiar word reading (e.g. *spin*, *rock*)

(7) Reading fluency (students read as much as they can of a story in a set time)

(8) Reading comprehension (students answer questions on what they have just read)

(9) Dictation (students write down a story as it is told slowly to them).

Versions of this test bank were prepared and piloted in five languages: English, Kuanua, Adzera, Hiri Motu and Tok Pisin (personal communication Machuca-Sierra, 5 Sep, 2020). EGRA assessors were to administer these tests during twenty-minute face-to-face interviews with each individual student.

Read-PNG also devised questionnaires for collecting qualitative data about each student, and his or her teacher in interviews, and for recording "classroom observation snapshots" (COS) for at least one classroom in each grade of each school surveyed.[2]

3.2 Read-PNG survey target selection.

The early grades targeted for survey were Elementary Preparatory, and Grades 1 and 2 in the elementary schools (abbreviated as EP, E1 and E2, respectively) and grades 3 and 4 in the primary schools (abbreviated as P3 and P4). It had been expected, according to the National Plan for Education at the time (Department of Education 2004:27–28), that children were being taught literacy in a local vernacular in EP, E1 and E2, and were then being bridged to English literacy in P3, and EGRA assessments are designed to test reading skills at these levels. Initial pilot surveys found, however, that most of the elementary schools surveyed in Madang, and all of such schools in NCD and WHP, were teaching in English instead. Consequently vernacular literacy testing was carried out only in ENB, using the Kuanua version of the EGRA instrument.

Even though the EGRA assessments were designed to test reading skill outcomes specified for EP and E1 in the elementary syllabus, the lower grades to be targeted for survey had to be dropped in some cases, because, as explained in the Madang report, "initial pilot studies found [that] 'floor effects' in Grade 1 indicated not enough students were able to complete the test. Therefore testing began from Grade E2 [in Madang]" (Machuca-Sierra & de Silva, 2014a: 36). The same was found in the other mainland provinces, and only in ENB were there sufficient numbers of children capable of performing well enough in the tests for E1 to be included in the survey there.[3]

[2] Most of the factors for which data were collected are listed in Appendix B.

[3] It could be argued that the delay in numbers gaining literacy skills in Madang, WHP and NCD might be because the teaching of literacy was delayed until E2 in those provinces. This is most unlikely; my own observations of elementary schools in another province, the Gulf Province, found that when elementary teachers stopped teaching literacy in a local vernacular because of community pressure or a change in policy, they started applying the same principles they had learnt in their vernacular literacy training to teach English literacy. That this effort largely failed in EP and E1 was not just because teachers lacked proficiency in English and had had no training to teach English literacy; rather it was very evident that it was also pedagogically unsound to attempt it – children simply were not able learn to read in a language that they could not yet speak. No matter how much

The essential details of the surveys in each of the four provinces are set out in Table 1. It should be noted that the surveys were *not* a longitudinal study (following the same students up through the grades); rather, the children of all three or four grades of each school were surveyed at the same time. It should also be noted, however, that surveys in two of the provinces were conducted in May, and in the other two in October. The possible affect of this timing difference on a comparison of survey results is discussed in section 4.4 below.

Table 1. Details of the four EGRA surveys

Province	Schools	COS*	Students	Grades	Survey date
Madang (MAD)	20	20, 45	1,279	E2, P3, P4	Oct 2011
National Capital District (NCD)	20	26, 56	1,266	E2, P3, P4	May 2012
East New Britain (ENB)	41	34, 50	1,696	E1, E2, P3, P4	Oct 2012
Western Highlands (WHP)	40	18, 34	1,323	E2, P3, P4	May 2013

* Under COS are the numbers of elementary and primary classrooms where classroom observation snapshots were made.

4 INTER-PROVINCIAL COMPARISON OF LITERACY TEST DATA

4.1 Methodology

The four Read-PNG provincial survey reports have similar structures and sets of data, since the same English EGRA instrument was used in all four provinces, and the Kuanua instrument used in E1 and E2 in ENB was analogous to the English one.[4] The data in the reports includes the numbers of students tested in each grade, the means for each grade, along with standard deviations, 95% confidence intervals, and maximum and minimum scores. Also included are the numbers of students who scored above zero in a test, the means with all the zero-scores removed, and sometimes medians. I collated the relevant data from the reports and prepared graphs that allow the results of each province to be compared.[5]

4.2 Results – test score means

In this section I compare the literacy skill levels in all four provinces using graphs that have four trend lines – one for each province. Each trend line shows how the level of each literacy

effort the teachers applied, it was only after the children had gained some proficiency in hearing and speaking English (by E2 or P3) that they could start to learn to read it.

[4] The survey reports show calculations of the EGRA instrument's reliability with Cronbach's Alpha coefficients of about 0.9, indicating "good internal consistency" (e.g. Machuca-Sierra & de Silva, 2014a: 19).

[5] I have done the comparisons using whole-sample means for this paper, but comparisons done using medians and means that omit the zero scores show very similar interrelationships between the provinces. (See Birney (2016) for several analogous graphs using medians.)

skill changes from grade to grade. I do this for each of the nine literacy skills in turn, but, first of all, I present a graph that summarises how the provinces compare over all the skills tested.

General summary

A general comparison of all four provinces appears in Figure 2. The trend line for each province shows how the average calculated for all nine test score means changes from grade to grade,[6] and gives a good impression of how each province was performing compared to the others in the teaching of literacy.

Note that the data points for the first two grades of ENB's trend line are marked with a V to indicate that the children had been taught to read in the local vernacular, and that the tests were carried out in that language. (This marking is used in all the relevant graphs below.)

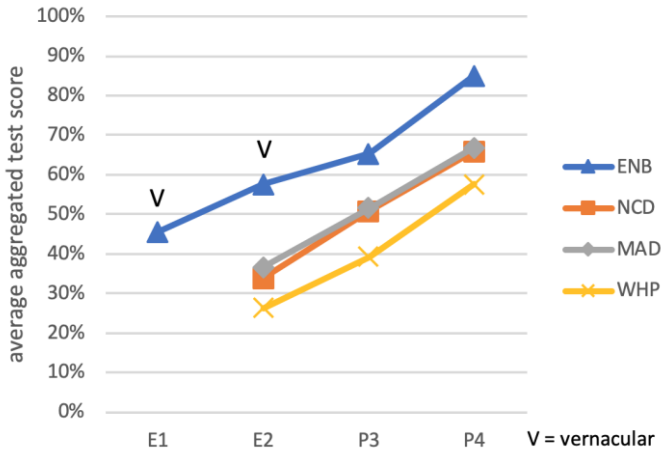

Figure 2. Summary of literacy progress through grades E1-P4, in four provinces

It can be clearly seen that, on average, children in ENB were surpassing those in other provinces in English literacy skills; they were generally at least a year ahead of those in Madang and NCD, and about two years ahead of those in WHP, even though ENB children were starting to learn English literacy three years later than those in the other provinces (counting the first year of school, Elementary Preparatory).

The results for each of the nine literacy skills tests will now be given, comparing all four provinces.[7]

[6] Some test data was in terms of fluencies (items per minute). For the graph in Figure 2, these fluencies were converted to percentages of the maximum mean fluency observed in each test, so that they would be compatible with other percentage-based test score data. The data points in the trend lines were then calculated by averaging the percentage score means for all nine tests. (See right hand column of Table A1 below.)

[7] Collated summaries of the actual test mean data from the reports is included in Appendix A.

Listening comprehension

Figure 3 shows that the ENB children had very good listening comprehension in their own language, as one would expect, and although they got lower average scores for listening comprehension in English in P3, the P4 results showed that they were progressing faster than the children in other provinces.

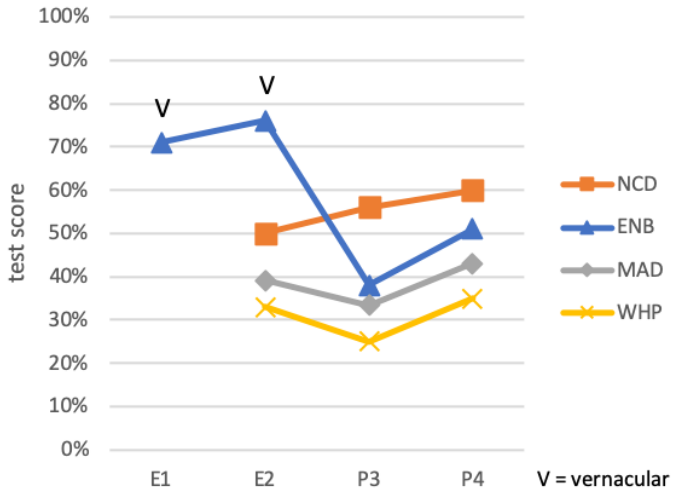

Figure 3. Listening comprehension test score means, grades E1-P4, four provinces.

The children in the NCD may have had an advantage in this skill for English in that they lived in or near the city of Port Moresby, where they may have had more exposure to spoken English.

Initial word sounds

Figure 4 shows that this phonemic awareness skill test was performed very well by ENB children in the Kuanua language, but children in all provinces found it more difficult for English – more than half of the E2 children in WHP scored scored zero in this test.

Figure 4. Initial word sounds test score means, grades E1-P4, in four provinces.

Children in NCD may have done better in this skill than those of Madang and WHP for the same reason explained above, but those from ENB clearly had an advantage by P4 in that they had already learnt this skill well in their vernacular in earlier grades.

English letter naming

Figure 5 shows that children in all provinces did equally well in naming English letters.

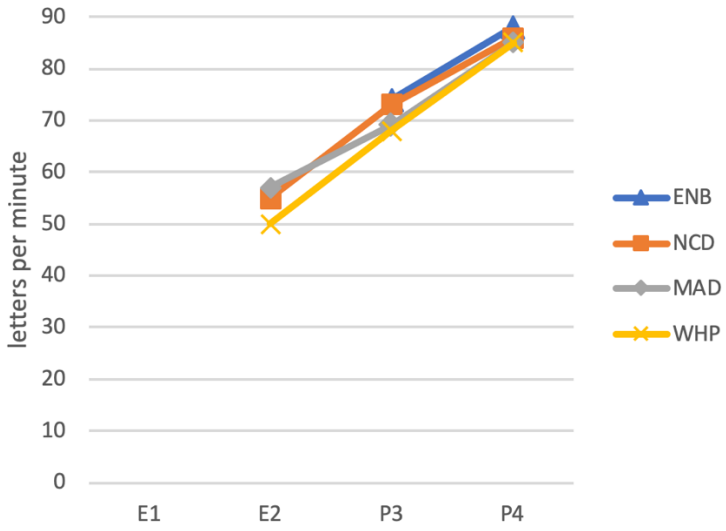

Figure 5. English letter naming fluency means, grades E1-P4, in four provinces.

This test was not carried out in ENB until grade P3, because the children in that province were learning only vernacular literacy in E1 and E2. Despite this, they still did very well at learning the English letters in the primary school grades.

Letter sounds

Figure 6 shows that, on average, ENB children did far better at letter sounds than those of other provinces. This skill, once developed in their own vernacular transferred very well to English after a slight drop in speed at first.

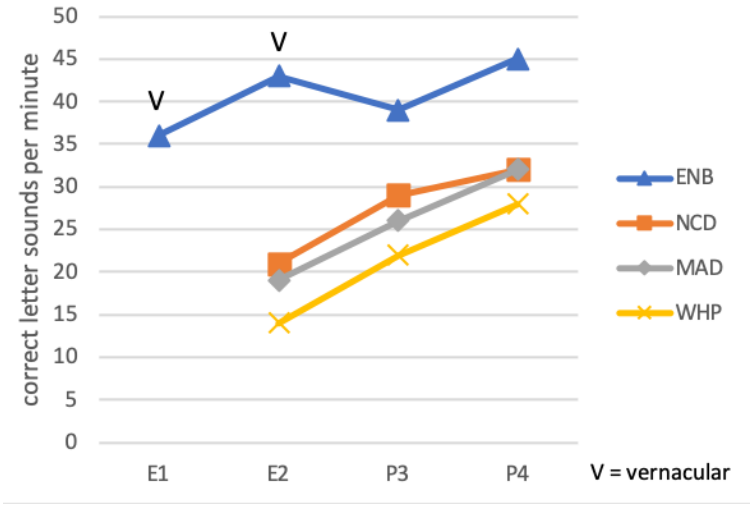

Figure 6. Letter sounds, grades E1-P4, in four provinces

Familiar word reading

Figure 7 shows that ENB children, having learnt how to read words in their vernacular, applied the same skill to reading English words, with no drop in speed at all.

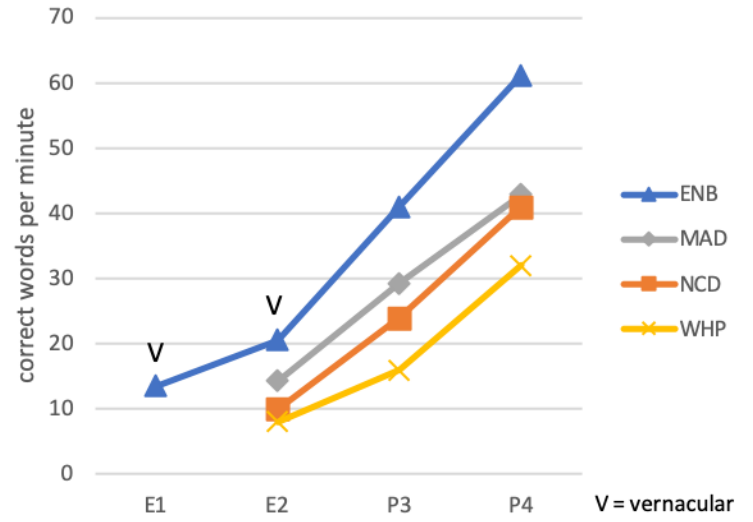

Figure 7. Familiar word reading means, grades E1-P4, in four provinces.

Invented (or unfamiliar) word reading.

Invented word reading requires a knowledge of letter sounds and the ability to blend them, and is an important word-attack skill. Figure 8 shows that ENB children, having acquired this skill in the vernacular, were able to apply it to English to great advantage.

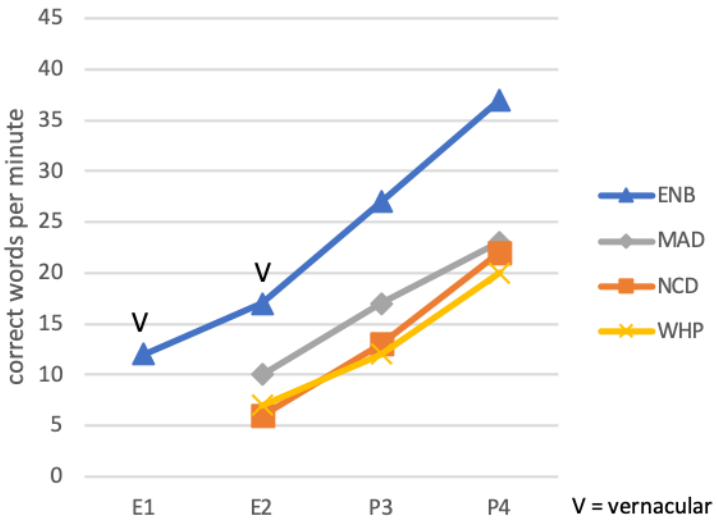

Figure 8. Invented word reading, grades E1-P4, in four provinces.

Reading fluency

Reading fluency is another skill that transfers well between languages. Figure 9 shows that the children of ENB did best in this.

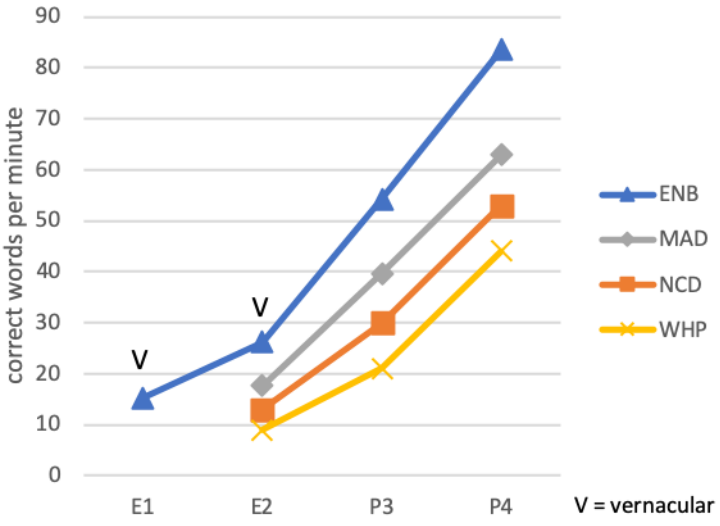

Figure 9. Reading fluency means, grades E1-P4, in four provinces.

Reading comprehension

This is a key skill, as it predicts the ability to understand text books in later grades. Figure 9 shows that, on average, the children in ENB also did best in reading comprehension, while

those in WHP did poorly, although the trends indicate that they would probably be able to catch up with NCD and Madang by P5.

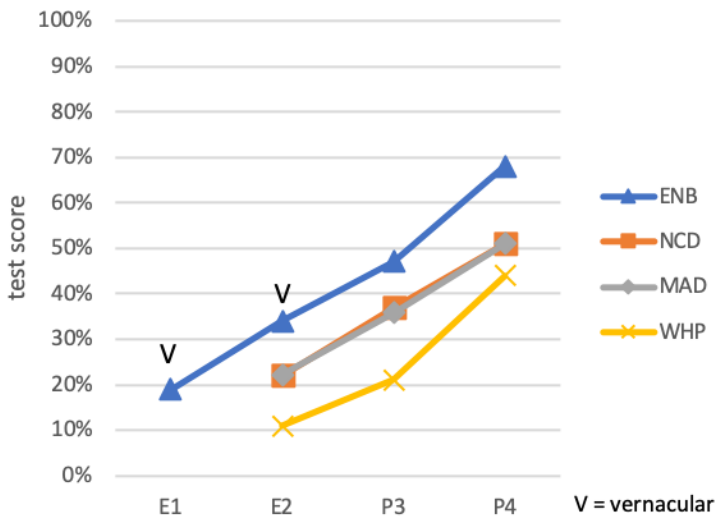

Figure 10. Reading comprehension test score means, grades E1-P4, in four provinces.

Dictation

Figure 11 shows that in elementary school the ENB children did best at dictation (in their Kuanua language), and the Madang children did best at this in English in P3; however, after bridging to English, ENB children had caught up with those in Madang by P4. This is also the only test where WHP clearly outperformed NCD in the lower rankings.

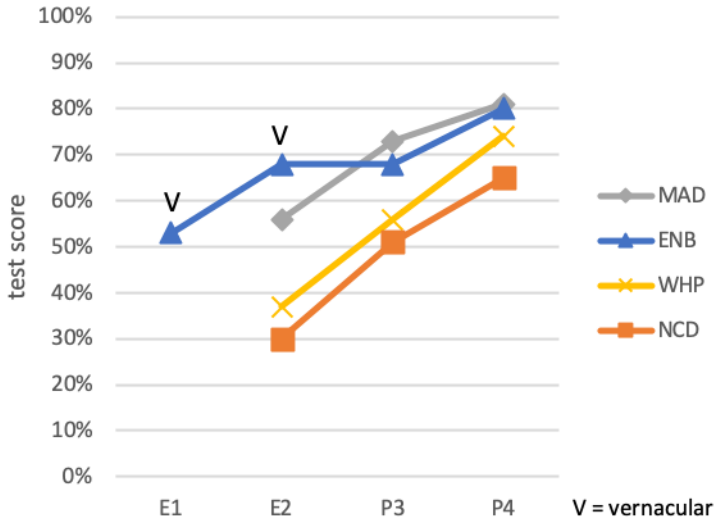

Figure 11. Dictation test score means, grades E1-P4, in four provinces.

In summary, comparing test results across all provinces, we can see that in one test, English letter naming, all did equally well; otherwise the children of ENB had by far the best results

in E1 and E2 in all the other eight skills; they also did best in P3 and P4 in six other skills, getting lower results only in tests that require aural English listening skills. The children of WHP, on the other hand, did poorly (on average) in nearly all the tests, and the children of NCD also generally did quite poorly, doing better than the others only in listening comprehension.

4.3 Proportions of zero-scorers

The proportion of children who score zero in a test can be used as a measure of difficulty that the teachers are having in imparting the skill being tested. Only letter naming was universally well taught, with hardly any zero scorers in any province. In many of the other skills tests, there were large numbers that failed in one or more provinces, especially in E2. This can be seen, for example, in Figure 12, which shows that as many as 64% of NCD students and 40% of WHP students in E2 scored zero in the invented word test.

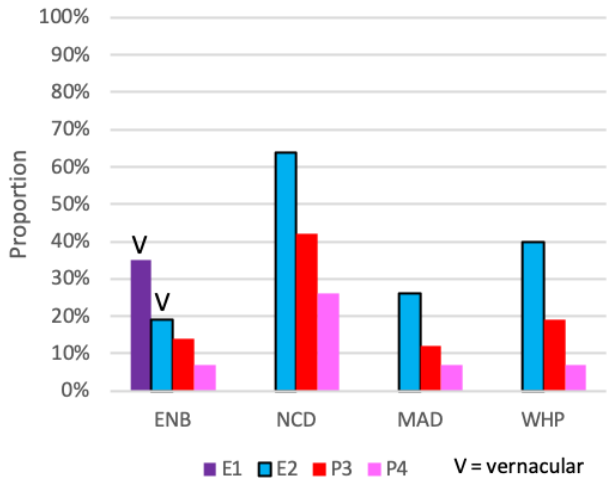

Figure 12. Zero-scorer proportions in the invented word test

The proportions of zero scorers usually decreases for each higher grade, of course, as more and more students grasp the skills being tested.

The skills tests that had particularly high numbers of zero scorers in E2 were reading comprehension (in four provinces), initial word sounds (in three provinces), invented words and reading fluency (both in two provinces). The other tests, except for English letter names, had high numbers of zero scorers only in one province (either NCD or WHP).

In order to understand more clearly the relative degree to which provinces were struggling with imparting literacy skills in schools, I simplified the zero-score data by counting the

number of instances for each province where a test had 30% or more zero scorers from grade 2 up. The results are shown in Figure 13.

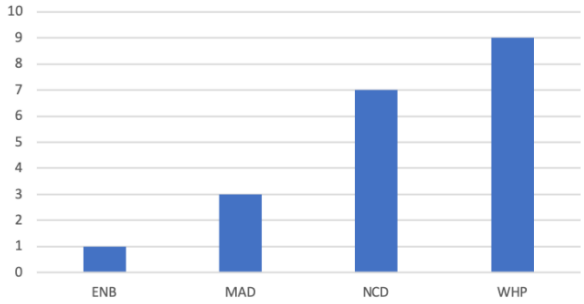

Figure 13. Numbers of tests in each province where 30% or more children scored zero (n=108 tests, grades E2 – P4)

This graph shows that WHP and NCD were the provinces where students struggled to learn the most, while in Madang and ENB, there were far fewer such problems. ENB was doing the best, with only one test with a zero-scoring rate greater than 30% – this was the reading comprehension test in E2, which was difficult for *all* provinces in that grade.

4.4 Statistical significance

The reports indicated statistical significance in terms of 95% confidence intervals. Non-overlapping confidence intervals are conservative evidence that a difference in means is statistically significant. These intervals amounted to between ±1.4% and ±4% of the test score means that were given as percentages. The confidence intervals of, for example, the P4 reading comprehension test score means are ENB (65%, 71%), NCD (48%, 55%), Madang (48%, 54%), and WHP (42%, 47%). In Figure 14, I-bars at the top of each column show the lower and upper limits of these confidence intervals for each province.

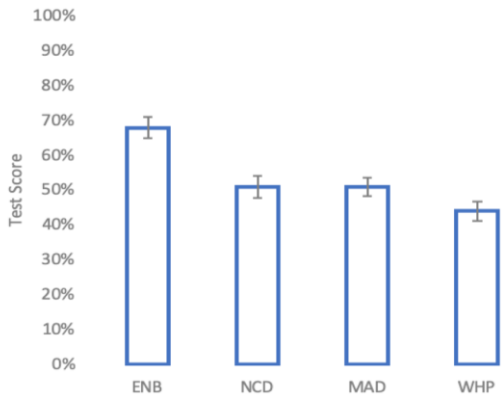

Figure 14. Mean scores and confidence intervals for the reading comprehension test in P4

When the I-bars overlap, as they do for NCD and Madang, this indicates that the confidence intervals are not distinct, and there is no significant difference between the means scores of these two provinces for this test. (They are both 51%.) The I-bar for ENB, on the other hand,

is far above the other I-bars, showing that the difference between its mean score (68%) and that of NCD (51%), for example, is statistically very significant.

Confidence intervals show that the ENB tests results are significantly better than those of other provinces in six of the nine tests at Grade 4 level.[8]

A criticism of this comparison between provinces could also be made concerning the fact that the surveys in ENB and Madang took place five months later in the year than those in NCD and WHP. This does not affect the comparison between ENB and Madang, of course, and it is not clear how much difference this would have made comparing these two provinces with the other two; but if we assume that the learning rate is linear over the 10.5-month school year, then we could calculate slightly increased estimates for the provinces tested in May by adding 5/10.5 (or 47.6%) of the average improvement from grade to grade of those two provinces.[9] As an exercise, this has been done for the general summary of data (depicted in Figure 2 above), and this is shown in Figure 15 below.

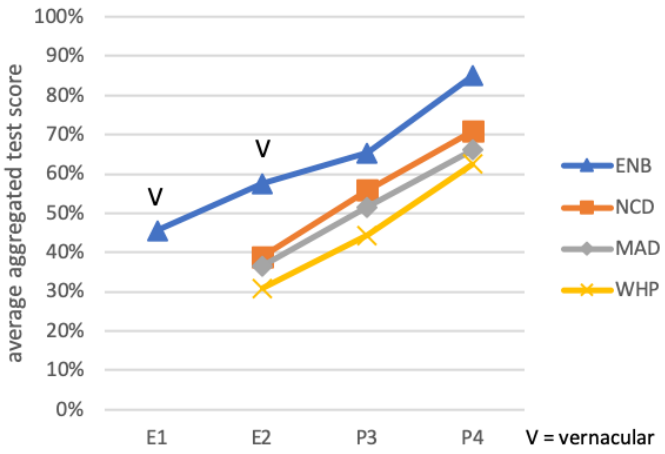

Figure 15. Summary of literacy progress through grades E1-P4, in four provinces, with 5-month estimated adjustment upwards for NCD and WHP

Even with this adjustment, ENB is still achieving well above the others, but NCD now appears to be doing slightly better (on average) than Madang.

[8] Appendix A lists the standard deviations and 95% margins of error used to calculate the confidence intervals for all nine test means for grade P4.
[9] The grade-to-grade average increase in scores amounted to 10.6% for both provinces.

5 INTER-PROVINCIAL COMPARISON OF STUDENT, TEACHER AND SCHOOL FACTORS

5.1 Methodology

As mentioned in §3.2, Read-PNG also explored qualitative factors to do with the student, the teacher, the classroom and the school which might be significantly associated with successful or unsuccessful literacy skill acquisition. There was a pool of more than 70 different factors that Read-PNG investigated in at least one province, and 29 of them were investigated in all four.

For the purpose of comparing between provinces, I had two aims: (1) to find out which factors had statistically significant effects on literacy acquisition in any province, (2) to try to rank those factors in terms of the size of the effect they had on test score means.

Read-PNG calculated the effect of each factor in terms of the expected increase (or decrease) in test performance, on average, when that factor was observed. For example, the Madang report states that "owning the textbook is associated with an additional 8 CWPM [i.e. correct words per minute in reading fluency], [a] 10% increase in reading comprehension and an 8% increase in dictation scores, on average" (Machuca-Sierra & de Silva, 2014a: 40).[10]

For the sake of simplicity, in this paper I have focused on analysing the effects of the factors on the means of just one test, reading comprehension, which is arguably the most indicative measure of successful literacy acquisition used in the surveys. I extracted from the data the effect sizes that Read-PNG regarded as statistically significant (i.e. those with $p \leq 0.10$), and ranked them.

At first I found that some factors were counter-intuitive, or were negative in one province while positive in another; but if I grouped them according to literacy teaching methodology, I found that many of those problems were reduced. The methodological groupings were: (1) vernacular literacy (in ENB), (2) English literacy after vernacular literacy (also in ENB), and (3) English-only literacy (in the other three provinces). In group (3), I calculated an average effect size for each factor over all three provinces before I ranked the factors.[11]

[10] In most cases Read-PNG calculated regression coefficients using least squares or Tobit regression calculations as a measure of the effects of these factors on test score means, and in some cases, especially in MAD, they carried out t-tests between the test score means when certain factors were observed and when not observed. (Tobit regression is used to calculate a regression coefficient when many of the data cluster around a hard limit of the data range, such as when there are many zero scores in the test data. This is explained in the Madang report, WBa: 38, footnote 11.)

[11] To further simplify the comparison I grouped together certain similar factors if they had similar effect sizes. I then averaged the effect sizes of each such grouping. There were three important clusters of factors that I so grouped; one was to do with which members of a student's family were literate (especially "father", "mother",

In the sections below I present the ranked factors and their strengths (as expected test score increases or decreases) for each of these three teaching methodologies in graphical form. The graphs show only the most influential factors for each of the three methodologies. Each graph has the same scale, so that comparison between the methodologies is easier.[12]

5.2 Positive (i.e. helpful) factors

The ten most helpful factors for vernacular literacy in ENB are shown in Figure 16.

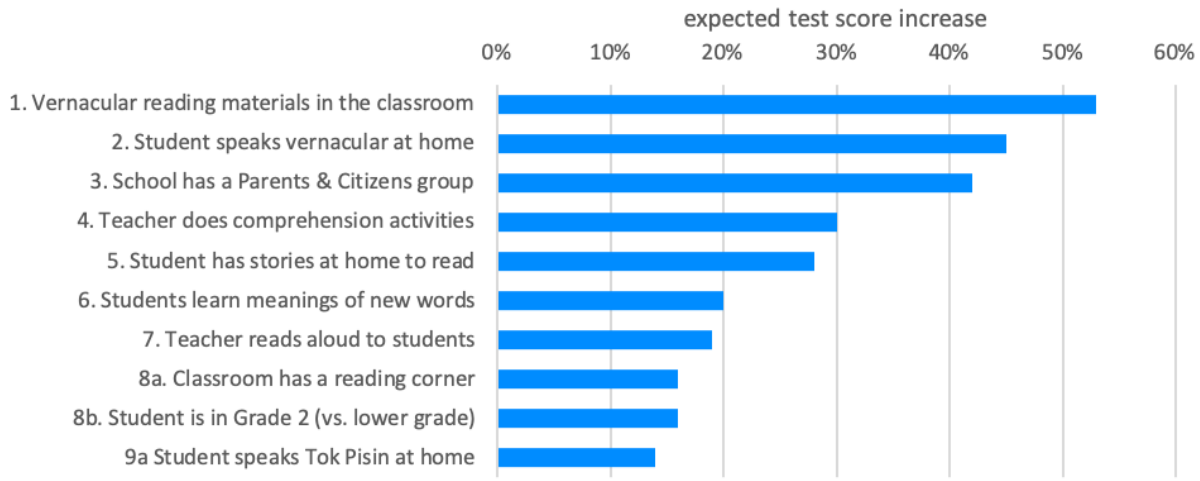

Figure 16. Factors helpful for successful teaching of vernacular literacy in ENB ranked by expected increase above the reading comprehension test score mean

The factors are all very sensible, with the most important being the use of vernacular reading materials in the classroom (observed in 88% of the classrooms of ENB – in some 12% of the E2 classrooms of ENB the focus was evidentially on English rather than vernacular). Having vernacular reading materials at home is also very helpful (factor 5).

Just five additional factors proved to have a positive association with English literacy acquisition following vernacular literacy (as found in ENB) are shown in Figure 17.

or "siblings" etc, which, in some reports, are recorded as just "siblings" or "other family member"); another was to do with the type of story material used in the classroom (i.e. the availability of "Shell and Big books", "graded readers", or "school journals" etc, which, in some reports, are grouped together simply as "reading books" or "printed stories"); and the third was to do with the medium used for story material (i.e. "stories in charts" and "stories on blackboard" which I combined as "stories on blackboard/charts").

[12] Complete lists of the statistically significant factors can also be found in Appendix B.

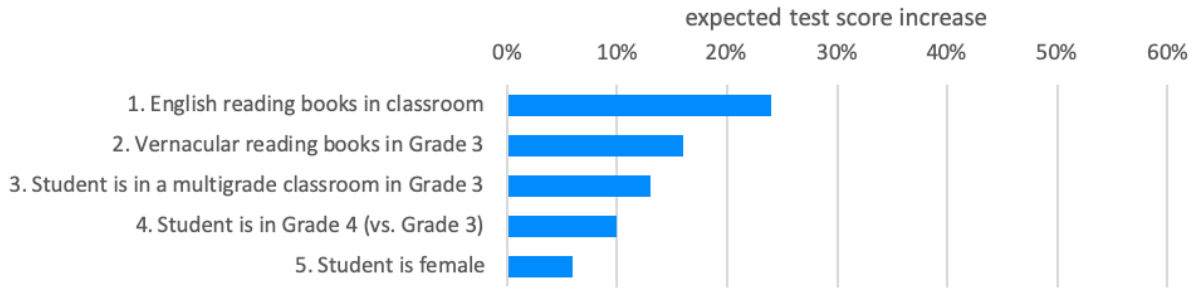

Figure 17. Factors helpful for successful teaching of English literacy after vernacular literacy in ENB, ranked by expected increase above the reading comprehension test score mean

Note the importance of reading materials in appropriate languages, including the use of vernacular reading materials in the bridging year (P3). Even though this was only rarely observed (Machuca-Sierra 2014b: 42), it was evidently effective where it did occur. The beneficial effect of having a multigrade class at this level should also be noted.

The ten most helpful factors for English literacy under an English-only policy (Madang, NCD, WHP) are shown in Figure 18.

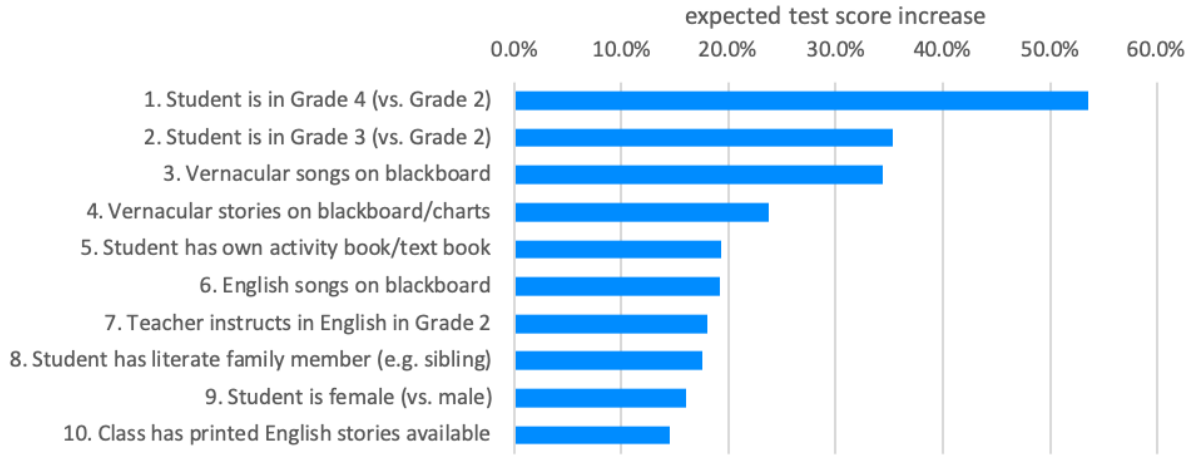

Figure 18. Factors helpful for successful learning of English literacy under a predominantly English-only teaching policy (in Madang, NCD and WHP) ranked by expected increase above the reading comprehension test score mean

Note that factors 3 and 4 in Figure 18 (which are also the first factors in that list that are related to teaching methodology) have to do with *vernacular* reading material. Amongst these three provinces these were observed as significant factors only in Madang.[13] (Vernacular

[13] When factors influencing some of the other tests (reading fluency and dictation) are examined, more detail on the use of vernacular in Madang comes to light; the use of vernacular stories and vocabulary words written on the blackboard and on charts in E2 and P3, and vernacular songs and prayers on charts and vernacular dictionaries and Big Books in P3 are also statistically significant. The use of such materials was reported for between two and six classrooms out of 65 observed in MAD (Machuca-Sierra & de Silva, 2014a: 53-55, 92-95). Birney (2016: slide 8) claims that "data from the PNG EGRA surveys [of Madang, NCD, and WHP] shows no evidence that students have an 'advantage' from learning to read in English vi-a-vis a language they are familiar

materials were almost completely absent from classrooms in the other two provinces.[14]) In spite of vernacular literacy being largely abandoned in 2009 according to official Madang provincial policy (Machuca-Sierra & de Silva, 2014a:17), *teachers in about 6 out of 65 classrooms observed have evidently continued teaching vernacular literacy in Madang, with huge benefits for the children concerned*. In those cases, the use of the vernacular materials (factors 3 and 4) appears to have a larger effect on English literacy than the use of certain kinds of English materials (factors 6 and 10)!

5.3 Negative (i.e. unhelpful) factors

Just five factors proved to have a negative association with vernacular literacy acquisition in ENB. These are shown in Figure 19.

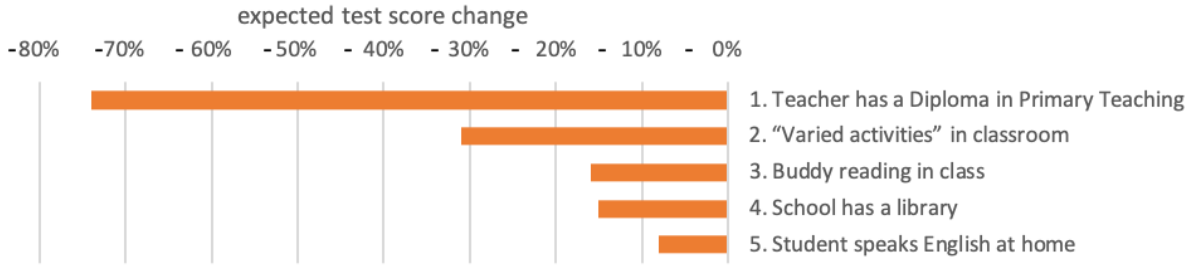

Figure 19. Detrimental factors for vernacular literacy in ENB, ranked by expected decrease below the reading comprehension test score mean

Note that the most important negative factor by far is the elementary teacher having a Diploma in Primary Teaching (rather than the appropriate Certificate in Elementary Teaching). An explanation for this came out in the teacher interviews, where teachers said that they had not learned synthetic phonics during their diploma studies (personal communication, Machuca-Sierra, 5 Sep 2020). This would have made it harder for them to teach vernacular literacy.

Factor 2 is probably too vaguely defined to be a useful measure for this study.[15]

with." On the contrary, it seems that there is indeed some evidence for that "advantage" in Madang (as well as the much stronger evidence in the ENB data).

[14] In WHP vernacular materials on charts were observed in just one classroom, and only English was observed on the blackboard (Machuca-Sierra & de Silva, 2014d: 42). In NCD no vernacular materials were observed at all (Machuca-Sierra & de Silva 2014a: 36).

[15] "Varied activites" is a general mix including spelling, verbal reporting, general discussions, reading with teachers, reading in groups, children reading on their own, building new words, creating new sentences & stories, etc (Machuca-Sierra & de Silva, 2014a: 68).

Factor 3 may be to do with the ineffectiveness of buddy reading (reading together in pairs) when both children in a pair may still be slow at reading. Group reading, on the other hand, does have a small positive effect (see Table B1).

Factor 4 seems counter-intuitive, but may indicate an inappropriate focus on English at this stage, as (in my observation) such libraries are typically stocked with English books.

Detrimental factors for bridging from vernacular to English literacy are shown in Figure 20. (It uses the same scale as Figure 19 so that the effect sizes involved can be easily compared between methodologies.)

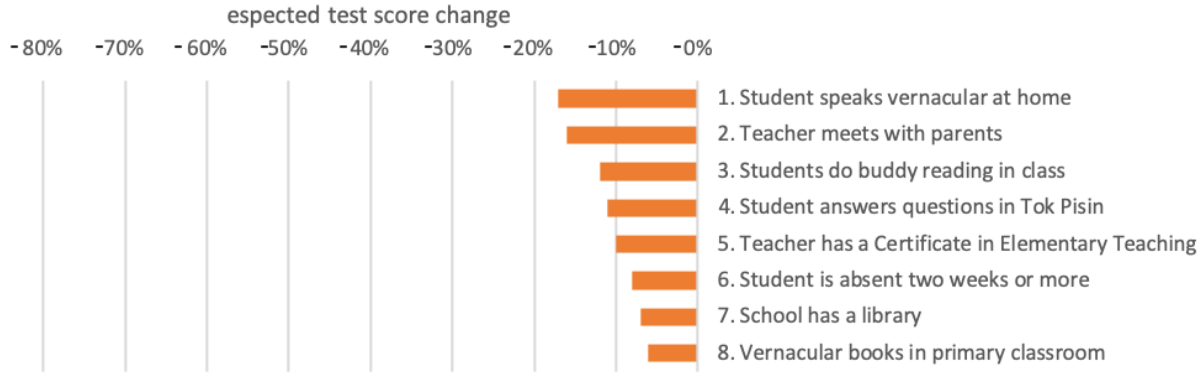

Figure 20. Detrimental factors for English literacy after vernacular literacy in ENB, ranked by the expected decrease below the reading comprehension test score mean

Factors 1, 4, 5, 6 and 8 in Figure 20 may make sense, but two seem counter-intuitive; the negative effect associated with teachers meeting with parents (factor 2) in this province might be because of problems, rather than the cause of problems; and buddy reading (factor 3) might be unsuccessful in ENB because children are still only beginning to learn to read English, a new language to them. The negative effect of having a library when children are learning to read English, however, seems counter-intuitive.

Even more counter-intuitive factors associated with poor literacy acquisition occur with the English-only methodology - see Figure 21.

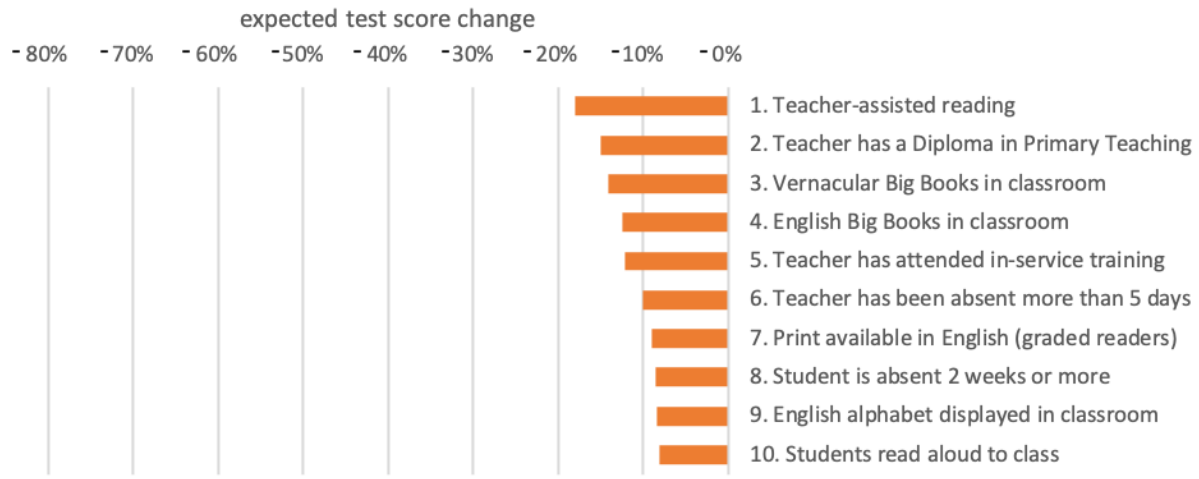

Figure 21.Detrimental factors for English literacy in an English-only system (as in Madang, NCD and WHP) ranked by expected decrease below the reading comprehension test score mean

The only factors that make sense in this set are factors 6 and 8, to do with teacher or student absence. The large numbers of counter-intuitive factors might actually be symptomatic of an inappropriate teaching methodology. Perhaps one way to explain how some of these could be negative would be to consider whether teachers might be pushing English literacy too hard onto children who are only starting to learn to speak English. The Read-PNG researchers themselves warn that "care should be taken in interpreting [such] results as they may have its [sic] basis in factors that could not be controlled for in the survey, namely socio-economic factors, rural/urban factors and so on" (Machuca-Sierra & de Silva, 2014d: 39).

6 DISCUSSION AND RECOMMENDATIONS

Although calls by education officials for the government to seriously scrutinise problems in the implementation of the education reforms in PNG went unheeded (Sinai, 2012), the Read-PNG school survey data has certainly shifted blame for the increasingly poor English reading standards away from the use of vernacular in schools.[16] Indeed, the real and detailed school literacy data explored in this paper provides very good evidence that the "vernacular-first" methodology, as practised in East New Britain [17] and a few schools of Madang at the time of the surveys, can be far more effective at helping children learn to read English, than a methodology that focuses on teaching English literacy from the beginning to children who

[16] Analysis of the complex causes of the lowering standards are beyond the scope of this study, but some important ones are discussed by Romanyshyn & Romanyshyn (2010), for example.

[17] Also acting in East New Britain's favour is its educational history; according to Litteral (2015:94), a successful mother tongue literacy program that had good community support was started in East New Britain as early as 1983. But East New Britain is not exceptional; other such programs were running before the reforms of the 1990s in the North Solomons Province (Bougainville), in Oro Province, and many language groups around the country.

are still learning to speak it, as practised in the National Capital District, Western Highlands Province and most of Madang.[18] There will, of course, be some unresearched socio-economic differences between these provinces that may affect the level of support and success of education efforts in all of these provinces, but the effect of language of initial literacy appears very clear.

In answering the question, "how best can children who speak a minority indigenous language at home become literate in a national or world language such as English?", the evidence strongly shows that a teaching methodology that includes the following points has been highly effective in practice in Papua New Guinea:

(1) teaching literacy in the children's vernacular first – this allows a much earlier start to literacy in school, and those literacy skills readily transfer to literacy in the national language;

(2) using vocabulary-building and comprehension exercises, reading aloud to the class, and having books for the children to read both at school and at home, in whichever the language the children are learning to read;

(3) giving the children's school good community support (e.g. a Parents and Citizens group);

(4) exploiting the children's vernacular literacy skills to bridge to national language literacy.

It also helps if a child has literate family members at home, and if teachers have had the appropriate training for the job.

7 CONCLUSIONS

The Read-PNG surveys using the EGRA tool have provided very good evidence that a child needs to learn to read in a language they already speak, before learning to read a national language that they have not yet mastered.

Minority languages have often been discounted, even in the minds of their own speakers, as hindrances to progress, but when they are valued as stepping stones to education for participation at a national or global level, this can give also give hope that these languages will be seen as treasures to be enjoyed and passed on to future generations.

Finally, it is more important to do research and examine the data thoroughly, than to rely on public opinion, especially when making decisions about education policy.

[18] The pedagogical problems of teaching English literacy to non-English speakers are discussed in footnote 3 in section 3.2 above.

ABBREVIATIONS

E1 = Elementary grade 1

E2 = Elementary grade 2

CLPM = Correct Letters Per Minute

COS = Classroom Observation Snapshot

CSPM = Correct Sounds Per Minute

CWPM = Correct Words Per Minute

EGRA = Early Grade Reading Assessment,

ENB = East New Britain,

MAD = Madang,

ME = Margin of Error

NCD = National Capital District,

NDOE = National Department of Education,

P3 = Primary grade 3,

P4 = Primary grade 4,

PNG = Papua New Guinea,

SD = Standard Deviation

WHP = Western Highlands Province,

REFERENCES

Birney, Grace H. (2016). *Applying EGRA in PNG: lessons learned* [PowerPoint slides]. Curriculum Development Division, National Department of Education of PNG. https://pubdocs.worldbank.org/en/625491457978448189/Session-10-Grace-Hiris-Birney.pdf **(**Accessed 2023-09-14.)

Department of Education. (2004). *Achieving a better future: A national plan for education 2005 – 2014*. Papua New Guinea: Department of Education. https://www.education.gov.pg/quicklinks/documents/edu-plans/national-education-plan-2005-2014.pdf (Accessed 2023-10-13.)

Dept told to act on directives. (2013, June 24). The National. http://edunewspng.com/ (Accessed 2014-12-9 – not available online in 2023.)

Dubeck, Margaret M. and Amber Gove. (2015). *The early grade reading assessment (EGRA): its theoretical foundation, purpose, and limitations*. International Journal of Educational Development 40 (pp. 315-322).

Litteral, Robert. (2015). Changes in mother tongue education policy in Papua New Guinea. *Language and Linguistics in Melanesia (LLM) 33(2)*, 93-99. https://www.langlxmelanesia.com/issues (Accessed 2023-12-03.)

Machuca-Sierra, Myrna and Steph de Silva. (2014a). *Madang Early Grade Reading Assessment (EGRA) survey: 2011 diagnostic results report.* Washington, DC: The World Bank. https://earlygradereadingbarometer.org/results/PNG/Madang-2011 (Accessed 2023-07-07.)

Machuca-Sierra, Myrna and Steph de Silva. (2014b). *National Capital District (NCD) Early Grade Reading Assessment (EGRA) survey: 2012 diagnostic results report.* Washington, DC: The World Bank. https://earlygradereadingbarometer.org/results/PNG/National-Capital-District-2012 (Accessed 2023-07-07.)

Machuca-Sierra, Myrna and Steph de Silva. (2014c). *East New Britain (ENB) Early Grade Reading Assessment (EGRA) survey: 2012 diagnostic results report.* Washington, DC: The World Bank. https://earlygradereadingbarometer.org/results/PNG/East-New-Britain-2012 (Accessed 2023-07-07.)

Machuca-Sierra, Myrna and Steph de Silva. (2014d). *Western Highlands Province Early Grade Reading Assessment (EGRA) survey: 2013 diagnostic results report.* Washington, DC: The World Bank. https://earlygradereadingbarometer.org/results/PNG/Western-Highlands-2013 (Accessed 2023-07-07.)

RTI International. (2015). *Early Grade Reading Assessment (EGRA) toolkit*. Second Edition. Washington, DC: United States Agency for International Development.

Romanyshyn, Bev and Vic Romanyshyn. 2010. Obstacles to student success in Papua New Guinea. *Rausim OBE long PNG*. https://rausimobe.files.wordpress.com/2010/11/obstacles-to-student-success-in-papua-new-guinea1.pdf (Accessed 2023-12-07.)

Sinai, Bernard. (2012, October 15). Education official thinks removal of OBE too soon. *The National*. http://edunewspng.com/2012/10/education-official-thinks-removal-of-obe-too-soon/ (Accessed 2015-01-15 – not available online in 2023.)

Taita, Luke. (2013). *Language policy in all schools*. Secretary's circular no. 04/2013. Waigani, Port Moresby: Department of Education.

The World Bank. (2016). *The implementation and results report (TF-98729) on a grant in the amount of US$19.2 million to the Independent State of Papua New Guinea for a Reading Education (Read-PNG) Project*. Report no: ICR00003637. Washington DC: The World Bank. http://documents1.worldbank.org/curated/en/554711467263195339/pdf/ICR3637-P105897-Box396255B-PUBLIC-disclosed-6-28-16.pdf (Accessed 2020-04-03.)

Wroge, Dianne. (2002). *Papua New Guinea's vernacular language preschool programme*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000137383 (Accessed 2023-12-07.)

APPENDIX A. LITERACY SKILL TEST DATA

The data from the reports (Machuca-Sierra & de Silva, 2014a–d) is collated and published in these appendices under implicit licence from the copyright holders, The World Bank, which states that their work "may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given."

The sources of the data in Table A1 are: for ENB, on pages 24–34; for Madang, on pages 69–73; for NCD, on pages 46–48; and for WHP, on pages 25–32 of the respective reports. The grade P4 standard deviation (SD) and margin of error (ME) are also shown so that 95% confidence intervals (as (mean-ME, mean+ME)) can be calculated. An overall average of all the test means is shown to the right, where each of the non-percentage scores has first been converted to a percentage of the maximum mean score for that test.

Table A1. Nine literacy test score means, by province and grade

Grade	Listening comprehension	Initial word sounds	English letter names CLPM	Letter sounds CSPM	Familiar words CWPM	Invented words CWPM	Dictation	Reading fluency CWPM	Reading comprehension	Overall average
ENB – East New Britain Province (P4, n=421)										
E1	71%	69%		36	14	12	53%	15	19%	45%
E2	76%	76%		43	21	17	68%	26	34%	58%
P3	38%	59%	74	39	41	27	68%	54	47%	66%
P4	51%	66%	88	45	61	37	80%	84	68%	85%
P4 SD	30%	37%	29	22.5	31.1	18.6	20%	41.1	31%	
P4 ME	±2.9%	±3.5%	±2.8	±2.1	±3.0	±1.8	±1.9%	±3.9	±3.0%	
MAD – Madang Province (P4, n=433)										
E2	39%	34%	57	19	14	10	56%	18	22%	37%
P3	33%	45%	69	26	29	17	73%	40	36%	51%
P4	43%	51%	80	32	43	23	81%	63	51%	66%
P4 SD	24.6%	34.6%	29.1	21.94	21.13	13.95	19.5%	31.2	27.8%	
P4 ME	±2.8%	±3.3%	±2.7	±2.1	±2.0	±1.3	±1.8%	±2.9	±2.6%	
NCD – National Capital District (P4, n=435)										
E2	50%	46%	55	21	10	6	30%	13	22%	34%
P3	56%	56%	73	29	24	13	51%	30	37%	51%
P4	60%	58%	87	32	41	22	65%	53	51%	66%
P4 SD	30%	36%	29	24	33	22	31%	46	34%	
P4 ME	±2.8%	±3.4%	±2.7	±2.3	±3.1	±2.1	±2.9%	±4.3	±3.2%	
WHP – Western Highlands Province (P4, n=438)										
E2	33%	25%	50	14	8	7	33%	9	11%	26%
P3	25%	42%	68	22	16	12	56%	21	21%	39%
P4	35%	47%	85	28	32	20	74%	44	44%	57%
P4 SD	26%	37%	24	19	16	11	17%	23	29%	
P4 ME	±2.4%	±3.5%	±2.2	±1.8	±1.5	±1.0	±1.6%	±2.2	±2.7%	

Maxima										
of means	60%	66%	88	45	61	37	81%	84	68%	
of all data	100%	100%	194	149	150	108	100%	221	100%	

Table A2. Numbers and proportions of zero scorers in nine literacy tests, by province and grade

Grade	Number tested	Listening comprehension	Initial word sounds	English letter names	Letter sounds	Familiar words	Invented words	Dictation	Reading fluency	Reading comprehension
ENB – East New Britain Province										
E1	418	3 0.7%	65 16%	–*	60 14%	121 29%	146 35%	73 18%	176 42%	235 56%
E2	427	6 1.4%	46 11%	–	40 9%	74 17%	82 19%	47 11%	95 22%	148 35%
P3	430	85 20%	109 25%	–	47 11%	30 7%	61 14%	31 7%	50 12%	89 21%
P4	421	45 11%	83 20%	–	21 5%	12 3%	31 7%	7 2%	17 4%	29 7%
MAD – Madang Province										
E2	400	42 11%	163 41%	6 1.5%	59 15%	46 12%	103 26%	53 13%	59 15%	147 37%
P3	446	82 18%	134 30%	10 2%	26 6%	23 6%	52 12%	21 5%	27 6%	91 20%
P4	433	36 8%	97 22%	7 1.6%	15 4%	12 3%	39 7%	11 3%	15 3%	35 8%
NCD – National Capital District										
E2	409	59 14%	122 30%	5 1.2%	64 16%	134 33%	261 64%	152 37%	154 38%	183 45%
P3	422	36 9%	75 18%	4 0.9%	35 8%	45 11%	177 42%	65 15%	61 14%	109 26%
P4	435	27 6%	79 18%	2 0.5%	32 7%	29 7%	115 26%	42 10%	40 9%	71 16%
WHP – Western Highlands Province										
E2	434	74 17%	256 59%	4 0.9%	152 35%	100 23%	174 40%	104 24%	143 33%	265 61%
P3	451	135 30%	162 36%	0 0%	99 22.0%	36 8.0%	86 19%	18 4%	50 11%	158 35%
P4	438	83 19%	131 30%	0 0%	48 11%	4 0.9%	31 7%	4 0.9%	9 2%	53 12%
Maxima and minima										
Max		30%	59%	2%	35%	33%	64%	37%	38%*	61%
Min		0.7%	11%	0%	4%	0.9%	7%	0.9%	2%	7%

* Not tested in E1 and E2 in ENB; numbers in P3 and P4 were zero or near zero.
NB: The numbers for WHP were calculated from percentages given in the report. The numbers for the other provinces were calculated from the numbers of *non*-zero scorers given in the respective reports.

APPENDIX B. STUDENT, TEACHER AND CLASSROOM FACTOR DATA

Tables B1–B3 list the factors investigated by Read-PNG that had statistically significant effects ($p \leq 0.10$) on Reading Comprehension test scores, along with the mean effect sizes (as regression coefficients, or, in a few cases, differences between means). Table B4 lists a few more factors that had statistically significant effects on other skills test scores. Factors are ranked according to effect size from most positive to most negative.

Table B1. Factors affecting vernacular reading comprehension in Grades 1 and 2 in ENB, ranked by effect size (source: Machuca-Sierra & de Silva, 2014c: 45, 49)

Positive factors	**Effect size**
Vernacular reading books in the classroom	53%
Student speaks vernacular at home	45%
School has a Parents & Citizens group	42%
Teacher does comprehension activities	30%
Student has stories at home to read	28%
Students learn meanings of new words	20%
Teacher reads aloud to students	19%
Classroom has a reading corner	16%
Student is in Grade 2 (vs. lower grade)	16%
Student speaks Tok Pisin at home	14%
Instruction is in vernacular	14%
Teacher meets with parents	11%
Student is female (vs. male)	9%
Group reading in class	9%
Negative factors	**Effect size**
Student speaks English at home	-8%
School has a library	-15%
Buddy reading in class	-16%
Varied activities in classroom	-31%
Teacher has a Diploma in Primary Teaching	-74%

Table B2. Factors affecting English reading comprehension (after learning vernacular literacy) in Grades 3 and 4 in ENB, ranked by effect size (source Machuca-Sierra & de Silva, 2014c: 46, 52, 59)

Positive factors	**Effect size**
English reading books in classroom	24%
Vernacular reading books in Grade 3	16%
Student is in a multigrade classroom in Grade 3	13%
Student is in Grade 4 (vs. Grade 3)	10%
Student is female	6%
Zero factor	**Effect size**
Number enrolled in class	0%
Negative factors	**Effect size**
Vernacular books in primary classroom	-6%
School has a library	-7%
Student is absent two weeks or more	-8%
Teacher has a Certificate in Elementary Teaching	-10%
Student answers questions in Tok Pisin	-11%
Buddy reading in class	-12%
Teacher meets with parents	-16%
Student speaks vernacular at home	-17%

The sources for the data in Table B3 are: for MAD, pp. 76, 88–96; for NCD, pp. 52 and 56, and for WHP, pp. 39, 40, and 53-56 of the respective reports (Machuca-Sierra & de Silva, 2014a–d). Most effect sizes included are statistically significant ($p \leq 0.10$); those that are not statistically significant are marked "NS" and the effect size is shown in parentheses. A few factors that are positive in one province and negative in another are in italics. An empty cell indicates that no data was collected for that factor in that province. The final column shows a simple average (or mean) for just the statistically significant factors of all three provinces. The factors in the table are ranked according to this average.

Table B3. Factors affecting English reading comprehension (with predominantly English-only education) in Grades 2–4 in MAD, NCD and WHP provinces, ranked by overall average effect size

Positive factors	**MAD**	**NCD**	**WHP**	**Average**
Student is in Grade 4 (vs. Grade 2)	42.7%	61.0%	57.0%	53.6%
Student is in Grade 3 (vs. Grade 2)	34.1%	45.0%	27.0%	35.4%
Vernacular songs on blackboard	34.4%			34.4%
Vernacular stories on blackboard/charts	23.7%			23.7%
Student has own activity book/text book	10.6%	28.0%	NS (4%)	19.3%
English songs on blackboard	19.2%			19.2%
Teacher instructs in English in Grade 2		18.0%		18.0%
Student has literate family member (e.g. sibling)	13.5%	20.0%	19.0%	17.5%
Student is female (vs. male)	NS (-1%)	NS (0%)	16.0%	16.0%
Class has printed English stories available		22.0%	7.0%	14.5%
Class does reading comprehension activities	13.9%			13.9%
Activity books are available	12.3%			12.3%
Students' written work displayed in classroom	12.1%			12.1%
Student's age	*-2.0%*	*24.0%*		11.0%
Teacher has a language guide	10.5%	NS (-1%)	NS (0%)	10.5%
Student is in multigrade classroom	8.0%		12.0%	10.0%
English dictionaries available	9.4%			9.4%
Sufficient space for organised group activities	9.2%			9.2%
Teacher meets with parents	5.1%	13.0%	NS (1%)	9.1%
Student does homework	NS (4%)	NS (8%)	9.0%	9.0%
English reading books/Shell books available	9.4%		7.0%	8.2%
Teacher reads to class	NS (9%)	8.0%%	8.0%	8.0%
Phonics activity/learning letter sounds			7.0%	7.0%
Students retell stories	5.9%			5.9%
Classroom has a reading corner	5.6%	NS (-3%)	NS (1%)	5.6%

Student sounds out words	5.3%			5.3%
Student recognises print as English	5.0%			5.0%
Student speaks English at home	3.0%	NS (3%)	5.0%	4.0%
Teacher had in-service training on phonics		*11.0%*	*-7.0%*	2.0%
Student repeated a grade			1.0%	1.0%
Zero factor	**MAD**	**NCD**	**WHP**	**Average**
Number enrolled in class		0.0%	0.0%	0.0%
Negative factors	**MAD**	**NCD**	**WHP**	**Average**
Students learn meanings of new words	*-7.9%*		*4.0%*	-2.0%
Student's family reads with him/her	NS (-1%)	NS (1%)	-5.0%	-5.0%
Teacher makes mistakes in English		NS (4%)	-5.0%	-5.0%
Student speaks vernacular at home	-5.0%	-6.0%		-5.5%
There is an active Parents and Citizens group	-7.2%	NS (-5%)		-7.2%
Student answers questions in Tok Pisin		-8.0%		-8.0%
Students read aloud to class	-8.1%			-8.1%
English alphabet is displayed in classroom	-8.3%			-8.3%
Student is absent two weeks or more	-7.0%	-10.0%	NS (-1%)	-8.5%
English reading books (graded readers) available			-9.0%	-9.0%
Teacher has been absent more than five days		-10.0%	NS (4%)	-10.0%
Teacher has attended in-service	-12.1%	NS (1%)		-12.1%
English Big Books in classroom	-10.9%		-14.0%	-12.5%
Vernacular Big Books in classroom	-14.1%			-14.1%
Teacher has a Diploma in Primary Teaching		-15.0%	NS (-8%)	-15.0%
Teacher-assisted reading	NS (-3%)	-18.0%		-18.0%

Table B4. Other factors that had statistically significant effects for other literacy tests

Positive Factors	**Context**
Students copy sentences	MAD: Dictation (+16%)
Silent reading (in primary school)	ENB: Oral Reading Fluency (+13 CWPM, or +15% of maximum mean)
Negative Factors	**Context**
Students repeat sentences in class	MAD: Dictation (-5%)
Teacher has a copy of the curriculum	MAD: Oral Reading Fluency (- 7 CWPM, or -8% of maximum mean)
Student ate breakfast	NCD: Oral Reading Fluency (-13 CWPM, or -15% of maximum mean)

Factors that had no statistically significant effect included: teacher or student using vernacular in the classroom, teacher's age and experience, print materials in Tok Pisin in classroom.