# FINAL REPORT

# THE EVALUATION OF THE LEARNING FOR LIVING PROJECT

# A PROJECT OF THE BUSINESS TRUST AND THE READ ORGANIZATION

## 2000 - 2004

## 24 February 2005

## ***ACKNOWLEDGEMENTS***

***This evaluation has been conducted in association with McCay Projects***

***Thanks to Professor W. Elley for the design of the Grade 3 and 4 test instruments***

***Thanks to Buyile and Lindiwe Mkhulise, Luisa Roscani, Kim Gregory, Jabulile and Sibongile Maratana for their contribution to the field work for this evaluation***

## EXECUTIVE SUMMARY

The READ/Business Trust Learning for Living Project (LFL) was launched in April 2000 and, from a high of 967, operated in 898 primary schools in all nine provinces of South Africa - with the support and cooperation of the Provincial and National Departments of Education (DoE).

The fundamental motivation of the project was the need to improve the quality of the outcomes of South African schooling. While much has been achieved in terms of access to schooling since the desegregation of the education system, much more remains to be done in terms of improving the quality of its outcomes.

The project was aimed at the areas of language, literacy and communication with the assumption that significant impact upon these areas would be achieved through the delivery of a comprehensive, cumulative and sustained programme with three major components:

- INSET for both classroom teachers and for senior school managers
- the provision of an extensive range of classroom-level material, including books and other types of teaching and learning material
- the provision of classroom-level monitoring and feedback.

The design for the evaluation of the Learning for Living Project used a longitudinal cohort-based approach combining both quantitative and qualitative research, organized through a quasi-experimental model, in both project and control schools.

Two impact indicators, in the form of predicted outcome statements, were selected for the measurement of the summative quantitative impact of the project:

- Project cohorts would show a significantly higher degree of gain between pre and post testing over the control cohorts on a battery of literacy and numeracy instruments. Scores are also obtained each year at Grades 5 and 7 levels.
- Project schools would show a significantly greater decline in their repeat rates, measured as the proportion of repeat pupils within total enrolment, over control schools.

The baseline report was presented in December 2000 and the data it contained clearly demonstrated that the project and control samples upon which research has been based were sufficiently equivalent for the requirements of the design. The second report was presented in December 2001 and began the process of systematically comparing the project and control groups for evidence of both quantitative and qualitative impact, the third and fourth reports continued that process and the final report, researched in late 2004, completes it.

Research for the final report was carried out in 80 schools, 50 project and 30 control, in all nine provinces in 22 different areas around the country. A total of 80 principal interviews, 152 teacher interviews and 152 lesson observations were completed, and 2 678 pupils tested, to yield the data upon which this report is based.

## FINDINGS

The design of the evaluation has been based on four generic questions and the key findings are organized against them.

**QUESTION ONE:** ***Has the project been delivered as designed?***

An annual comparison of project reports with the operational delivery schedule of the project indicated that the different components of the programme have largely been delivered as planned; the most notable exception is a continued shortfall in the level of classroom monitoring noted in previous reports. Though, in most cases, explanations exist for this shortfall it remains of some concern, especially insofar as the monitoring relates to the achievement of quality of usage of project methods and materials at classroom level.

Nonetheless, a very high level of training, materials supply and monitoring has been sustained. In all, **13 164** individual teachers have been trained in a total of **957** schools, a mean of **13.75** per school. Each teacher has received a mean of **9.6** INSET courses and **6.85** classroom-level monitoring visits.

A great deal of teaching and learning materials of different kinds has been supplied to each teacher:

- a mean of **6.8** kits, each typically consisting of a teacher copy and six pupil copies of a number of different titles
- a mean of **2.3** kits of reusable pupil activity books, consisting of packs of 25 books each
- a mean of **4.05** storage boxes and easels.

This level of training, materials supply and classroom-level support far exceeded anything available to the teachers in the control schools. The logistics of delivering, administering and managing a project of this kind on such a large scale probably exceeds anything previously attempted by a single NGO in this country. The overall total of all types of sets of material delivered to schools, for example, is **172 972**, and very much more if we consider individual packs making up larger sets; LFL reports reflect delivery of an astonishing total of **4 002 103** individual books of different types – a mean of **304** per teacher. Control over such an extended exercise in materials processing and handling, especially one in which delivery of materials must be coordinated with provision of INSET, is no simple matter and the READ department responsible for its successful achievement deserves recognition.

When we consider INSET and monitoring as well as materials supply, the fact that the LFL has been able to coherently sustain such a high level of project activity is a noteworthy achievement in its own right.

In our view, the delivery system operated effectively enough to ensure that the project was delivered in full and that, consequently, we are evaluating the application of the strategic design of the project.

**QUESTION TWO:** ***Has the project achieved its predicted qualitative impact (the process means of the project) on a consistent basis in the project schools?***

All of the reports of this evaluation have noted that teachers routinely use the materials on a regular basis, most of them along with the project-supplied teacher guides, and the consequence is that a LFL-based lesson, in its various forms, is recognizably such in all of the project schools. There was

clear evidence that the INSET and monitoring provided by the project succeeded in achieving a consistent impact on the classroom behaviour of teachers across the sample. This was the central classroom-level finding of the third and fourth reports, and it remains so into the final report.

Principals and teachers are virtually unanimous in their approval of the LFL and have consistently reported that positive changes have occurred in classrooms throughout the study. Lesson observations have confirmed that project impacts have occurred against a number of specific indicators:

- An increased number of observations of displays of pupil work.
- An increased number of observations of the use of teacher-made materials in project classrooms.
- An increased proportion of overall lesson time spent reading by pupils.
- An increased proportion of overall lesson time devoted to reading to pupils.
- An improved level of quality in the use of group work.
- An increase in the use of comprehension-type questions.
- The LFL has proved remarkably successful in selecting materials that are universally welcomed and used in all of the different local and regional contexts in which research takes place, even if to varying degrees of both teacher and pupil comprehension.
- The project methods and, especially, materials are generally supporting the presentation of higher quality lessons.

It is again noteworthy that the level of attention dedicated to reading continued to increase throughout the study in the project schools. This is clearly a project impact, and a very significant one in explaining the degree of pupil impact reflected by the study. It is obviously made more possible in the first place by the much greater supply of books available to teachers and pupils in project schools, though there is still little evidence of the extension of the practice of reading in a more qualitative sense, especially for sustained individual reading.

It should also be said that, while children in project schools have improved relative to those in control schools, the quantity and level of writing remains generally very low indeed in both types of schools. Few children can correctly complete simple sentence stems, even less have control of anything but the simplest and most limited vocabulary, or can spell words with any degree of accuracy.

**QUESTION THREE:** ***Have the predicted quantitative impacts (the product ends sought by the project) occurred in the project schools?***

Cohort One enjoyed the longest research life and the longest exposure (5 years) to the effects of the LFL. It, consequently, is the most reliable of the measures contained in the report and the one that should carry the most weight in the measurement of the pupil impact of the project.

The discussions that follow report the raw data in terms of covariance - the difference in gain in score between project and control groups between pre and post-testing. The gain in score of the control group over this period is subtracted from the gain of the project group to yield the measure of project impact – the resulting figure is referred to as the covariant. The covariant is reported for the project group in relation to the control group; a *plus* sign indicates the degree to which the gain of the project group *exceeded* that of the control group, a minus sign the opposite.

The second report reflected a mean covariant score +1.7% after one year in favour of the project group for Cohort One. Since the error factor of that data was 1% (a function of the relative sizes of

the sample and target populations), this figure indicated empirical evidence of impact at a significance level of 70%. The rigour of the design, and of the application of the data collection process, supported the conclusion that we were seeing the first systematic and empirical evidence that the project was achieving its core impact objective, the acceleration of the rate at which pupils were learning. The report went on to comment:
*'If this conclusion is correct, we can predict that Cohort I will demonstrate an increased rate of learning in successive tests, and that Cohort II will demonstrate the same pattern of increasing covariant gain of project over control as Cohort I. Further, since Cohort I will be exposed to the programme for the longest period, we can predict that it will demonstrate the largest absolute gain of the cohorts.'*

The third and fourth reports subsequently confirmed that the project group of Cohort One increased its mean literacy covariant score; from +1.7% to +3.7% to +6.4%. It will be noted that it achieved an increase in the *rate* at which the difference between project and control groups was widening over the first four years. In other words, it was not only getting better, but was doing so at an accelerating pace as its participation in the programme unfolded. Secondly, as predicted, Cohort Two provided evidence of a significantly increased rate of gain in the project group to **+3.45%** by its exit point at the end of 2003.

In the final year of the study, the mean literacy covariant of Cohort One grew by just +0.45% to an overall project total of **+6.85%**. The acceleration noted over the previous four years disappeared as the LFL came to an end - the project group has only very slightly increased its gain over control since 2003. Nonetheless, the eventual figure is still significant in terms of project impact - there is a 100% statistical certainty that the difference has been caused by the LFL.

It should be noted that the mean literacy covariant obscures the difference in outcome when writing and reading are compared. The impact for reading at **+8.4%** is +3.1% higher than the figure of **+5.3%** recorded for writing. This is consistent with the scores contained in previous reports and with the data derived from lesson observations throughout the study. Interviews with teachers and principals, as well as lesson observations, support the conclusion that there has been a real improvement in the reading abilities of the children in project schools. This is partly due to the greatly increased supply of books and other materials enjoyed by the project schools and partly to the INSET and monitoring with which they have been provided. On the other hand, there is little doubt that writing remains a very significant problem in all of the schools, both project and control. In neither type of school are children generally required to write more than a couple of sentences and the vast bulk of all the writing they do is of the 'fill in the missing word' variety. The fact that the LFL *has* achieved significant impact in improving the writing scores of project pupils is genuinely meaningful as a project outcome but it has to be acknowledged that the problem is not yet solved.

In terms of the *oral* use of English, there is no doubt that the first goal of all schools, parents, teachers and pupils is a sufficient level of verbal mastery of English. So long as the majority of School Governing Bodies and the DOE continue to opt for this language as the medium of instruction at Intermediate Phase and beyond, education cannot continue in any meaningful form without this ability, even if 'code switching' is employed.

An improved level of ability to understand and communicate verbally in English has been cited as a major project impact throughout the evaluation by both principals and teachers. This ability has also been evident in lesson observations in project classrooms where the difference with control schools can be startling. Even very young Foundation Phase pupils in some rural schools have been able to

sustain simple conversations with researchers in English, and the pupil level of comprehension of the lessons themselves was noticeably higher in many project schools, urban or rural. That this was not achieved to the same degree on the conventional indicators of post-basic literacy, reading and writing, should not detract from this outcome even though it could not be measured by the instruments used in the study.

Finally, in terms of project impact, it should be noted that, in 2001 after the first year of the LFL, the evaluators were asked to provide a comment on factors that could affect the eventual outcome of Business Trust projects. One of our comments then read:
*'Previous READ projects have generally been delivered at classroom level within the context of a formal language programme, for example the MAPEP or Day-by-Day series. With the decreased emphasis upon standardized learning programmes in general, and upon text books in particular, this context is altering toward a situation in which READ materials are sometimes used in their own right as the primary language programme in Foundation Phase.'*

The point, in our opinion, remains valid today. There are no longer any standardized language programmes employed by all schools but rather a wide variety of different programmes. While all may be approved by the relevant selection committees, there is an enormous range in methods, approaches, emphases and, most importantly, quality in these programmes. As a consequence, in many of the project schools the LFL was carrying the burden of providing both initial basic oral literacy at all grade levels *and* its evolution into the more developed range of competencies called literacy, of which verbal proficiency is a necessary, but not sufficient, precondition.

In our opinion, the LFL strategic approach would have had more impact on reading and writing if children were generally more proficient in oral English in the project schools at all grade levels in the first place.

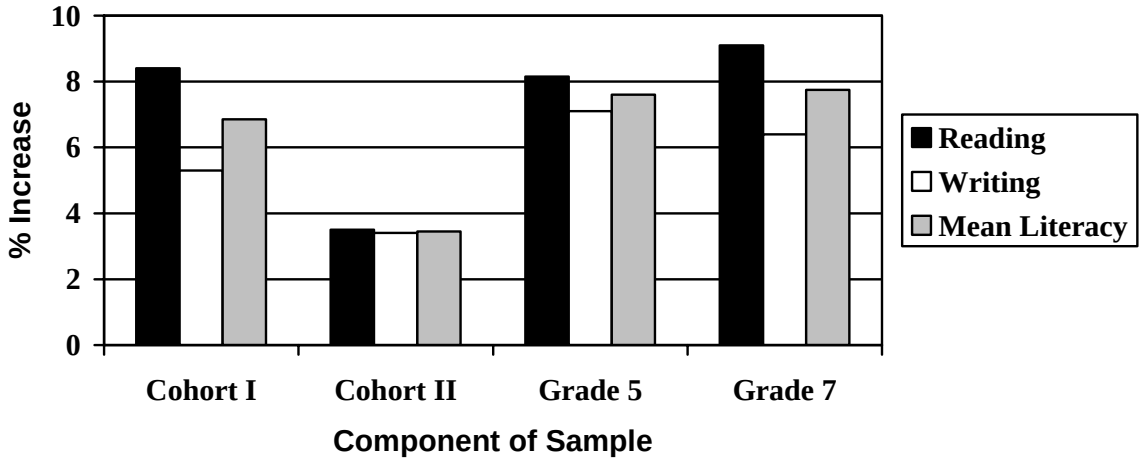
**Summary of Covariant Scores**

The table presents all of the scores obtained by both cohorts and by both grade levels.

| | **Writing** | **Reading** | **Mean Lit.** | **Numeracy** | **Mean L & N** |
|---|---|---|---|---|---|
| **Cohort I** | **+5.3** | **+8.4** | **+6.85** | **+1.0** | **+4.9** |
| Cohort II | +3.4 | +3.5 | +3.45 | +0.5 | +2.0 |
| Grade 5 | +7.1 | +8.15 | +7.6 | +2.85 | +6.0 |
| Grade 7 | +6.4 | +9.1 | +7.75 | +4.8 | +6.8 |

The consistency of the data is unmistakable; all of the twenty covariant scores are positive, even if two of them – numeracy for both cohorts – are too low to be statistically significant. The data clearly indicates that the LFL has achieved significant impact upon project pupils and that this impact has been caused by the project to a 100% degree of certainty. The data for Cohort I is the most reliable of the different comparisons and, in the opinion of the evaluators, it is the measure most likely to reflect the actual impact of the project. The rest of the comparisons confirm that positive impact has occurred; Cohort I provides the fine measurement.

The chart below illustrates the relative gains achieved by the components of the sample for language. Each column reflects the degree of improvement in the scores of the project group over those achieved by the controls between pre and post-testing.




The social significance of the absolute degree of change achieved by the Learning for Living project in relation to the needs of the national education system, ***in the absence of a comparative and longitudinal study combining data from a variety of language and book-based programmes***, is essentially a political and economic question. The question is obviously bound up with the issue of cost. While this evaluation was not asked to consider the issue, it appears that the impacts achieved by the LFL have been achieved at a low input cost. The project cost R153 million in total and READ estimates that 875 000 children attend the schools participating in the LFL. On these figures, the per capita cost is **R175** per pupil over the whole project, including project staff salaries and administration.

**Question Four:** ***Can process and product impacts measured or observed in the schools be ascribed to the project?***

All of the data and, especially, the use of the analysis of covariance within a quasi-experimental model for the collection and interpretation of both quantitative and qualitative information, continues to support the argument that the changes that have been measured or observed have occurred as a result of participation in the project.

In conclusion, it should be acknowledged that it is very unusual for an evaluation to have the opportunity to use five data collection rounds: one before the project commenced delivery, three during delivery and one after project completion. Three rounds over three years (pre, mid and post) is normally the best that can be hoped for. Indeed, in many cases researchers are asked to evaluate a project only after its completion - one round of data collection without baselines - in others they are asked to work without a control group against which to compare project impact.

The end result is that this study has been based upon the use of the strongest available design and both READ and the Business Trust deserve mention for their recognition of not only the need for an evaluation of the LFL, but also of the need to provide the research with the most reliable of foundations.

# SECTION ONE

# BACKGROUND TO THE FINAL REPORT

This is the last of five annual reports of the evaluation of the READ/Business Trust 'Learning for Living' (LFL) Project.

## 1. Objectives and Design of the Project

The LFL was operated with the support and cooperation of the Provincial and National Departments of Education (DoE) and was launched in April 2000 in 957 primary schools, in all nine provinces of South Africa, with 13 164 teachers and approximately one million pupils. Since then a number of schools were withdrawn from the project with the approval of the DoE due to the amalgamation of a number of schools in the Northern Cape, and a number of severely dysfunctional schools in other provinces. As a result there were 898 schools participating in the LFL by project end.

The fundamental motivation of the project is the need to improve the quality of the outcomes of South African schooling. While much has been achieved in terms of access to schooling since the desegregation of the education system, much more remains to be done in terms of improving the quality of its outcomes. The areas of literacy and language in general are universally acknowledged to be critical in achieving this goal, especially in a multilingual country in which the vast majority of School Governing Bodies and parents have opted for the use of a second language, English, as the medium of instruction in Intermediate Phase and beyond – sometimes even in Foundation Phase.

The objectives and operational design of the project were described in detail in the baseline report and, for the interest of readers who have not read that report, are reproduced in Appendix One.

In short, the project was aimed at the areas of language, literacy and communication with the assumption that significant pupil impact upon these areas would be achieved through the delivery of a comprehensive, cumulative and sustained programme with three major components:

- INSET for both classroom teachers and for senior school managers
- the provision of an extensive range of classroom-level material, including books and other types of teaching and learning material
- the provision of classroom-level monitoring and feedback.

The INSET courses provided guidance in the effective use of a book-based approach to learning and, specifically, in the use of the classroom teaching and learning materials provided by the project which, in turn, provided the opportunity for the guided practice of the concepts introduced during the courses. Classroom monitoring, based upon open checklists of expected and observable competencies supplied to teachers at INSET courses, was intended to provide feedback firstly to teachers in their application of the materials and concepts, and secondly to the organization itself as formative information about the developing impact of the project.

The programme was delivered through a national network of regional offices in the different provinces, each of which consisted of longer-term professional READ staff along with new appointments trained as they were recruited. The regional offices were supported by the national office which provided a centralized source of specialist, logistic, training, administrative and other support.

To provide reliable and useful formative information to the management of the delivery system of the project as a whole, a systematic internal quality control programme was developed yielding direct feedback about qualitative impact at classroom-level.

## 2. Design of the Evaluation

The research design was discussed in detail in the baseline report and is reproduced in Appendix Two. In short, the design for the external evaluation of the Learning for Living Project used a longitudinal cohort-based approach organized through a quasi-experimental model, combining quantitative, qualitative and contextual research in project and control schools.

The design was based upon four generic questions:

1 Has the project been delivered as designed?
2 Has the project achieved its predicted qualitative impact *(the process means of the project)* on a consistent basis in the project schools?
3 Have the predicted quantitative impacts (*the product ends sought by the project*) occurred in the project schools as a result?
4 Can process and product impacts measured or observed in the schools be ascribed to the project?

These questions are answered through:

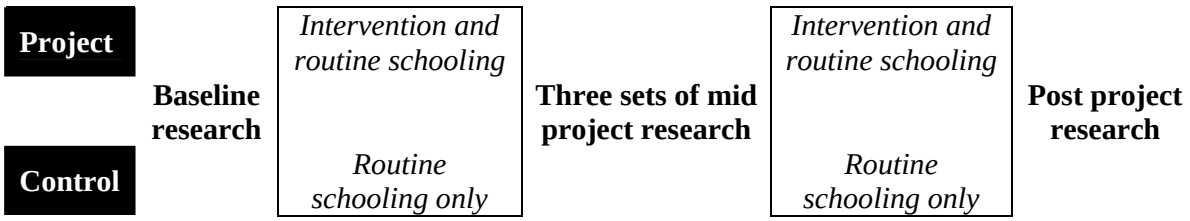
1 Comparison of formal project delivery reports against the delivery schedule, triangulated against principal and teacher reports of project contact.
2 A combination of principal and teacher interviews triangulated against lesson observations. Two researchers are used.
3 The application of a longitudinal quasi-experimental design using the analysis of covariance between mean pre and post-test scores on literacy and numeracy instruments for three cohorts of project and control pupils.
4 The combination of all of the above and, especially, the use of the analysis of covariance within a quasi-experimental model for the collection and interpretation of both quantitative and, where sensible, qualitative information.

Two impact indicators, in the form of predicted outcome statements, were selected for the measurement of the summative quantitative impact of the project:

- Project cohorts will show a significantly higher degree of gain between pre and post testing over the control cohorts on a battery of literacy and numeracy instruments. Scores are also obtained each year at Grades 5 & 7 level for the same analysis.
- Project schools will show a significantly greater decline in their repeat rates, measured as the proportion of repeat pupils within total enrolment, over control schools.

To investigate project impact upon pupils, a quasi-experimental design using tests in writing, reading and numeracy for project and control groups was applied. The testing process was designed to yield two types of data; the first longitudinally tracks the performance of the *same* groups of children (cohorts) as they progress through different grade levels, the second records the performance of *different* groups of children as they reach Grade 5 and Grade 7 levels.

The first type is much more reliable as a measure of the impact of an intervention than the second; tracking the same groups of children over time as they progress from one grade to the next controls for many of the variable and contextual factors that influence children's performance. Conversely, testing different groups of children at specific grade levels controls many fewer variables, and the results are more open to external and contextual influences like demographic changes, different teachers, or local-level DoE and other initiatives. Nevertheless, this type of testing yields data concerning the effect of intervention projects on the general 'standards' of the schools, a factor routinely used in discussions about educational transformation.




Baseline testing took place in early 2000, with three sets of annual mid-project testing, and post testing in late 2004. Cohort One had a research life of five years to reflect project impact, starting at Grade 3 in 2000 and reaching Grade 7 by 2004. Cohort Two had a life of four years, from Grade 4 in 2000 to Grade 7 in 2003; the final data set for this cohort was also contained in the fourth report. Children from Grades 5 and 7 were tested on each research round, and data from these sources both reflect impact over five years. The raw test data has been analyzed in terms of the covariance of the pre and post scores of the project and control groups.

It is interesting to note that the children in the project group who were tested in Grade 5 in 2004 were in Grade 1 in 2000. They, therefore, started their education in schools in which the LFL project was operating.

The design originally intended the use of school-based repeat rates as an indicator of project impact but that proved unworkable. Consequently, the existing study data has instead also been used to calculate the relative pass rates of the project and control groups.

### 3. Application of the Research Design

The baseline study was carried out in 90 schools in all, 54 project and 36 control, in 22 research sites in all nine provinces. The schools were selected as a stratified sample to reflect as much as practicable the national and provincial variety of socio-economic and educational contexts. Just over half of them are in rural, many remote rural, areas and the rest in or near regional or major urban areas in both formal and informal settlements. Controls were matched as closely as possible, primarily with regard to socio-economic status and involvement with other major development initiatives. Pupils were subsequently randomly selected at each of the schools.

As some of the more remote 'deep rural' areas may be unfamiliar to the general reader, the table below generally lists the research sites in terms of the nearest town; the Tzaneen schools in the Limpopo Province, for example, are actually in the Bolebedu District to the north of the town of Tzaneen itself. The list does, however, portray the variety of the 22 research sites.

**Table 1: Research sites**

| | | | |
|---|---|---|---|
| Limpopo | Tzaneen | KwaZulu/Natal | Bergville |
| | Louis Trichardt | | iXopo |
| Mpumalanga | Badplaas | Eastern Cape | Motherwell |
| | Malelane | | Queenstown |
| Gauteng | Soweto | | Peddie |
| | Vosloorus | Western Cape | Piketberg |
| North West | Potchefstroom | | Worcester |
| | Orkney | | Khayelitsha/Athlone |
| | Mabopane | Northern Cape | Kimberley |
| Free State | Bloemfontein | | Koopmansfontein |
| | Welkom | | Posmasberg |

The final report contains data from 80 schools, 50 project and 30 control, 88.9% of the original samples. The change in these numbers from baseline (90 schools) was due to a number of factors beyond the control of the field researchers:

- School closed during the course of the study.
- School Governing Body(SGB) meeting scheduled during research visit.
- Two schools on strike.
- Four schools in which the principal was absent and rest of staff not informed of visit.
- Two schools not informed of visit by LFL staff.

In all, 80 school manager interviews, 152 teacher interviews and 152 lesson observations took place for the final report.

### 3.1. Attrition of the Sample

Previous reports have commented upon the extraordinary attrition rates evidenced by the cohorts. The second report (2001) noted that they were already 14.5% smaller one year after baseline. In order to ensure that the eventual sample size at post-testing, four years after baseline, was large enough to support reliable extrapolation to the whole project population, the design originally assumed a total attrition rate of around 25%. Consequently samples of 60 pupils per school (20 per cohort) were selected at baseline to produce an eventual expected minimum of 45 per school, sufficient for the overall needs of the design. However, at an attrition rate of 14.5% in the first year alone, we were faced with the possibility of an eventual sample too small to support extrapolations.

While the testing of newly selected grade 5 and 7 children each year compensated to some extent, we were forced to defend the size of the cohorts as the primary impact measurement groups. In the event, we selected five extra children in 2001 when the original sample size for a cohort at any school had fallen below 15. As a consequence, we reduced the attrition rate for the first year to a more manageable, but still high, 11.2% - the exercise was not repeated as it defeats the purpose of cohort testing.

**Table 2: Attrition rate of Cohort I (2000 - 2004)**

| | **Project** | **Control** | **Sample** |
|---|---|---|---|
| Baseline size | 1 049 | 678 | 1 727 |
| Final Size | 666 | 388 | 1 054 |
| Total attrition | 383 | 290 | 673 |
| Total attrition (%) | 36.5 | 42.8 | 39.0 |
| Average attrition p.a. (%) | 9.1 | 10.7 | 9.7 |

**Table 3: Attrition rate of Cohort II (2000 - 2003)**

| | Project | Control | Total |
|---|---|---|---|
| Baseline n | 1 054 | 674 | 1 728 |
| Final n | 727 | 423 | 1 150 |
| Total attrition | 327 | 251 | 578 |
| Total attrition % | 31.0 | 37.2 | 33.4 |
| Average % attrition p.a. | 10.3 | 12.4 | 11.1 |

Since the overall annual enrolment figures remained reasonably stable and the daily attendance rates of the schools is over 90%, it is now evident that most of the children missing from the original cohorts left the schools, and that many new pupils arrived at all grade levels, few of whom had any previous exposure to LFL-trained teachers.

The second report commented that discussion between READ and the schools on this issue suggested that the problem was unlikely to significantly affect the project in the immediate future. Unfortunately, that suggestion has now clearly proved to be unfounded. We are left with the obvious conclusion that many of the schools in the country are effectively replacing their pupil enrolment over the period of primary schooling, a conclusion which may help to explain why both transformational and NGO-generated change is taking as long as it is to produce significantly better outcomes – pupil mobility is so high it tends to defeat the purposes of projects based upon sustained cumulative change.

### 3.1.1. Reliability of Post-Test Sample

The pupil impact data contained in this report has been derived from 2 678 children in all.

**Table 4: Summary of quantitative research for the final report (2004)**

| Grade | Project | Control | Total |
|---|---|---|---|
| **5** | 1 019 | 605 | 1 624 |
| **7** | 666 | 388 | 1 054 |
| **Total** | **1 685** | **993** | **2 678** |

According to the literature (Krejcie and Morgan, 1970), a sample size of 2 678 for a population of one million yields data at a confidence level of 95% and a precision just better than 2%. The post-test sample is, therefore, large enough for the purposes of covariant analysis.

Secondly, since the pre-test scores of the project and control groups within the sample indicated almost complete equivalence between the two groups - the mean difference for all three instruments was 0.5% - we can safely assume that the impact test data reported in this study is reliable.

## SECTION TWO

## DELIVERY OF THE PROJECT

### 1. Project Statistics

The second and third reports contained a detailed comparison of planned versus actual project delivery over the first two years. These reports concluded that, despite a shortfall in the number of monitoring visits per teacher, the Learning for Living Project had very largely succeeded in implementing its complete strategic design. This provided confirmation that it was likely the project would be delivered to a sufficiently complete degree to allow us to conclude that we would be evaluating the application of the strategic design of the project.

That situation has essentially been maintained into the fourth and final reports, and it appears redundant to provide as detailed an examination of the issue in this report. The table below, therefore, reports total project activity over each year of the full project life span rather than comparing actual and planned annual totals item by item. It will be evident that this analysis confirms the conclusions of the previous reports; that the first of the generic evaluation questions has been answered in the positive.

**Table 5: Summary of total project delivery**

| | **•Schools** | ***Teachers** | **#Delegates** | **♣Monitoring** | **Materials** | | |
|---|---|---|---|---|---|---|---|
| | | | | | ⊙*Teachers* | ✦*Pupils* | ◈*Equip* |
| Year One | 957 | 5 500 | 18 679 | 10 717 | 28 150 | 4 000 | 11 350 |
| Year Two | 907 | 5 500 | 26 659 | 16 104 | 49 388 | 19 000 | 35 548 |
| Year Three | 928 | 1 000 | 26 957 | 19 130 | 6 445 | 0 | 6 445 |
| Year Four | 898 | 1 164 | 30 712 | 23 541 | 5 222 | 7 424 | 0 |
| Year Five | 896 | 0 | 22 801 | 20 748 | 0 | 0 | 0 |
| **Total** | **957** | **13 164** | **125 808** | **90 240** | **89 205** | **30 424** | **53 343** |

•*The figure for each year refers to the total number of schools involved in the project. Over the life of the LFL a number have been withdrawn with the agreement of the DoE, chiefly due to amalgamation of schools in the N. Cape and some severely dysfunctional schools in other provinces.*

*****Teachers refer only to new teachers entering the project each year; e.g. in Year 3, 6 500 teachers were trained but only 1 000 were new to the project – subject teachers doing the non-fiction course.*

*#A delegate is defined as a teacher attending a course –for example, if one teacher attends three courses he/she will be reflected as three delegates in this column.*

♣*Monitoring visits refer to classroom observations by project trainers of lessons by teachers after training courses which result in the production and submission to the LFL of a formal feedback form.*

⊙*These materials largely consist of packs, themselves usually made up of different titles with multiple, usually six, copies of each title. If we had included all of these sets the eventual total would have been very much higher. These figures, therefore, refer to more global items like Big Books, Story Kits, Sunshine Pack, Theme Packs etc.*

✦*Pupil material consists of reusable pupil activity books in packs of 25 per class provided to teachers.*

◈*Equipment consisted of storage boxes, both wood and cardboard, as well as 5 500 easels.*

**Table 6: Project delivery as the mean of each component**

| **Teachers trained/school** | **Courses /teacher** | **Monitoring visits/teacher** | **Materials** | | |
|---|---|---|---|---|---|
| | | | *Teachers* | *Pupils* | *Equip.* |
| **13.75** | **9.6** | **6.85** | **6.8** | **2.3** | **4.05** |

The project managed to sustain a very high level of delivery. The overall total of all types of packs of material delivered to schools, for example, was **172 972**, and very much higher if we consider

individual packs making up larger sets. In all, LFL reports reflect procurement, processing, packaging, dispatch and delivery of an astonishing **4 002 103** individual items – a mean of **304** per teacher. Control over the planning, administration and logistics of such an extended exercise, especially one in which delivery of materials must be coordinated with provision of INSET, is no simple matter and the READ department responsible for its successful achievement deserves recognition. When we consider INSET and monitoring as well as materials supply, the fact that the LFL has been able to coherently sustain such a high level of project activity is a noteworthy achievement in its own right. The level of training, materials supply and classroom-level support continues to far exceed anything available to the teachers and children in the control schools.

However, since the project plan called for the provision of one classroom-based monitoring visit after each course for each teacher, the cumulative course versus monitoring visit shortfall has continued to widen from the –1.7 visits per teacher noted in the second report to **-2.75** for the current report. Previous reports noted that there were wide variations in this regard between the different provinces and that has remained true into the final year of research. There appear to be four main explanations for the overall shortfall:

- Teachers only receive monitoring visits after substantive programme courses but not after motivational and other general workshops. Consequently one would not in any case expect a perfect match between the mean numbers of courses attended and visits received per teacher.
- In some cases, researchers were at schools shortly after a course had been presented and there had been no time for the associated visits to take place. There is a 'lag' time, therefore, between courses reported and visits completed.
- In some cases, shortfalls have occurred through normal project staff turnover, while newly appointed staff are being trained.
- The effort to ensure that newly-appointed teachers in project schools received INSET and monitoring to support their ability to use project-supplied materials 'inherited' from departed teachers who had already been trained, inevitably placed stress upon the ability of LFL to keep up with its routine schedule of monitoring.

In general, despite the fact that it would indisputably have been better to achieve complete monitoring coverage according to project plans, these four factors go some way to explaining the existing shortfall. Given the already large figure of over 90 000 monitoring visits completed, it would clearly be unreasonable to assume that the project delivery plan had been seriously compromised by the shortfall to the extent that the evaluation is no longer measuring the impact of the strategic design of the LFL. It should also be remembered that a number of visits by project staff to schools did not involve formal classroom monitoring and feedback; information gathering and dissemination, problem-solving, checking materials supply and issuing, planning for motivational events (e.g. Festival of Books and Readathon), etc were among other reasons for school visits.

### 1.1. Per Capita Cost

The total cost of the LFL project was R153 million and the project estimates 875 000 children in the schools. This yields an overall project per capita cost of **R175** per pupil. This figure includes all LFL costs including staff salaries and project administration as well as INSET, materials and monitoring costs.

## 2. Participant Delivery and Utility Reports

The study was designed to triangulate project delivery reports with reports from school participants. This, however, proved to be increasingly difficult as the research continued – by the final report, many participants were unable to immediately identify INSET and materials with which they had been supplied, even when they were provided with copies of the original project delivery plan (*See* Appendix One, Extract from the Baseline Report, Table One – Summary of Delivery Schedule for Teacher INSET and Materials Supply.)

Nevertheless, since participants were provided with these alternatives, it seems reasonable to assume that the data will be relatively more accurate than that obtained by open-ended questions in the previous reports. Further, the research for the final report also included a question asking principals and teachers to nominate the INSET and materials they had personally found most useful at school/classroom level. Consequently, although the delivery figures cannot be directly compared with project-supplied figures, they do reflect the relative utility value placed on each item by project participants; they may be of interest to the project designers.

The questions were necessarily restricted to teachers in project schools. Interviewees were allowed to nominate a maximum of three categories in terms of usefulness to ensure a high degree of value-based choice in the responses, making it possible to construct a rank order of the utility of materials and INSET. Without this sort of device, teachers almost always provide a blanket endorsement of all of the project components, making it impossible to establish the actual relative value of each set.

### 2.1. Materials

**95.8%** of teachers confirmed that they had received materials from LFL, the remainder were all newly arrived teachers.

**Table 7: Materials provision: teachers (%)**

| Type | Utility Rank | Received |
|---|---|---|
| New heights | 1 | 31.2 |
| Story kit | 2 | 91.7 |
| Home language starter pack and story kit | 3 | 13.5 |
| Accelerated pack | 4 | 70.8 |
| Group Readers | 5 | 86.5 |
| Sunshine extension pack | 6 | 81.2 |
| Story Kit little books | 7 | 74.0 |
| Book education pack | 8 | 85.0 |
| Classroom resource collection | 9 | 80.2 |
| Learners books: accelerated pack | 10 | 66.7 |
| Core books | 11 | 18.7 |
| Learners books for Sunshine | 12 | 78.1 |
| Learning to read and write | 13 | 30.2 |
| Sunshine science books | 14 | 9.4 |
| *Other: Writing challenge, Fluency packs, Shared reading and writing, Charts, Posters* | *n/a* | *15.6* |

### 2.2. INSET

**97.9%** of teachers confirmed that they had attended INSET courses presented by the LFL project.

**Table 8: INSET: teachers (%)**

| Course | Utility Rank | Attended |
|---|---|---|
| Shared reading and writing | 1 | 78.1 |
| Group and guided reading | 2 | 79.2 |
| Reading aloud | 3 | 75.0 |
| Reading and writing for real purposes | 4 | 55.2 |
| Using stories for language development | 5 | 71.9 |
| Book education | 6 | 67.7 |
| Reading for information skills | 7 | 50.0 |
| Learning to read and write | 8 | 32.3 |
| Teaching phonics and spelling | 9 | 24.0 |
| Using a non-fiction book for lesson planning | 10 | 61.5 |
| Using a classroom resource collection | 11 | 66.7 |
| Classroom management | 12 | 64.6 |
| Choosing and using books | 13 | 37.5 |
| Independent reading | 14 | 65.6 |
| Introducing a literacy programme | 15 | 54.2 |
| Learning to read: home language literacy | 16 | 8.3 |
| *Other: Assessment, story telling, Skill development, Methodology, About using the books* | *n/a* | *8.3* |

### 2.3. Principal Training

**100%** of principals confirmed that they had attended LFL training courses and they were more able to recall specific INSET courses than were teachers. It is likely, therefore, that the delivery figures are more likely to be accurate, though our sense was that they are still incomplete.

**Table 9: INSET: principals (%)**

| Course | Utility Rank | Attended |
|---|---|---|
| Financial management and budgeting | 1 | 77.5 |
| Marketing and fundraising for the school | 2 | 79.6 |
| Strategies for building a language programme | 3 | 18.4 |
| School governance and management | 4 | 75.5 |
| School manager's leadership, mentoring and monitoring | 5 | 44.9 |
| Curriculum management and development | 6 | 69.4 |
| Management of human resources | 7 | 69.4 |
| School administration | 8 | 71.4 |
| Insights to success | 9 | 44.9 |
| Communication skills and conflict management | 10 | 81.6 |
| Parental education and involvement | 10 | 20.0 |
| Management of physical resources | 10 | 20.0 |
| Strategic planning for school leaders | 11 | 67.3 |
| Planning the last year of the LFL project | 12 | 10.2 |

### 2.4. Classroom Monitoring

**90%** of teachers agreed that they had received monitoring visits from project staff. The project design calls for a classroom visit after each course but, when asked what triggered or caused a classroom visit, only 21.9% of them could provide this answer; 25.0% could not hazard a guess and

the rest of the responses included phases like *'they come to help us', 'they are here so often', 'they come to check we use the books'*, etc without any indication of awareness of a systematic pattern to the visits.

Since the design stresses the value of classroom feedback based upon the use of course-specific checklists to support the development of competence in the skills dealt with by each course, these are curious comments. It was clear initially that many teachers confused general school visits or book/materials delivery with monitoring visits but, even when researchers indicated that they were referring to classroom visits which concluded with a discussion of a checklist, most still did not articulate the connection between attending a course and receiving a formal classroom monitoring visit.

It was even more difficult to establish how many such visits teachers had received over the life of the programme. We obtained a reported mean of **17.9**, a plainly impossible figure given the mean of **6.85** reported by the project itself. Although, theoretically at least, teachers should have each received a roughly similar number of formal visits, the answers we received to this question were wildly inconsistent, ranging from none to over 80!

### 2.5. Leader Teachers

According to the project design, the most enthusiastic and able teachers in schools were to be identified as leader teachers. The fundamental role of these teachers was to provide both programme and motivational support to the rest of the staff in between visits from READ staff. They are also, importantly, available to orient teachers who are newly appointed to the schools. The intention was to train them a little in advance of the rest of the teachers so that they were able to provide support immediately after general INSET courses were completed for other teachers.

Of the teachers in project schools who were interviewed, and whose lessons were observed, for this report, **21.0%** were leader-teachers. All teachers in project schools were asked what the primary roles of these leader teachers in relation to the project were.

| **Table 10: Primary roles of leader teachers in schools** | **Teachers (%)** |
|---|---|
| General motivation/encouragement of staff | 68.6 |
| Report back to staff on own training/project meetings | 53.2 |
| Maintain contact between school and LFL | 52.1 |
| Administration of LFL in school e.g. materials supplied | 46.8 |
| Provide classroom support/advice to other teachers as needed | 34.0 |
| Actively coordinate programme in school | 18.1 |
| Organize events like Readathon | 14.9 |
| Train teachers e.g. new arrivals | 12.8 |
| Supervise classroom appearance | 5.3 |
| Nothing | 5.3 |

## SECTION THREE

## PUPIL IMPACT

### 1. Raw Scores of the National Sample

The baseline report noted that the national sample forms the unit of analysis in this evaluation. It was selected to provide an analogue of the national population with which the LFL was dealing, rather than to provide a sample with provincially reliable sub-samples. This option would have resulted in a very much larger sample, along with significantly increased costs, and it was not necessary to adopt it in order to provide the evaluation with a nationally representative sample.

**Table 11: Project group post-test scores (%)**

| Grade | n | Writing | Reading | Mean Lit. | Numeracy | Mean L&N |
|---|---|---|---|---|---|---|
| **5** | 1 019 | 19.0 | 40.95 | 30.0 | 39.45 | 33.1 |
| **7** | 666 | 38.1 | 60.7 | 49.4 | 52.9 | 50.6 |
| **Total n** | 1 685 | | | | | |

**Table 12: Control group post-test scores (%)**

| Grade | n | Writing | Reading | Mean Lit. | Numeracy | Mean L&N |
|---|---|---|---|---|---|---|
| **5** | 605 | 12.0 | 30.6 | 21.3 | 37.1 | 26.6 |
| **7** | 388 | 33.4 | 53.3 | 43.35 | 52.0 | 46.2 |
| **Total n** | 993 | | | | | |

### 2. Analysis of Covariance

The tables that follow analyze the raw data in terms of covariance - the difference in gain in score between project and control groups between pre and post-testing. The gain in score of the control group is subtracted from the gain of the project group to yield the measure of project impact – the resulting figure is referred to as the covariant. The covariant is reported for the project group in relation to the control group; a *plus* sign indicates the degree to which the gain of the project group *exceeded* that of the control group, a minus sign the opposite. (Where possible, figures have been rounded to one decimal point and this sometimes creates small arithmetic discrepancies when combining mean scores.)

### 2.1. Cohort One

We have already noted that Cohort I has had the longest research life and the longest exposure to the effects of the LFL. It, consequently, is the most reliable of the measures reported below and it is the one that should carry the most weight in the measurement of the impact of the project.

**Table 13: Cohort I: covariance 2000 – 2004 (%)**

| | n | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|---|
| **Project** | | | | | | |
| Baseline | 1 049 | 3.2 | 17.5 | 10.35 | 23.6 | 14.8 |
| Final | 666 | 38.1 | 60.7 | 49.4 | 52.9 | 50.6 |
| Change | | +34.9 | +43.2 | +39.05 | +29.3 | 35.8 |
| **Control** | | | | | | |
| Baseline | 678 | 3.8 | 18.5 | 11.15 | 23.7 | 15.3 |
| Final | 388 | 33.4 | 53.3 | 43.35 | 52.0 | 46.2 |
| Change | | +29.6 | +34.8 | +32.2 | +28.3 | +30.9 |
| **Covariant** | | **+5.3** | **+8.4** | **+6.85** | **+1.0** | **+4.9** |

The second report reflected a mean covariant score +1.7% after one year in favour of the project group in this Cohort. Since the error factor of that data was 1% (a function of the relative sizes of the sample and target populations), this figure indicated empirical evidence of impact at a significance level of 70%. The rigour of the design, and of the application of the data collection process, supported the conclusion that we were seeing the first systematic and empirical evidence that the project was achieving its core objective, the acceleration of the rate at which pupils were learning. The report went on to comment:

*'If this conclusion is correct, we can predict that Cohort I will demonstrate an increased rate of learning in successive tests, and that Cohort II will demonstrate the same pattern of increasing covariant gain of project over control as Cohort I. Further, since Cohort I will be exposed to the programme for the longest period, we can predict that it will demonstrate the largest absolute gain of the cohorts.'*

The third and fourth reports subsequently confirmed that the project group of Cohort I increased its mean literacy covariant score; from +1.7% to +3.7% to +6.4%. It achieved an increase in the *rate* at which the difference between project and control was widening. In other words, it was not only getting better, but was doing so at an accelerating pace as its participation in the programme unfolded. Secondly, as predicted, Cohort II provided evidence of a significantly increased rate of gain in the project group to +3.45% by 2003.

The fourth report concluded that, even if this increase in acceleration disappeared and the existing rate was maintained up to the post-tests (i.e. covariance grew by the same +2.8%), we could predict an eventual impact upon this cohort of +9.35%. If, on the other hand, the rate of covariant gain continued to increase, eventual impact would be over 10% for Cohort I.

It is now evident that this did not occur in the final year – the mean literacy covariant grew by just +0.45% to a project total of **+6.85%**. In other words, the acceleration noted in the previous reports has disappeared as the LFL came to an end - the project group has only very slightly increased its gain over control since 2003. Nonetheless, the eventual figure is still significant in terms of project impact - there is a 100% statistical certainty that the difference has been caused by the LFL.

It will also be noted that the mean literacy covariant obscures the difference in outcome when writing and reading are compared. The impact for reading at **+8.4%** is +3.1% higher than the figure of **+5.3%** recorded for writing. This is consistent with the scores contained in previous reports and with the data derived from lesson observations throughout the study. There is little doubt that writing remains a very significant problem in all of the schools, both project and control. In neither type of school are children generally required to write more than a couple of sentences and the vast bulk of all the writing they do is of the 'fill in the missing word' variety. The fact that the LFL *has* achieved significant impact in improving the writing scores of project pupils is genuinely meaningful as a project outcome but it has to be acknowledged that the problem is not yet solved.

On the other hand, the higher gain for reading is much more significant as a project impact. Interviews with teachers and principals, as well as lesson observations, support the conclusion that there has been a real improvement in the reading abilities of the children in project schools. This is partly due to the greatly increased supply of books and other materials enjoyed by the project schools and partly to the INSET and monitoring with which they have been provided.

In few of the control schools, irrespective of the intentions or capacities of the teachers, do children ever have even the opportunity to read anything more than a textbook. Virtually none of the

schools, project or control, reported that they had been supplied with, or themselves purchased under Section 21, reading books in any quantity; the emphasis is on attempting to re-supply children with the basic teaching and learning books and materials for the various subjects for the new curriculum. The situation is severely exacerbated by the fact that the majority of all schools have stored books and materials supplied in the past that they do not use because they are 'from the old curriculum'. Amongst these materials are often multiple copies of familiar language texts like 'English – The Active Way' or 'Day-by-Day', etc, virtually all of which contain a wide variety of reading and comprehension passages that remain readily usable in classrooms today.

In the project schools, by contrast, there are multiple copies of all kinds of reading books available to children through the LFL, while teachers were trained and monitored by project staff to ensure that they were, in fact, regularly used.

In terms of the oral use of English, there is no doubt that the first goal of all schools, parents, teachers and pupils is a sufficient level of verbal mastery of English. So long as the majority of School Governing Bodies and the DoE continue to opt for this language as the medium of instruction at Intermediate Phase and beyond, education cannot continue in any meaningful form without this ability, even if 'code switching' is employed.

An improved level of ability to understand and communicate verbally in English has been cited as a major project impact throughout the evaluation by both principals and teachers. This ability has also been evident in lesson observations in project classrooms where the difference with control schools can be startling. Even very young Foundation Phase pupils in some rural schools have been able to sustain simple conversations with researchers in English, and the pupil level of comprehension of the lessons themselves was noticeably higher in many project schools, urban or rural.

The lesson observations throughout the evaluation, and all the other research projects ***ESA*** has carried out, have indicated that children still spend most of their time in classrooms, whether project or control, either speaking or listening, much less reading and very little writing. Although there are clear differences between project and control schools that are reported and discussed in the following section of the report, this situation remains generally true – education in the majority of South African schools remains essentially a verbally-based activity.

Much of an LFL-based lesson is also verbally-based with a strong emphasis on questioning, whether for comprehension or, less often, content. When supported by an extra supply and variety of books on which to base lessons and class, group or individual discussions, the result is evidently a significant improvement in the oral comprehension and use of English. That this was not achieved to the same degree on the conventional indicators of post-basic literacy, reading and writing, should not detract from this outcome even though it could not be measured by the instruments used in the study.

Finally, in terms of project impact, it should be noted that, in 2001 after the first year of the LFL, the evaluators were asked to provide a comment on factors that could affect the eventual outcome of Business Trust projects. One of our comments then read:

*'Previous READ projects have generally been delivered at classroom level within the context of a formal language programme, for example the MAPEP or Day-by-Day series. With the decreased emphasis upon standardized learning programmes in general, and upon text books in particular, this context is altering toward a situation in which READ materials are sometimes used in their own right as the primary language programme in Foundation Phase. The longer-term*

*consequences are unclear but we have already suggested that the organization consider the implications, especially in terms of specific work sequences, linked to sequential language content, for the materials.'*

This point, in our opinion, remains pertinent today. There are no longer any standardized language programmes employed by all schools but rather a wide variety of different programmes. While all may be approved by the relevant selection committees, there is an enormous range in methods, approaches, emphases and, most importantly, quality in these programmes – some could hardly be called a programme at all.

READ has developed or supplied language materials as part of the LFL; for example, the 'New Heights' series, and most of the other courses have at least some language exercises built into them. Nonetheless, the fact is that, in many of the project schools, the LFL was carrying the burden of providing both initial basic oral literacy at all grade levels *and* its evolution into the more developed range of competencies called literacy, of which verbal proficiency is a necessary, but not sufficient, precondition. In my opinion, the LFL would have had more impact on reading and writing if children were generally more proficient in oral English at all grade levels in the first place.

### 2.2. Cohort Two

Cohort II completed its life in 2003 and the analysis contained in the fourth report reflected its final impact. That analysis is repeated here for the sake of completeness.

**Table 14: Cohort II covariance 2000 – 2003 (%)**

| | n | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|---|
| **Project** | | | | | | |
| Baseline | 1 054 | 8.4 | 28.5 | 18.45 | 22.8 | 20.6 |
| Final | 727 | 34.7 | 55.0 | 44.85 | 54.7 | 49.8 |
| Change | | +26.3 | +26.5 | +26.4 | +31.9 | +29.15 |
| **Control** | | | | | | |
| Baseline | 674 | 8.4 | 27.5 | 17.95 | 23.6 | 20.8 |
| Final | 423 | 31.3 | 50.5 | 40.9 | 55.0 | 47.95 |
| Change | | +22.9 | +23.0 | +22.95 | +31.4 | +27.2 |
| **Covariant** | | **+3.4** | **+3.5** | **+3.45** | **+0.5** | **+2.0** |

The table confirmed that Cohort II continued to reflect evidence of significant project impact. However, it did not achieve an acceleration in its rate of gain to the same degree as did Cohort I up to the fourth report. The covariant difference between baseline and the first mid-tests was –0.6%, by the second mid-tests the difference had widened by +3.25% to +2.65%, while by the final tests it had slowed to +0.8% to result in an eventual impact of **+3.45%** for the mean literacy covariant. Interestingly, this cohort did not demonstrate the same differential impact upon writing and reading as did Cohort I.

The fourth report went on to comment that it was evident that the pattern of covariant scores predicted in the second report had been maintained throughout the study – both cohorts provided evidence of significant project impact with the longest lived cohort, beginning at the lowest grade level, reflecting the greatest impact. The clear conclusion, and one with important implications for further applications of the LFL model, is that impact will be greatest when the project is implemented over the longest possible period *and* begins at the lowest possible grade level. Results

obtained in this study suggest that the older children are, and the longer they have been exposed to non-project-based schooling, the lower the eventual level of impact is likely to be.

In any event, the positive covariant obtained from this cohort, and the fact that its behaviour had been predicted from the first report, provides support to the impact data derived from Cohort I.

### 2.3. Grades Five and Seven

We have already noted that data drawn from different groups of children at the same grade level is more open to influence from uncontrolled variable factors than is that drawn from the longitudinal tracking and is, hence, inherently less reliable as a measure of project impact. Nevertheless, discussion about rising or falling 'standards' are effectively based upon this form of comparison.

**Table 15: Grade 5 covariance 2000 – 2004 (%)**

| | n | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|---|
| **Project** | | | | | | |
| Baseline | 1 037 | 12.6 | 24.4 | 18.5 | 32.5 | 23.2 |
| Final | 1 019 | 19.0 | 40.95 | 28.0 | 39.45 | 33.1 |
| Change | | +6.4 | +16.55 | +11.5 | +6.95 | +10 |
| **Control** | | | | | | |
| Baseline | 667 | 12.7 | 22.2 | 17.45 | 33.0 | 22.6 |
| Final | 605 | 12.0 | 30.6 | 21.3 | 37.1 | 26.6 |
| Change | | -0.7 | +8.4 | +3.85 | +4.1 | +3.9 |
| **Covariant** | | **+7.1** | **+8.15** | **+7.6** | **+2.85** | **+6.0** |

**Table 16: Grade 7 covariance 2000 – 2004 (%)**

| | n | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|---|
| **Project** | | | | | | |
| Baseline | 986 | 29.0 | 41.6 | 35.3 | 49.2 | 39.9 |
| Final | 666 | 38.1 | 60.7 | 49.4 | 52.9 | 50.6 |
| Change | | +9.1 | +19.1 | +14.1 | +3.7 | +10.6 |
| **Control** | | | | | | |
| Baseline | 587 | 30.7 | 43.3 | 37.0 | 53.1 | 42.4 |
| Final | 388 | 33.4 | 53.3 | 43.35 | 52.0 | 46.2 |
| Change | | +2.7 | +10.0 | +6.35 | -1.1 | +3.9 |
| **Covariant** | | **+6.4** | **+9.1** | **+7.75** | **+4.8** | **+6.8** |

All of the five comparisons for both grade levels result in a covariant that indicates significant project impact. Even if we double the threshold for significance from 2% to 4% to compensate for the inherent unreliability of this form of comparing pupil performance, nine of the ten covariant scores remain statistically significant – there is a 100% certainty that they have been caused by the LFL project.

If we consider only the two literacy scores for both grades, the mean covariant is **+7.7%**, very similar to the positive covariant of **+6.85%** obtained by Cohort I. Again, this provides clear support to the accuracy of the data for that cohort.

## 2.4. Repeat Rates

The original design for the evaluation used relative changes in pass rates, as defined by internal school assessments, as the second of the impact indicators. In terms of the LFL itself, it was logically assumed that any increase in language performance should be reflected by a corresponding general improvement in performance in subjects taught in that language and that, as a consequence, overall pupil repeat rates should decline in project schools. The systemic reasoning was that an improvement in these rates by the LFL would indicate an ability of the project model to enhance the cost effectiveness of the education system; each pupil who repeats a year consumes an extra year of scarce resources.

The reported figures were constructed by recording pupils who had failed the previous year as the repeats for the current year (i.e. 2003 failures were counted as 2004 repeats) on the grounds that if they were not repeating at the particular school, they should be repeating at another. This method was adopted because pupils who fail at one school can transfer to another school and not be counted as a repeat in either.

**Table 17: Pupil repeat rates 2000 – 2004 (% of enrolment)**

| | **Rate** |
|---|---|
| **Project** | |
| Baseline | 8.5 |
| Final | 5.9 |
| Change | -2.6 |
| **Control** | |
| Baseline | 7.8 |
| Final | 5.5 |
| Change | -2.3 |
| **Covariant** | **+0.3** |

As all the previous reports have noted, there is no evidence of significant project impact on this indicator. However, previous reports have also noted that repeat rates, already low at baseline, have been falling steadily over the last four years in *both* project and control groups, and that one must assume there is a 'floor' to them unless state policy dictates otherwise. The real question, of course, is the degree to which the generalized decline in the repeat rate is due to a generalized increase in the level of pupil performance. The reasons given by school-level interviewees for the general decline in repeat rates are:

- a belief that 'in OBE nobody fails', combined with a universal feeling that schools must reduce failures
- the restriction of failures to one per phase for each pupil
- the reviewing of school-level failures by the DoE, or by the school itself on the basis of DoE-based guidelines, resulting in the passing of a number of children who had been failed by internal school assessments – a practice referred to as 'condoning' their promotion.

While the effect of the first on repeat rates is impossible, and the second very difficult, to quantify it is possible to do so for the third factor. Figures were collected for 2002 and 2004.

**Table 18: Effect of condoned passes on repeat rates: 2002 (% of enrolment)**

| | **Failed by school** | **Actual repeats** | **Difference** |
|---|---|---|---|
| Project | 9.5 | 8.0 | -1.5 |
| Control | 11.0 | 9.8 | -1.4 |

**Table 19: Effect of condoned passes on repeat rates: 2004 (% of enrolment)**

| | Failed by school | Actual repeats | Difference |
|---|---|---|---|
| Project | 7.5 | 5.9 | -1.6 |
| Control | 7.9 | 5.5 | -2.4 |

We now have clear evidence that condoning is the most significant factor sustaining the generalized decline in repeat rates. Whatever the consequences for the education system in general, virtually all of the decline in both project and control groups over the period of the study can be accounted for by this practice. There is, in other words, an external factor affecting both groups that obscures any potential project impact and renders the indicator unusable for its measurement.

Rather than abandon the indicator altogether, and since we did still have available all of the scores obtained throughout the evaluation on the three instruments, it was decided to use the existing data to simulate the measurement of pass rates. Three measures of central tendency were taken for the baseline scores; arithmetic mean, median and mode. Of the three measures, the median - the score at which exactly the same number of children attain either a higher or lower figure - provided the best distribution of scores for the purpose and provided the 'pass mark' standard against which post-test scores could be compared.

**Table 20: Pass marks for the instruments (%)**

| | n | Writing (30 items) | Reading (68 items) | Numeracy (48 items) |
|---|---|---|---|---|
| Grade 5 | 1 037 | 5.0 | 18.4 | 28.1 |
| Grade 7 | 986 | 30.0 | 44.8 | 54.2 |

It is important to keep in mind that the instruments were not originally intended to measure performance against any particular normative standard but were designed to measure covariance over a period of time, five years in the case of the LFL. As a consequence, they were designed to yield baseline scores over a wide range of grade levels that are low enough to allow sensitive measurement of eventual project impact on the same group of children. On the other hand, a norm-based instrument would typically yield a spread of scores around an arithmetic mean point of approximately 55/65% while the study instruments yielded a mean in reading, for example, of 24.4% and 41.6% for Grades 5 and 7, respectively, at baseline.

Secondly, the sensitivity of the instruments is increased by converting scores into percentages when they contain less than one hundred items. Each extra item correct, in other words, is worth more than one extra percentage point in score.

**Table 21: Percentage increase in score for each extra item correct**

| Writing | Reading | Numeracy |
|---|---|---|
| 3.3 | 1.4 | 2.1 |

The implication of all this is that the tables that follow must be read and interpreted with caution. They measure pass rates against their own internal standards and the figures should not be regarded as measurements against a national standard, or any other kind of external standard. Further, pass rates are based on the full set of subjects taken at each grade level whereas there are only three instruments used in the study.

The tables record the percentages of children who achieved a score equal to, or higher than, the median score obtained by that grade level at baseline.

**Table 22: Grade 5 pass rates 2000 – 2004 (% of pupils)**

| | n | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|---|
| **Project** | | | | | | |
| Baseline | 1 037 | 49.9 | 50.8 | 50.35 | 57.8 | 52.8 |
| Final | 1 019 | 64.6 | 83.8 | 74.2 | 79.1 | 75.8 |
| Change | | +14.7 | +33.0 | +23.85 | +21.3 | +23.0 |
| **Control** | | | | | | |
| Baseline | | 53.4 | 49.4 | 51.4 | 57.7 | 53.5 |
| Final | | 52.2 | 68.3 | 60.25 | 73.0 | 64.5 |
| Change | | -1.2 | +18.9 | +8.85 | +15.3 | +11.0 |
| **Covariant** | | **+15.9** | **+14.1** | **+15.0** | **+6.0** | **+12.0** |

**Table 23: Grade 7 pass rates: 2000 – 2004 (% of pupils)**

| | n | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|---|
| **Project** | | | | | | |
| Baseline | 986 | 41.4 | 43.8 | 42.6 | 39.3 | 41.5 |
| Final | 666 | 57.8 | 70.65 | 64.2 | 50.8 | 59.75 |
| Change | | +16.4 | +26.85 | +21.6 | +11.5 | +18.25 |
| **Control** | | | | | | |
| Baseline | 587 | 49.0 | 46.7 | 47.85 | 51.4 | 49.0 |
| Final | 388 | 48.7 | 59.1 | 53.9 | 46.2 | 51.3 |
| Change | | -0.3 | +12.4 | +6.05 | -5.2 | +2.3 |
| **Covariant** | | **+16.7** | **+14.45** | **+15.6** | **+16.7** | **+15.95** |

The effects of the LFL are clearly evident in these figures. Given the discussions above, they all reflect significant positive covariance in favour of the project group and there is no statistical doubt that they reflect project, rather than other, effects. It will be noted that, with the exception of numeracy at Grade 5 level, all of the covariant scores are above 11%.

## 3. Conclusion

The consistency of the data is unmistakable; all of the twenty covariant scores presented in the analysis are positive, even if two of them – numeracy for both cohorts – are too low to be statistically significant. The data clearly indicates that the LFL has achieved significant impact upon project pupils and that this impact has been caused by the project to a 100% degree of certainty.

The data for Cohort I is the most reliable of the different comparisons and, in the opinion of the evaluators, it is the measure most likely to reflect the actual impact of the project. The rest of the comparisons establish that positive impact has certainly occurred; Cohort I provides the fine measurement.

**Table 24: Summary of all covariant Scores**

| | Writing | Reading | Mean Lit. | Numeracy | Mean L. & N. |
|---|---|---|---|---|---|
| **Cohort I** | **+5.3** | **+8.4** | **+6.85** | **+1.0** | **+4.9** |
| Cohort II | +3.4 | +3.5 | +3.45 | +0.5 | +2.0 |
| Grade 5 | +7.1 | +8.15 | +7.6 | +2.85 | +6.0 |
| Grade 7 | +6.4 | +9.1 | +7.75 | +4.8 | +6.8 |
| Pass Rate G5 | +15.9 | +14.1 | +15.0 | +6.0 | +12.0 |
| Pass Rate G7 | +16.7 | +14.45 | +15.6 | +16.7 | +15.95 |

The social significance of the absolute degree of change achieved by the Learning for Living project in relation to the needs of the national education system, ***in the absence of a comparative and longitudinal study combining data from a variety of language and book-based programmes***, is essentially a political and economic question. Analysis and statistics have taken us as far as they are capable without a comparative study; they have provided the measure of project impact as accurately as possible – it is up to others to decide the social meaning of the figures.

## SECTION FOUR

## PRINCIPAL AND TEACHER INTERVIEWS

### 1. Class Size and the Management of Human Resources

The baseline report noted that there was a very wide range of actual class sizes in the different provinces and this has continued to be the case throughout the study. In particular, those provinces affected by extensive shortages of classrooms, like Mpumalanga and Limpopo, continue to have very large class sizes.

What is of particular interest to this study is the actual usage of human resources in the schools; how effectively do schools use their teaching staff to produce class sizes that are close to the theoretical possible sizes (i.e. pupils/teacher ratio)? It is understood that some teachers have special duties and responsibilities which take them away from classroom teaching and we would not, therefore, expect a perfect match between the two figures. However, it can reasonably be assumed that the closer actual class sizes are to the theoretical figure, the more likely the school is managing its human resources effectively.

**Table 25: Comparison of theoretical pupils/teacher ratio and actual class size**

| | #P/T Ratio | *Enrolled class size | Difference |
|---|---|---|---|
| **Project** | | | |
| Baseline | 38.3 | 45.4 | +7.1 |
| Final | 36.6 | 43.7 | +7.1 |
| Change | -1.7 | -1.7 | 0.0 |
| **Control** | | | |
| Baseline | 36.9 | 44.7 | +7.8 |
| Final | 35.8 | 45.0 | +9.2 |
| Change | -1.1 | +0.3 | +1.4 |
| **Difference** | **-0.6** | **-2.0** | **-1.4** |

**#** *Total school enrolment divided by total number of classroom teachers*

***** *Mean enrolled class size of lessons observed*

The usage of human resources at both project and control schools is clearly not great – there are differences of +7.1 and +9.2 more pupils per class respectively than the overall pupil/teacher ratios. However, while the pupils/teacher ratio has improved in both types of schools, only in the project schools have class sizes decreased correspondingly. In control schools, by contrast, actual class sizes have actually *increased* despite an improvement in their pupils/teacher ratio and, as a consequence, in project schools the class sizes are 1.4 pupils per class closer to the overall pupils/teacher ratio.

The comparative difference between the two groups, though not large, suggests that the management training component of the LFL has resulted to some degree in a more effective use of human resources in project schools.

### 2. Amount and Payment of Fees

The baseline report pointed out that the amount of fees charged, and the level of collection of those fees, is of interest to this study in two ways. First, without a significant income from fees, schools

are unable to purchase materials and other teaching aids from their own income. Second, it is possible that perceptions by parents and guardians of increased quality in the schools as a result of participation in the project could result in either a willingness to accept fee increases or in the total proportion of fees that are collected.

It has been argued that a low level of fees and a low payment rate indicates, other than the obvious factor of poverty, a reluctance on the part of parents to pay for what they perceive to be low quality education in schools. Many principals and teachers point to parents who would not pay fees at their schools but were willing to pay very much higher fees at ex-Model C or township private schools, or to pupils who regularly spend far more than their school fees on buying food and drinks from the ubiquitous 'mamas' outside virtually all of the schools. The implication is that should parents feel their children are receiving higher quality education in project schools their willingness to pay higher fees would increase – thus providing a potential indicator for the perceived improvement of higher quality as a result of the Learning for Living Project.

**Table 26: Payment and cost of fees**

| | Mean cost of fees p.a. | % of fees paid |
|---|---|---|
| **Project** | | |
| Baseline | R37.40 | 66.1 |
| Final | R56.90 | 64.4 |
| Change | +R19.50 | -1.7 |
| **Control** | | |
| Baseline | R45.91 | 72.3 |
| Final | R78.20 | 68.4 |
| Change | +R32.29 | -3.9 |
| **Difference** | **-R12.79** | **+2.2** |

In both types of schools, the level of fees payable has increased since 2000 though the increase is R12.79 p.a. less in project schools; it may be that the supply of materials to these schools by the LFL has reduced their perceived need to purchase their own materials. What is disturbing in the general sense is that in both types of schools the proportion of parents prepared to pay these fees has actually *declined* since baseline. However, in the project specific sense, the decrease has been comparatively lower in project schools by 2.2%. While the difference is too small to be considered really significant, there is at least a suggestion that the presence of the LFL in project schools has slowed the general decline in willingness to pay school fees.

### 3. Availability of Teaching and Learning Materials

The principal and teacher interviews for all of the reports for this study have included a question asking participants whether their schools had sufficient teaching and learning material - defined as sufficient material for *each* pupil for *all* of the subjects/learning areas presented by the school. The baseline report commented that in many of the schools where participants answered that they did not have sufficient materials, and there are extraordinarily high numbers of such schools, storerooms contained discarded and unused material not considered when giving their answer to the question. Certainly some of this material, most of it provided over the last ten years, was outdated, especially historical and social studies type material, but the vast bulk of it appeared readily usable. In other schools, storage was chaotic and existing material, even if in use, was degenerating rapidly and visibly. These conditions have remained stable throughout the evaluation.

The figures reported for these items in each report of this study have always been estimates by respondents and it has become increasingly evident that the definitions upon which the estimates are based have changed from year to year; schools now have materials, provided over the last ten years, that are based on three different curricula – the original pre-1994 curriculum, C2005 and the Revised National Curriculum Statement (RNCS). This has resulted in widely differing figures, though always reflecting significant shortages, and the obvious lack of consistency has made it unrealistic to report figures in terms of a precise measure like covariance. Consequently, the table below is based upon responses for 2004 only and reflects a simple project/control comparison for that year.

**Table 27: Available materials in schools and classrooms (%)**

| | Principals | | Teachers | |
|---|---|---|---|---|
| | **Fully supplied** | **Insufficient** | **Fully supplied** | **Insufficient** |
| Project | 40.8 | 58.2 | 51.6 | 48.4 |
| Control | 23.3 | 76.7 | 35.1 | 64.9 |
| **Difference** | **+17.5** | **-17.5** | **+16.5** | **-16.5** |

While the effects of the large-scale supply of materials by the LFL to project schools are clearly evident in the table, it continues to reflect a continuing shortage of materials reported by principals and teachers in both types of schools. Significant as the scale of materials supply by the project is, it clearly cannot make up for the shortages across the whole range of Learning Areas. The project materials are universally regarded as fundamentally dealing with LLC and, especially, English and, while the various collections do include materials relevant to other subjects, it would obviously be unrealistic to expect the project to have achieved, in terms of both cost and logistics, a scale of delivery many times larger – only the state is capable of supply at this level, or of ensuring that schools use *all* of the existing materials at their disposal.

### 4. Perceptions of School Quality

Principals and teachers were asked to rate the quality of their schools in relation to other schools in their local communities. The question provided a 5-point response scale where 5 = excellent and 1 = very poor. The mean of the responses was calculated for each group; a score of 3 indicates a perception of average quality whereas a score over 4 indicates a belief that schools are well above average.

**Table 28: Perception of school quality: Principals**

| | Principals | Teachers |
|---|---|---|
| **Project** | | |
| Baseline | 4.1 | 3.7 |
| Final | 4.2 | 4.2 |
| Change | +0.1 | +0.5 |
| **Control** | | |
| Baseline | 3.7 | 3.4 |
| Final | 3.8 | 3.5 |
| Change | +0.1 | +0.1 |
| **Difference** | **0.0** | **+0.4** |

Principals and teachers in both types of schools have always held a high opinion of the relative quality of their schools as compared to comparable schools in their local communities. It should be said at the outset that neither the pupil testing nor the lesson observations suggested that the quality of the majority of project and control schools were or are, objectively, so significantly above average.

In terms of relative change over the life of the project, the difference in the level of perception of quality of project and control principals pre-dated the project and there exists no evidence of post-project change in this regard – most project principals, in other words, already considered the quality of their schools to be very high in the first place and this perception was confirmed by their selection (recognition) for participation in the LFL. Teachers in project schools, on the other hand, have become more optimistic about the quality of their schools. This is a significant change and is clearly explained by the LFL; English and, to a lesser degree, other language teachers, in particular, are convinced that they are offering a much better quality of education than are their counterparts in control schools. The increased ability of project pupils to understand and use English, especially in the verbal, allows schooling in all subjects to take place in that language, and makes much more accessible other materials which are almost always in English from Intermediate Phase onwards.

Principals and teachers regularly cite the efforts of parents to enrol their children in project schools as evidence that the school in general is providing a better education – because of the presence of the LFL and its perceived effects upon the understanding and use of English by the school as a whole. Motivational events, like Readathon, have a significant effect in supporting these perceptions and play a major role in generating and confirming the local-community support of the LFL.

**5. Attitudinal Reponses to the Project**

The tables that follow were all originally constructed for previous reports from focus questions which did not provide set answer codes. Responses were, rather, unprompted and later collated into broad categories. The method was adopted because the intention was to provide READ with useful and reliable information about the attitudes and opinions of project participants as they evolved over the life of the evaluation; providing set response options can result, for example, in participants appearing to have more information and opinions about the project than is actually the case. For the final report, however, all of the previously used categories were provided as a set of closed responses with an *other* option. Interviewees were allowed to nominate a maximum of five categories to ensure a degree of value-based choice in the responses. The figures reflect the percentage of interviewees that nominated each category, and the questions were necessarily restricted to principals and teachers in project schools.

| **Table 29: Most significant impacts of the project** | Principals (%) |
|---|---|
| Improved teacher methodology | 79.6 |
| Materials supply | 69.4 |
| Improved school management | 61.2 |
| Improved oral communication of pupils | 59.2 |
| Improved reading by pupils | 40.8 |
| Improved pupil attitudes/confidence | 36.7 |
| Improved teacher attitudes/confidence | 32.6 |
| Improved understanding/delivery of OBE/C2005 | 26.5 |
| Increased joint planning by staff | 22.4 |
| Pupils improved in all subject areas | 16.3 |
| Increased community support | 12.2 |
| Dramatizations by pupils | 12.2 |
| Improved writing by pupils | 10.2 |
| None | 2.0 |
| *Other: Helped with incoming rural pupils, Became local leaders in education, Improved appearance of classrooms* | *10.2* |

Principals have remained thoroughly and universally positive about the impact of the project throughout the study. Given the reported shortages of materials in schools, the supply of such materials is very highly valued. They are generally convinced that they are managing their schools better, that teachers are presenting better quality lessons in terms of classroom methodologies and are almost as convinced that pupils are much more able to verbally communicate in English – understandably a highly valued skill, given the near universal choice of this language as the medium of instruction. Many, though still a minority, asserted that pupils' reading ability had also improved though few believed that pupil writing had significantly improved.

| **Table 30: Most significant impacts of the project** | Teachers (%) |
|---|---|
| Materials supply | 74.7 |
| Improved teacher methodology | 72.6 |
| Improved oral communication of pupils | 57.9 |
| Improved reading by pupils | 52.6 |
| Improved pupil attitudes/confidence | 31.6 |
| Improved teacher attitudes/confidence | 28.4 |
| Pupils improved in all subject areas | 22.1 |
| Improved understanding/delivery of OBE/C2005 | 17.9 |
| Improved school management | 16.8 |
| Improved writing by pupils | 16.8 |
| Increased joint planning by staff | 14.7 |
| Improved classroom management | 11.6 |
| None | 4.2 |
| *Other: Improved teachers' English, Dramatizations by pupils, Increased community support, Pupils work on their own, Can use English as MOI from Grade 1.* | *10.5* |

Teachers, like principals, are almost universally positive about LFL. The materials that have been supplied by the project have found strong approval from them, partly because of the shortages of similar materials in schools noted elsewhere, but primarily because of the relevance, quality and level of the books and other learning materials. They are seen to make the practice of OBE more possible and, importantly, are explicitly linked to the INSET courses – a quality which much of the rest of the materials they receive does not share - almost as many teachers believe that their own teaching methodology has improved as a consequence.

Improved oral communication of pupils is seen as the most significant pupil impact of the project and this is discussed in both Section Three and Five of this report. Many teachers also report that they believe that the reading levels of their pupils have improved, a belief that found confirmation in testing, though it is clear that there is still a long way to go before genuine higher-order reading competencies are likely to be achieved. As for principals, very few mentioned improved writing by pupils as a project – a conclusion confirmed by both testing and lesson observations.

All of the reports to date have mentioned that principals and teachers are reluctant to appear critical of the project. Because the question has the potential to provide useful formative information to project management and implementers, we made some effort to press interviewees to, at least, make suggestions for project improvement. Even then almost 20% of them simply insisted that there were no problems of any kind and nothing that needed improvement.

| **Table 31: Most significant problems experienced during project** | **Principals (%)** |
|---|---|
| Clashes with DoE systems, especially programme/lesson planning | 38.7 |
| Teacher resistance | 28.6 |
| No support from DoE | 24.5 |
| Materials/methods apply primarily to language teachers | 22.4 |
| Organizational confusion/strained relation with trainer | 20.4 |
| Too much work | 18.4 |
| None | 18.4 |
| *Other: Disruptive to school programme, No certificates yet, Nowhere to store books, Poor school admin., Materials too difficult, Not enough monitoring, Not enough classrooms, Need more resources* | *16.0* |

| **Table 32: Most significant problems experienced during project** | **Teachers (%)** |
|---|---|
| None | 33.7 |
| Clashes with DoE systems, especially programme/lesson planning | 24.2 |
| Too much work | 20.0 |
| Need more resources | 15.8 |
| Materials/methods apply primarily to language teachers | 14.7 |
| No support from DoE | 13.7 |
| Need more monitoring | 11.6 |
| Organizational confusion/strained relation with trainer | 11.6 |
| *Other: Poor school admin., Materials too difficult, Disruptive to school programme, Not enough on writing skills, Clash with other NGO methods, Need more materials in other languages.* | *13.7* |

There do appear, in some areas, to have been problems with administrative arrangements; notice of courses being received late or the incorrect teachers nominated for attendance, etc., and a number mentioned clashes with DoE methods and planning systems.

The DoE, in some provinces, applies a different format for course and lesson planning than did the LFL. In some cases, LFL trainers insisted on the use of project methods which meant that teachers were forced to produce two sets of plans, one for the DoE and one for the LFL. The disjunction was remedied in some of the provinces or districts in which LFL operated but it was still an issue for schools and teachers in others by the end of the project.

Just under a quarter of the teachers reported that they continue to have problems with insufficient supplies of materials. In general, this is more a reflection of the large shortage of state-supplied materials than any problem that can be ascribed to the LFL but, in other cases, teachers have pointed out that if more copies of individual titles were supplied the project, methods would be more usable in large classes.

**5. Sustainability of the project**

There are two dimensions to the question of sustainability. The first is concerned with the ability of the schools that participated in the LFL to continue using the materials and skills they have gained from the project now that it has completed its life. The second is concerned with the possibility that the effects of the project can be spread to other schools, chiefly through the benefits gained by the local DoE during project implementation, supported by the existing project schools.

In regard to the first, both principals and teachers are overwhelmingly confident that the project effects can be sustained in their own schools; **85.7%** and **88.6%** respectively believe that it is very likely or likely to be so. On the other hand, they are equally pessimistic about the possibility of project spread lead by the DoE with only some **12.2%** expressing a contrary view. Except for a few exceptional cases, local officials of the department have not had sufficient exposure to the project training and materials, nor sufficient experience in monitoring its implementation, to reproduce it in other schools – even when they are supportive of the project in principle. It is true that school clusters have been set up by project staff and that some of these clusters share materials and classroom visits with schools that did not participate in the LFL, but we do not believe that project effects can be widely spread without the leadership of the DoE.

## SECTION FIVE

## LESSON OBSERVATIONS

### 1. Grades and Subjects

In all, 152 lesson observations were completed, 94 in project and 58 in control schools.

Since the project design involved the inclusion of non-language teachers at project schools, previous reports have tried to include these teachers in field research while reserving at least half of the interviews and lesson observations for language teachers. However, it became clear during the course of the study that many, if not most, non-language teachers have had minimal contact with INSET and the materials - field research for the previous reports indicated that they knew little about the project and seldom used the project materials in classrooms.

Consequently, for the final report we requested project schools to nominate two teachers, both of whom should have been directly involved with the LFL Project - i.e. received INSET, materials and classroom monitoring - and one of whom had to be a language teacher. The other could be from any learning area, preferably numeracy or mathematics, but if none of these teachers had been directly involved in LFL it would be preferable to instead select two language teachers. In the event, almost three quarters of the teachers nominated by project schools were either LLC or English teachers.

**Table 33: Lesson observations**

| Grade | Project | Control |
|---|---|---|
| Multi-grade | 4.3 | 3.4 |
| 1 | 4.3 | 8.6 |
| 2 | 24.5 | 13.8 |
| 3 | 9.6 | 8.6 |
| 4 | 13.8 | 15.5 |
| 5 | 14.9 | 8.6 |
| 6 | 20.2 | 24.1 |
| 7 | 8.5 | 17.2 |

| Subject | Project | Control |
|---|---|---|
| LLC | 13.8 | 13.5 |
| English | 58.6 | 41.4 |
| Numeracy | 19.1 | 36.2 |
| *Other* | *8.5* | *8.6* |
| | *A&C, EMS* | *EMS, NS,* |
| | *HSS, NS* | *HSS* |

### 2. Use of READ Methods and Materials

The baseline report commented that it was important that project teachers develop the classroom methods, activities and behaviours promoted by the LFL on a sufficiently consistent basis. If this did not occur, and if these kinds of qualitative changes were inconsistent across the national sample, it was unlikely that the predicted pupil impact which is dependent on these behaviours would be consistently achieved.

All of the subsequent reports have noted that teachers routinely used the materials on a regular basis, most of them along with the project-supplied teacher guides, and the consequence is that a READ-based lesson, in its various forms, is recognizably such in all of the project schools. There was clear evidence throughout the study that the INSET and monitoring provided by the project succeeded in achieving a consistent impact on the classroom behaviour of teachers across the sample. This was the central qualitative process finding of all of the previous reports with regard to classroom level teaching and it remains true.

As we have already seen, these changes have succeeded in achieving a significant quantitative impact upon the performance of children taught by LFL-trained and supplied teachers. As a consequence, the second of the evaluation questions upon which the research design is based; *'Has the project achieved its predicted qualitative impact (the process means of the project) on a consistent basis in the project schools?'*, has been answered in the affirmative.

The third and fourth reports went on to remark that, while a remarkably even and consistent spread of project methods and materials had been achieved, the degree of *quality* to which these methods and materials were used varied widely. Secondly, it commented that, by the final phase of the project, it would become increasingly interesting to see whether teachers also showed development toward 'transference' of these methods/activities/behaviours to their routine teaching activity.

By the end of the project, it was evident that, while many teachers *are* generally presenting LFL-based lessons as such to a higher level of quality, there were still few who could do more than directly reproduce them step-by-step. Few appear able to integrate the LFL-based methods with materials from other sources, and even fewer use them as part of a mix of methodologies and approaches to be used as appropriate. In this sense, *substitution* rather than *transference* has been more likely to occur.

The problem of 'transference' is certainly not unique to the LFL project and the clear implication is that it is unlikely to be achieved within the system as a whole in the short to medium term. The evidence available to this study clearly indicates that teachers who are presenting set lessons with materials and guidance provided are presenting better quality lessons than are those who do not have similar resources.

**3. Reading and Writing**

All of the previous reports have commented that a very clear emphasis was placed by project schools and teachers upon oral competence with relatively less attention paid to reading and, especially, writing. That has remained true into the final report.

Reading in the classroom almost always takes one of three forms:

- Teachers read stories to children who listen silently, usually following the text as pointed out by the teacher in a Big Book while he/she reads.
- Teachers read and children repeat what has been read in unison as a whole class, usually paragraph by paragraph, following the teacher's pointer.
- Children read in groups, either in unison or in sequence, usually aloud.

While there is nothing wrong with any of these forms *per se*, and while in control schools even these forms are relatively much less frequently observed (chiefly due to lack of materials with which to apply them), it will be noted that the emphasis remains upon the verbal and that children are virtually never observed reading silently as individuals. Certainly, teachers may not consider this form of reading as *'sexy'* enough for an observation, a comment on the educational value they place upon it in the first place, but it is absolutely certain that it occurs very much less frequently than do the other forms.

In terms of writing, children are seldom asked to produce anything more than single words *(fill in the missing word)* or, less often, generate complete sentences. Virtually never are they asked to produce sustained paragraphs and their workbooks contain next to no examples of such in project or control schools. It is true that children in project schools are more often likely to be asked to

produce complete sentences, usually in comprehension exercises following reading, but any paragraphs they attempt are no more than disconnected short sentences.

Tests have consistently demonstrated that the level of the ability to write in English of the children in both project and control schools is very low indeed. It is not uncommon, in fact quite the opposite, for children to be unable to complete even the simplest sentence stems understandably. Very few of them have control of anything but the simplest and most limited vocabulary, even less can spell words with any degree of accuracy.

On the other hand, there is little doubt that the level of oral comprehension of English of pupils in project schools has improved. Principals and teachers regularly assert that this is so and the difference between project and control schools, especially at Foundation Phase in rural and remote schools can be dramatic.

**4. Most Significant Problems Experienced in Teaching**

This item deals with the problems reported by teachers. In previous reports, the question was open-ended but, for the final report, the categories distinguished in previous interviews were provided as possible responses with interviewees allowed to select a maximum of five categories.

| **Table 34: Most significant problems experienced in teaching** | **Project** | **Control** | **Diff.** |
|---|---|---|---|
| Environmental context of school/pupils | 45.8 | 49.1 | -3.3 |
| Poor preparation of basic skills of pupils | 32.3 | 33.3 | -1.0 |
| Shortage of books, materials, equipment | 31.2 | 59.6 | -28.4 |
| Big classes and no individual attention | 29.2 | 28.1 | +1.1 |
| OBE/C2005 concepts, methods and/or assessment | 28.1 | 38.1 | -10.0 |
| Medium of instruction not understood by pupils | 24.0 | 26.3 | -2.3 |
| Poor pupil motivation and discipline | 14.6 | 33.3 | -18.7 |
| Inadequate support from the DoE | 13.5 | 12.3 | +1.2 |
| No problems | 8.3 | 1.7 | +6.6 |
| *Other: Project – Shortage of classrooms, Homework seldom completed, No remedial teachers. Control – same plus No feeding scheme* | *8.3* | *8.8* | *-0.5* |

Both project and control teachers nominate the environmental context of the schools and the children as the most significant cause of the problems they experience. This was especially true for teachers in rural and, especially, remote areas in which English is seldom heard or spoken, few if any books other than the Bible are available in any language and many children are cared for by a grandparent who herself received very little formal education (the baseline report indicated that 50% of the whole sample did not live with their parents). All of this means that the pupils have very little home support and that homework is seldom completed. Add to this the familiar problems of poverty and underdevelopment in much of the country and it is understandable that environmental context plays such a large role in determining the effects of schooling.

In terms of school-level problems, many of the teachers complain that children arrive in their classes without adequate control of the basic language and numeracy skills required for them to progress, others that they cannot rely on any degree of consistency in what children know or are able to do – *'I must start again with each new class'* in the words of one of the teachers. The result is huge ranges in the levels of ability of children in one class - *"from remedial to very advanced'*. Increasing numbers of teachers express frustration that C2005 does not contain any real guidance about the content matter which pupils are expected to learn, as they do about the assessment

methods and criteria of C2005 – the effects of group assessment, in particular, are seen as both unhelpful and unfair to more able pupils; *"How is it possible for some learners to reach Grade 7 when they cannot read or write and cannot do the simplest sums? All they can do is copy the work of the others."*

There are three areas in which the project appears to have had significant effect on changing the experience of teachers:

- 28.4% less project teachers complained of a lack of teaching and learning materials. This is clearly a direct project impact.
- 18.7% less experienced problems with pupil motivation and discipline. They explained that the methods and materials of the LFL generated more interest and willingness to become involved in pupils.
- 10.0% less reported problems in their understanding and implementation of OBE and C2005. Again, this was directly linked to the INSET, methods and materials in that they are seen as embodying, and making possible, the new approaches.

**5. Use of LFL Materials in Lesson Observations**

Of lessons observed in project classrooms, 75.5% involved the use of project materials while 24.5% did not. The lessons that did not involve the use of these materials were almost all presented by non-language teachers; only 3.1% of these lessons were based on project materials.

The Big Books are by far the most popular of the project materials - 69.0% of all the lessons observed in which LFL materials were used involved the use of these books. The next most popular were the group readers used in 16.9% of the observations. In the remaining 14.1% of the lessons, the box library collection and the New Heights series were most commonly observed in use.

**6. Displays of Pupil Work**

The lack of displays of pupil work in classrooms was one of the most immediately noticeable factors in the schools at baseline, whether project or control. While some of the schools were festooned with different types of pupil work, the vast majority were not, especially beyond the junior primary level. In most of the urban schools there were posters and charts aplenty on the walls but very few examples of pupil work; in the rural areas few schools had either. The baseline report concluded that it would not be an exaggeration to say that there were no more than a couple of schools in the whole sample in which pupils' work is regularly displayed and regularly changed on a meaningful and routine basis.

Successive reports contained clear evidence of project impact, from a position of near equivalence at baseline to a difference of 21.6% more project classrooms in which pupil work was displayed. The second report commented that the change was evidently a project impact promoted and encouraged by the programme. Secondly, since the project commenced delivery at Foundation Phase level, it was logical to assume that differential change should be even greater as project delivery reached higher levels. By the final report the project/control distinction has, indeed, continued to widen.

At Foundation Phase level, in particular, the difference that can be ascribed to project impact since baseline was remarkable - some of the teachers have virtually plastered the walls with pupil drawings, word lists, posters and, in higher grades, book and group reports and short retold stories.

**Table 35: Classrooms in which pupil work is displayed (%)**

| | Yes | Some | No |
|---|---|---|---|
| **Project** | | | |
| Baseline | 13.8 | 23.3 | 62.9 |
| Final | 41.5 | 21.3 | 37.2 |
| Change | +27.7 | -2.0 | -25.7 |
| **Control** | | | |
| Baseline | 13.2 | 13.2 | 73.5 |
| Final | 12.1 | 10.3 | 77.6 |
| Change | -1.1 | -2.9 | +4.1 |
| **Difference** | **+28.8** | **+0.9** | **-29.8** |

**7. Evidence of Teacher-made Materials**

Given the reported, as well as the actual, shortages of teaching and learning material in the schools, and given the demands of the new curriculum for teachers to assemble their own learning programmes from a variety of materials, the ability of teachers to do so is critical if meaningful education is to take place. There was at baseline very little evidence that they were able to do so.

Succeeding reports provided clear evidence of project impact which had started in and spread beyond Foundation Phase, though it was still most evident at this level. There were very many more examples of teacher material, and a much greater confidence in teachers about making such material. While by far the most commonly teacher-produced materials were still posters illustrating things like the days of the week or types of transport, there were also decidedly more examples of teacher-worksheets, fact sheets and task instructions. Most importantly, there were many more examples of Big Books which had been produced by teachers. Most of these books were still very simple, illustrating, for example, the stages of growth of a plant or providing a short and simple summary of a story, but nothing similar was evident in any of the control classrooms. In a few cases, even pupils had made their own books, illustrating what they learn on different days of the week, for example, or describing some event. Again, nothing similar was evident in any of the control classrooms. These were hopeful signs that the materials would continue to provide both an inspiration to create more classroom-level materials and a model for doing so.

**Table 36: Classrooms in which teacher-made materials were evident (%)**

| | Yes | Some | No |
|---|---|---|---|
| **Project** | | | |
| Baseline | 37.1 | 20.7 | 42.2 |
| Final | 47.9 | 34.0 | 18.1 |
| Change | +10.8 | +13.3 | -24.1 |
| **Control** | | | |
| Baseline | 20.6 | 27.9 | 51.5 |
| Final | 13.8 | 36.2 | 50.0 |
| Change | -6.8 | +8.3 | -1.5 |
| **Difference** | **+17.6** | **+5.0** | **-22.6** |

There continues to be clear evidence of project impact, though there is still very little evidence that teachers are able to design complete units of work combining multiple sources and texts. While teachers in project schools continue to develop many more materials than do the teachers in control schools, the bulk of these materials is still either copied directly from LFL materials – chiefly the pupil exercises contained in the teacher guides - or is very clearly based directly upon them –

chiefly worksheets. While a significant number of teachers do continue to produce Big Books, and pupils in a number of schools continue to make small books – chiefly summaries of stories they have read – it does not appear likely that teachers will be able to move on to the development of more complex learning programmes. In my opinion they are likely to remain dependent upon supplied materials and texts.

Nonetheless, the disparity in the production of any kind of materials in favour of the project schools does continue to confirm that the LFL has had a significant impact in this important area of modern South African schooling.

### 8. Modes of Pupil Activity

The baseline report argued that it has long been of concern in South African education that, for the majority of our pupils, schooling has consisted almost exclusively of verbal activity. Children spent most of their time listening to the teacher, responding to his/her questions or talking to each other in groups; they were seldom asked to read and even less were they asked to produce any form of sustained writing. Much of the emphasis in recent years has been upon altering these patterns of classroom activity toward a more balanced mix of the basic modes of pupil activity.

**Table 37: Modes of pupil activity by % of the lesson**

| | Listen | Speak | Read | Write | Activity |
|---|---|---|---|---|---|
| **Project** | | | | | |
| Baseline | 54.5 | 21.8 | 6.1 | 10.4 | 7.1 |
| Final | 55.0 | 18.6 | 8.2 | 10.0 | 4.8 |
| Change | +0.5 | -3.2 | +2.1 | -0.4 | -2.3 |
| **Control** | | | | | |
| Baseline | 54.5 | 23.5 | 8.3 | 10.6 | 3.0 |
| Final | 59.1 | 19.6 | 3.7 | 12.2 | 5.2 |
| Change | +4.6 | -3.9 | -4.6 | +1.6 | +2.2 |
| **Difference** | **-4.1** | **+0.7** | **+6.7** | **-2.0** | **-4.5** |

By end-project it remains evident that verbal activity remains by far the most common mode of classroom activity in both project and control schools, absorbing over three quarters of lesson time, and that reading, writing and pupil activities are still relatively neglected. However, there is clear evidence in the table that the frequency of pupils listening to the teacher and in the amount of reading, noted in the third and fourth reports, has altered in favour of the project schools.

This is clearly a project impact, and a significant one, obviously made possible by the much greater supply of books available to teachers and pupils. If anything, it is surprising that the differences are not much higher; the main reason appears to be that most of the reading observed in classrooms consists of the teacher reading to the pupils who listen as a whole - they very seldom read silently on their own.

### 9. Instructional Focus

The baseline report noted that over recent years the increased use of group work was invariably included in the objectives of virtually every developmental initiative, whether of the DoE or NGO's, sometimes to the virtual exclusion of other forms of classroom instructional focus. As a consequence, classrooms arranged in group seating patterns, usually six to a group, are now the norm in South African primary schools all the way to Grade 7.

However, all of the subsequent reports have found that most teachers still fundamentally direct their lessons to the whole class. All that has changed, in effect, is the seating arrangement, a distinct disadvantage to those pupils who find their backs to the teacher and the chalkboard. Consequently, only instances in which groups were asked to work as a social unit, for even a short period, to perform some substantive task were recorded as observations of group work.

The baseline, second and third reports went on to comment that much of the group work we observed was undirected and unfocussed with minimal tasks predominating and much of the time wasted on trivial activity-based tasks used for their own sake as group work - like cutting-and-pasting at Grade 7 level. In the decidedly fewer observations of effective group work, each member of the group was far more likely to have been given a specific, highly structured and understandable task and teachers had sufficient materials for all of the pupils. They, consequently, had the basis on which to keep all of the pupils meaningfully engaged. These lessons were never based on open ended discussion type tasks and they never dealt with the learning of new concepts. The second report commented that, as project delivery moved progressively into the more senior levels and as teachers become more confident in the use of project materials and methods, we would expect to see a more even spread of instructional focus in classrooms, as well as a greater control and understanding of the use of group work if the project was achieving its own internal qualitative classroom-based objectives.

The fourth report found that there had been a clear shift in favour of 'genuine' group work in the project schools made possible by the provision of materials and the associated teacher guides and after-reading model worksheets. Teachers were more likely to be able to set specific tasks and provide essential materials because the project was *both* guiding them *and* providing them with the means to do so. While it could not be said that there was evidence that teachers are generally able to design and present effective group-based lessons without directly producing READ examples, there were many fewer instances of ineffective group work in project than control classrooms for the this report. The report concluded that the LFL was having an impact on improving the effectiveness of group-based lessons. Observations for the final report have confirmed this conclusion.

**Table 38: Instructional Focus by % of Lesson Time**

| | Whole class | Group | Pair | Individual |
|---|---|---|---|---|
| **Project** | | | | |
| Baseline | 77.1 | 7.5 | 2.4 | 13.0 |
| Final | 79.9 | 9.0 | 0.9 | 10.2 |
| Change | +2.8 | +1.5 | -1.5 | -2.8 |
| **Control** | | | | |
| Baseline | 79.5 | 5.2 | 1.7 | 13.6 |
| Final | 77.4 | 10.4 | 0.9 | 11.3 |
| Change | -2.1 | +5.2 | -0.8 | -2.3 |
| **Difference** | **+4.9** | **-3.7** | **-0.7** | **-0.5** |

While the table does not reflect a significant *increase* in the time spent dedicated to group work, in fact whole class teaching continues to absorb over three quarters of the time in both types of schools, the observations themselves once again provided evidence that project-trained teachers continue to use the method more effectively because of the effects of project materials and training.

The fourth and final rounds of research also suggest that many teachers have recognized the aimlessness of most group work and are increasingly reluctant to use the method. Many teachers in both project and control schools commented that group work allowed weaker pupils to 'hide' or just

coast along without engaging themselves in substantive effort that, in the vast majority of cases, is done by one or two of the stronger pupils.

The problem of assessment is inevitably bound up with this situation, and many teachers find it difficult to understand why all the children in a group should be given the same assessment category when the work produced is the fruit of the efforts of one or two members. It is inadequate to say that differentiation should take place through the observation by teachers of the involvement of individual members; in large classes with many groups containing between six and nine (or more) members, this is simply unrealistic.

It appears that the common practice of assigning functional roles to group members - 'scriber', 'facilitator', 'helper', 'fetcher', 'reporter' , etc. - significantly contributes to the problem of minimal involvement, the 'scriber' simply does most of the work while the others watch or, more seriously, disengage altogether.

Finally, it remains evident that many pupils naturally tend to work in pairs even when they are seated in groups; quite a number of lessons involved teachers trying very hard to get children who had divided into pairs to instead work as a whole group. This ended whatever productive work was going on in the pairs and substituted a sort of mechanistic turn-taking.

In our view, the use of individual and pair-based work is seriously neglected in all of the schools, both project and control.

**10. Types of Teacher-Initiated Learning Activity**

We have already seen that, in general, verbal methods continued to prevail in the classrooms. In an effort to record in more specific detail the variety, and relative frequency of use, of the teaching activities used in the lesson observations, the evaluation has also recorded the proportions of each type of activity in relation to the total observed for each report. The intention was to reflect the average effective 'weighting' given to each activity by teachers in the classrooms across the whole sample. Consequently, a single peripheral use of an activity was not scored and the emphasis was placed upon recording the activities upon which teachers appeared to place the most reliance. As such the table reflects a numerical summary of qualitative observations recorded both during and after each lesson observation. Despite the inherent subjectivity of the data, we were trying to establish a degree of systematic rigour in the observation of what were likely to be fairly small changes in classroom behaviour.

Little in the data contained in the previous reports suggested that observable changes had occurred in the relative distributions of these types. The overall impression gained through the lesson observations for these reports was that Foundation Phase teachers were presenting better quality lessons through their enriched training and materials, rather than that major changes in types of classroom activities had occurred.

**Table 39: Types of teacher-initiated learning activity by % of total observations.**

| | Project | | | Control | | | Diff. |
|---|---|---|---|---|---|---|---|
| | Baseline | Final | Change | Baseline | Final | Change | |
| Verbal re method | 11.2 | 10.7 | -0.5 | 12.4 | 10.3 | -2.1 | **+1.6** |
| Verbal re content | 14.0 | 9.6 | -4.4 | 15.5 | 15.4 | -0.1 | **-4.3** |
| Chorus response | 12.1 | 8.9 | -3.2 | 16.6 | 10.7 | -5.9 | **+2.7** |
| Read non-fiction | 3.0 | 4.9 | +1.9 | 4.5 | 5.6 | +1.1 | **+0.8** |
| Read fiction | 7.2 | 16.3 | +9.1 | 3.8 | 3.4 | -0.4 | **+9.5** |
| Works from chalkboard | 12.2 | 3.4 | -8.8 | 17.0 | 9.9 | -7.1 | **-1.7** |
| Circulates round groups | 7.5 | 2.3 | -5.2 | 2.3 | 2.6 | +0.3 | **-5.5** |
| Applied activity | 4.0 | 3.9 | -0.1 | 2.6 | 4.7 | +2.1 | **-2.2** |
| Demonstration | 0.2 | 0.3 | +0.1 | 0.0 | 0 | 0.0 | **+0.1** |
| Project work | 0.2 | 0.3 | +0.1 | 0.0 | 0 | 0.0 | **+0.1** |
| Reports on work done | 1.7 | 0.8 | -0.9 | 0.0 | 1.3 | +1.3 | **-2.2** |
| Research a topic | 0.0 | 0.3 | +0.3 | 0.0 | 0 | 0.0 | **+0.3** |
| Worksheets/exercises | 7.4 | 8 | +0.6 | 9.8 | 10.7 | +0.9 | **-0.3** |
| Uses teaching aid | 5.9 | 9.8 | +3.9 | 4.5 | 7.3 | +2.8 | **+1.1** |
| Question and answer | 13.2 | 20.5 | +7.3 | 10.9 | 18 | +7.1 | **+0.2** |

Given the nature of this data, we regard differences over time between project and control (covariance) as potentially significant (worthy of explanation) only when they are at 5% or over (much higher than the 2% of the true quantitative data). The table contains only one such indicator – the time teachers spend reading to children that has increased by **+9.5** in project over control groups. The difference in this area has been evident in all of the study reports and there is no doubt that it is an impact of the project. Interestingly, the difference applies to the reading of fiction books only – because whatever reading teachers do in control schools is dictated by the fact that they only have available the non-fiction subject or learning area 'texts' from which to read to children.

### 11. Types of Questions Asked by Teachers

The baseline study noted that pupil questioning - to discover what pupils know about a topic, to stimulate thought or discussion, or to reinforce what has been learned - is inherent to virtually all forms of education. In particular, the use of participative, active or exploratory/discovery methodologies depend upon the asking of questions well beyond the simple recall and factual questions so common in the chalk-and-talk classrooms of the past. Consequently, the evaluation has recorded the types, and relative frequency of use, of the questions that teachers typically use during the lesson observations. Again, the figures refer to the proportion of each type to the total rather than to the percentage of lessons; the peripheral use of an activity was not scored but the emphasis was again placed upon recording the types of questions upon which teachers appeared to place the most reliance.

Nothing in the data contained in the previous reports suggested that substantive changes had occurred in terms of the types of questions asked by teachers. Indeed, factual and recall questioning remained the mainstay of teacher technique in both project and control classrooms. The only significant change noted in the third report was an increase in the level of comprehension questioning in the project schools; clearly an artefact of the project-supplied materials. Teachers could ask questions about the meaning of texts they have read in class primarily because they have the book supply to enable them to do so.

**Table 40: Types of questions asked by teachers by % of total observations**

| | Project | | | Control | | | Difference |
|---|---|---|---|---|---|---|---|
| | Baseline | Final | Change | Baseline | Final | Change | |
| Opinion | 3.5 | 4.9 | +1.4 | 3.5 | 7.7 | +4.2 | **-2.8** |
| Judgement | 2.2 | 2.0 | -0.2 | 2.2 | 2.4 | +0.2 | **-0.4** |
| Predict | 0.9 | 2.9 | +2.0 | 0.9 | 0.6 | -0.3 | **+2.3** |
| Explain | 6.7 | 7.2 | +0.5 | 6.7 | 10.7 | +4.0 | **-3.5** |
| Summarize | 1.3 | 0.3 | -1.0 | 1.3 | 0.0 | -1.3 | **+0.3** |
| Factual | 18.1 | 8.8 | -9.3 | 18.1 | 11.9 | -6.2 | **-3.1** |
| Rote | 21.9 | 20.3 | -1.6 | 21.9 | 20.2 | -1.7 | **+0.1** |
| Comprehension | 11.1 | 16.4 | +5.3 | 11.1 | 11.2 | +0.1 | **+5.2** |
| Sequence | 0.6 | 2.0 | +1.4 | 0.6 | 0.6 | 0.0 | **+1.4** |
| Recall | 8.9 | 8.8 | -0.1 | 8.9 | 7.1 | -1.8 | **+1.7** |
| Calculate | 9.2 | 2.0 | -7.2 | 9.2 | 1.8 | -7.4 | **+0.2** |
| Observe | 5.7 | 10.1 | +4.4 | 5.7 | 5.2 | -0.5 | **+4.9** |
| Identify | 9.8 | 13.4 | +3.6 | 9.8 | 9.5 | -0.3 | **+3.9** |
| Describe* | | 7.2 | 7.2 | | 4.8 | +4.8 | **+2.4** |
| Classify* | | 2.0 | 2.0 | | 0.6 | +0.6 | **+1.4** |
| Estimate* | | 0.7 | 0.7 | | 0.0 | 0.0 | **+0.7** |

**** Note:** These items were added in 2003 – therefore there are no baselines.*

In the final round of observations, only one item, 'comprehension', again exceeds the 5% significance mark – for the same reason. However, it could also be noted that, although the changes are all small, the fact that twelve of the sixteen (75%) of them are positive suggests that teachers in project classrooms are beginning to ask a wider range of questions. What remains disappointing is that teachers seem to take a long time to use the materials to enable them to move on to these more sophisticated kinds of questions, particularly summarizing, sequencing, judging, predicting, estimating, etc. It will be noted that factual, rote, comprehension and identification questioning still constitute well over half of all the questions asked in both project and control classrooms.

## 12. Questions Asked by Pupils

The baseline report commented that it is a source of continuing interest in South African education that so few questions are asked by pupils of their teachers. When questions are asked, they seldom deal with the content of the lesson, but with the methods used or for permission to do something (*must we write the answers or should we just discuss them?*, for example). Even though teachers frequently ask *'are we together?'* or *'does everyone understand?'*, the answer is invariably a whole class chanting *'yes'* in unison.

Previous reports noted no significant differences in this regard in project or control classrooms. One might assume that this was because pupils were thoroughly at ease with the content of the lesson, however lesson observations and test results indicated that this was not the case. The reports commented that it was more likely that the situation reflected a general continuing passivity of pupils in many South African schools that has not yet been significantly affected by reform or by the LFL. Greater activity, in other words, tended to be more physical than mental.

The baseline conditions have remained virtually constant into the final year of the study and there is no indication of any project impact on this indicator. The types of questions which pupils do ask are still primarily about the method upon which the lesson is based, or for permission to do something;

in both samples well under half of the questions asked by pupils were based upon the content with which the lesson was dealing.

**Table 41: Proportion of lessons in which pupils asked questions**

| | % of lessons |
|---|---|
| **Project** | |
| Baseline | 12.9 |
| Final | 17.2 |
| Change | +4.3 |
| **Control** | |
| Baseline | 8.8 |
| Final | 12.1 |
| Change | +3.3 |
| **Difference** | **+1.1** |

### 13. Estimated Quality of Materials

The fourth report presented an item dealing with the estimated effectiveness of project materials in classroom use which concluded that the materials are generally both high quality and high interest. The children manifestly enjoyed the materials and their use, and the teachers have no other similar materials in similar quantities. The appropriateness of the language level is related to the different areas and types of schools; what might be too difficult in remote rural areas of Limpopo may be too simple for established urban townships on the Reef. This is true of the pupils' understanding of the content and methods of the materials as well.

The report commented that the variety of schools on a national scale is an enduring feature of our education system and one with which all educational agencies, from the DoE to NGOs, have to deal. The LFL, in its attempt to do so, has proved remarkably successful in selecting materials that are universally welcomed and used in all of the different local and regional contexts in which research takes place, even if to varying degrees of both teacher and pupil comprehension. These conclusions remained true into research for the final report and the figures reflected in the fourth report are reproduced below. In this table 5 equals excellent and 1 equals very poor.

| **Table 42: Estimated effectiveness of LFL materials observed in use (2004)** | **5** | **4** | **3** | **2** | **1** | **Mean [5]** |
|---|---|---|---|---|---|---|
| Is the appearance of the material/s appealing to learners? | **11.8%** | **73.5%** | **11.8%** | **2.9%** | **0.0** | **4.0** |
| Is the language level appropriate for the class? | **1.5%** | **34.2%** | **47.8%** | **16.4%** | **0.0** | **3.2** |
| How well did the learners understand the content of the material/s? | **4.5%** | **37.9%** | **40.9%** | **16.7%** | **0.0** | **3.2** |
| How well did the learners understand the methodology of the material/s? | **6.0%** | **32.8%** | **52.2%** | **6.0%** | **1.5%** | **3.3** |
| Did the materials provide the basis for an effective lesson at this level? | **1.5%** | **66.1%** | **18.5%** | **12.3%** | **1.5%** | **3.5** |
| Was the teacher able to use the materials well enough to present an effective lesson? | **1.5%** | **22.7%** | **45.4%** | **25.8%** | **4.5%** | **2.9** |

## 14. Quality of Lesson Preparation and Presentation

The effectiveness and appropriateness of the materials having been established, research for the final report substituted a new item. It must be immediately admitted that the table is based on the subjective judgements of the observers but it was felt that it was necessary to provide some kind of general reflection of the relative quality of lessons presented by project and control teachers. In this table, each item was scored from 5 to 1 where 5 equals excellent and 1 equals very poor. The mean score is reflected for each item; 3, therefore, equals 'average'.

| **Table 43: Estimated quality of lesson preparation & presentation (2005)** | **Project** | **Control** | **Diff.** |
|---|---|---|---|
| The lesson has a specific objective | 3.5 | 3.3 | +0.2 |
| The teacher is fully prepared? | 3.4 | 3.1 | +0.3 |
| The lesson has a logical structure & process | 3.2 | 3.0 | +0.2 |
| The lesson methodology is appropriate to the topic | 3.0 | 2.9 | +0.1 |
| Pupils' level of understanding of content | 2.8 | 2.8 | 0.0 |
| Teacher assessment of pupils' understanding of lesson content | 2.9 | 2.7 | +0.2 |
| Learner involvement | 3.5 | 3.1 | +0.4 |
| Lesson conclusion | 2.9 | 2.7 | +2.0 |

Seven of the eight indicators reflect a positive, if small, difference in favour of the project schools and there is little doubt that the project methods and, especially, materials are generally supporting the presentation of higher quality lessons.

However, it should be said that most of the lessons observed in project classrooms were direct replications of a LFL lesson, usually with the help of a step-by-step teacher guide. There is no problem with this during the period of training as newly learned skills are practised but, by project end, we would have expected more teachers to be able to present lessons developed by themselves, using LFL-derived skills and materials, rather than still be replicating lessons with which they had been provided. As a consequence, we often saw the same lessons presented at different grade levels at different schools all over the country.

# APPENDIX ONE

## Extract from the Baseline Report

## Objectives and Operational Design of the READ/Business Trust Project

### 1. Background

The READ Organization commenced the 'Learning for Living' Project in 2000 as a result of a successful submission to the Business Trust (BT). That submission proposed the establishment of a five-year programme, in co-operation with the National and Provincial Departments of Education (DOE) to *improve the communication, language, cognitive and literacy skills of learners from Grade 1 to Grade 7* in 1 000 schools (this has since become 967 schools) in all nine of the provinces.

The fundamental motivation of the project was the need to improve the quality of the outcomes of South African schools. While much has been achieved in terms of access to schooling since the desegregation of the education system, much remains to be done in terms of quality of outcomes. The Inaugural Report of the BT includes figures from the Third International Maths and Science Study (TIMMS) to illustrate that South African children are seriously under-performing in relation to international standards; recent figures in the press have confirmed the perception. The same BT report bluntly refers to the 'crisis in schooling'.

There is little doubt that one of the most serious causes of the endemic under-performance of children in our schools has been the failure to teach/learn language and literacy adequately, especially when one considers that the vast majority of children are educated in English, a second, third or fourth language for both teachers and pupils.

This evaluation was commissioned to provide the systematic external assessment required by the BT. This is the first report of the evaluation and it is primary concerned with the systematic establishment of the baseline quantitative and qualitative conditions prevailing in the schools prior to the inception of the project. It is against these baselines that project impact will be measured and observed in future reports.

### 2. Description of the READ Programme

#### 2.1. Objectives

In their proposal to the BT, READ provided a general statement of the objectives of the project.

*"The project will aim to provide a measurable, sustainable and cost-effective programme in the learning area of Language, Literacy and Communication…This project, while focusing on the acquisition of reading, writing and cognitive skills will also facilitate effective integration of all the learning areas, to ensure adoption of cross-curricular, skills-based teaching and learning. This will in turn impact on learners' outputs by increasing abilities to demonstrate outcomes (knowledge, skills, values and attitudes), and on teachers' abilities to deliver the new methodology more effectively."*

## 2.2. Programme

The programme is applied through a combination of INSET for teachers and principals, materials supply and classroom-level monitoring. The INSET courses give guidance in the use of the classroom materials which, in turn, provide the opportunity for the guided practice of the concepts introduced during the courses. Classroom monitoring aims at providing feedback firstly to teachers in their application of the materials and concepts, and secondly to the organization itself as formative information about the developing impact of the project.

### 2.2.1. Training and Materials Supply

*Classroom Teachers*

In all, Foundation Phase teachers will receive 11 INSET courses and Intermediate Phase teachers will receive 13. The table below, summarizing the schedule for the delivery of these courses and their associated materials, is based on the latest schedule available from READ (December, 2000). Where this schedule explicitly specified whether material was for Foundation or Intermediate Phase, that information has been recorded in the table. One assumes the rest can be deduced from the Level designations of the materials; e.g. that Level 1, 2 & 3 refers to material for the Foundation Phase, and that Level 4, 5, 6 and 7 to the Intermediate Phase.

**Table One: Summary of Delivery Schedule for Teacher INSET and Materials Supply**

| | **INSET** | **For** | **Classroom Materials** | **For** |
|---|---|---|---|---|
| **Year 1** | | | | |
| 1-3//00 | Introducing a literacy programme | F | Book education pack | F |
| 4-6/00 | Using stories for language development | F | Storykit 2005: Level 1, 2 & 3, Stage A & B | F |
| 8-9/00 | Shared reading and writing | F | | |
| **Year 2** | | | | |
| 9-12/00 | Guided reading | F | Sunshine extension pack: Pack A & B | F |
| 1-3/01 | Introducing a literacy programme | F/I | Book education pack | I |
| | Using stories for language development | I | Storykit 4, 5, 6 & 7 | I |
| | Introduction to learning to read/home Language literacy | F | Home language starter pack and story kit<br>Core books: Level 1 to 7<br>Learners books to complement Sunshine | F<br>F/I<br>F |
| 4-6/01 | Shared reading and writing | I | Accelerated pack: Selection A: Level 1, 2 & 3<br>Selection B: Level 1, 2 & 3 | |
| | Group and guided reading and using group readers | I | Accelerated learners books to complement above<br>Group readers: Stage 4, 5, 6 & 7 | I |
| | Group reading and using group readers | F | Learning to read and write | I |
| 7-9/01 | Learning to read and write | F/I | Storykit little books | F/I |
| **Year 3** | | | | |
| 1-3/02 | Using a classroom resource collection and independent reading | I | Classroom resource collection: Level 4, 5, 6 & 7<br>Core books: Level 4, 5, 6 & 7 | |
| 4-6/02 | Using a story for language development and for language across the curriculum | F/I | Storykit little books: Stage 1 to 6 | |
| 7-9/02 | Using a non-fiction book for lesson planning | I | Depending on budget: Sunshine Science books: level 4, 5, 6 & 7 | |
| **Year 4** | | | | |
| 1-3/03 | Reading for information skills: reading and writing across the curriculum | I | Depending on budget: Sunshine Science books: level 4, 5, 6 & 7 | |
| | Classroom management | I | | |
| 4-6/03 | Teaching phonics and spelling | F/I | | |
| **Year 5** | | | | |
| 1-3/04 | Reading and writing for real purposes | I | | |
| 4-6/04 | Choosing and using books/other material | F/I | | |

Most of the courses last from 1 to 4 days, many of them are 3 days. The schedule represents a sustained programme of training and the provision of a large number of materials to schools. This report notes elsewhere that shortages of material are common in the schools participating in the project and there is no doubt that the materials alone will amount to a significant change in their classrooms. The scope and sustained nature of the INSET will also be new to teachers and, alone, amounts to a significant change in their professional experience.

*Principals*

Alongside the teacher INSET, training is also provided for principals and senior school staff. READ has learned by experience that change in schools is a systemic process and that it is essential to engage them on levels beyond the classroom alone. Teachers need support in planning and implementing learning programmes based on new methods and materials and school managers need to know how to provide this support. Indeed, effective school management and leadership is now generally recognized to be an essential component of the effort to transform schools into viable institutions of learning. The change from regarding principals as a combination of the DoE's 'postman and policeman' to entrusting them with the managerial, financial and professional leadership of schools requires skills which were not required when many of them were promoted, and which many of them are taking time acquiring.

**Table Two: Summary of Principals Training** (Continued overleaf)

| | |
|---|---|
| **Year 1** | |
| 4-6/00 | Management skills for school leaders |
| 7-9/00 | Governance skills for school leaders |
| **Year 2** | |
| 10-12/00 | Governance skills for school leaders – continued |
| 1-3/01 | Financial management and budgeting |
| | School manager's leadership, monitoring and mentoring in the BT Project |
| 4-6/01 | The seven habits of highly effective people |
| 7-9/01 | Strategic planning for school leaders |
| **Year 3** | |
| 10-12/01 | Planning the school year |
| 1-3/02 | Principle-centred leadership |
| 4-6/02 | Communication skills and conflict management |
| 7-9/02 | Marketing your school |
| **Year 4** | |
| 1-3/03 | Parental education |
| 4-6/03 | Fundraising for the school: part one |
| 7-9/03 | Fundraising for the school: part two |
| **Year 5** | |
| 10-12/03 | Planning the last year of the BT Project |
| 1-3/04 | Strategies for building a sustainable language programme |
| 4-6/04 | Depends upon needs expressed by principals: many courses available |

*Leader Teachers*

The most enthusiastic and able teachers in schools will be identified and provided with extra training as leader teachers. The fundamental role of these teachers is to provide both programme and motivational support to the rest of the staff in between visits from READ staff. They are also, importantly, available to orient teachers who are newly appointed to the school. The intention is to

train them a little in advance of the rest of the teachers so that they are able to provide support immediately after general INSET courses are completed.

**Table Three: Summary of Leader Teachers Training**

| **Year 2** | | **For** |
|---|---|---|
| 4-6/01 | Shared reading and writing | F/I |
| | Group and guided reading | F/I |
| **Year 3** | | |
| 10-12/01 | Using a classroom resource collection for teaching and learning, and independent reading | F/I |
| 1-3/02 | Teaching writing | F/I |
| **Year 4** | | |
| 1-3/03 | Teaching information skills across the curriculum | F/I |
| 7-9/03 | Reading and writing for real purposes | I |
| **Year 5** | | |
| 1-3/04 | Choosing resources for learning and teaching in the classroom | F/I |

**Table Four: Summary of Numbers of Courses by Participants by Year**

| | FP teachers | IP teachers | Principals | LeaderTeachers |
|---|---|---|---|---|
| Year 1 | 3 | 0 | 2 | 0 |
| Year 2 | 5 | 5 | 5 | 2 |
| Year 3 | 1 | 3 | 4 | 2 |
| Year 4 | 1 | 3 | 3 | 2 |
| Year 5 | 1 | 2 | 3 | 1 |
| Total | 11 | 13 | 17 | 7 |

### 2.2.2. Classroom-level Monitoring

Classroom-level monitoring is intended to provide direct feedback to teachers (and to the organization itself) about the classroom-level application of the READ approach whether in terms of the INSET, or of the books/materials teachers receive. The internal monitoring system (*See* Quality Control, below) is based upon checklists of expected outcomes which are provided to teachers at each course – they are used to structure observation and feedback during READ classroom visits, and to provide opportunities for self and peer-assessment between them.

Though it is possible that READ staff may visit teachers in their classrooms more than once if they are having problems with the INSET/material, or for other incidental reasons, the delivery design commits the project to at least one visit for each course he/she attends. Extrapolating from the above tables, this implies the following minimum distribution of classroom visits per teacher.

**Table Five: Summary of Classroom Visits by Year and by Phase**

| | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Total |
|---|---|---|---|---|---|---|
| Foundation Phase | 3 | 5 | 1 | 1 | 0 | 11 |
| Intermediate Phase | 0 | 5 | 3 | 3 | 2 | 13 |

The leader teachers also provide support and guidance to other teachers between READ visits, and are given extra training to enable them to do so (above).

### 2.2.3. Quality Control

In a programme as extensive and sustained as that outlined above, the issue of quality control becomes essential if the whole system is to achieve consistent impact upon school, teacher and, ultimately, pupil performance. The project proposes to deal with the issue through two processes:

- Two internal training programmes running ahead of the school-level programme have been scheduled, one for mentors and the other for the trainers they are to support. These programmes are intended to ensure the teacher-trainers are themselves adequately trained and supported by senior staff.

- READ's existing internal monitoring system, based on the use of checklists for monitoring visits and self/peer assessment, has been extensively re-designed and re-developed with the assistance of international academic experts in the field. It aims to provide formative feedback to the organization about the early and developing impact of the progressive implementation of the delivery plan upon teachers and pupils. While data from the external evaluation is available to the internal system, and *vice versa*, it is important to realize that the two are separately designed, implemented and reported.

Finally, READ Head Office provides a centralized source of expertise and support to the regional staff. Essentially the organization aims to avoid the creation and expense of senior specialist and support units in each region by using the mobile resources of its Head Office staff, all of whom are highly skilled and experienced.

# APPENDIX TWO

## Extract from the Baseline Report

## Design of the Evaluation of the READ/Business Trust Project

### 1. Purpose of the Baseline Study

It must be understood that the baseline study is intended to be the reporting and systematic recording of data about schools, teachers and pupils reflecting the conditions prevailing before the inception of the project, and which will be used in later reports to begin the process of evaluating the achievement of impact, or otherwise, of the READ/BT project. It is *not* intended to be a general national survey of conditions except insofar as the schools which have been selected for the project are considered to be representative of *all* schools in South Africa.

Even less is it intended to be a comparative study of conditions prevailing in the different provinces. There are two main reasons why provincial and grade level scores cannot be used for comparative purposes on any systematic level.

- The testing/school visits took place at different times of the year in different provinces. The first visits took place in March in KZN and the last in August in the Northern Cape, a difference of six months extra schooling for the Northern Cape. Since the indicators of impact use the analysis of covariance of mean gains, whether at grade or cohort level, and *not* the comparison of absolute scores, this has no relevance to the study; but it does mean that direct comparisons of provincial scores are likely to be misleading.

- The provincial sub-samples are not intended to be representative of each province (*See* discussion of the sample, below). Essentially, the mix of urban and rural schools in each province was decided by the needs of national sample.

### 2. Evaluation Questions

The design of the evaluation uses a mixed model applying both quantitative and qualitative methods. It is based upon four generic questions:

- Has the project been delivered as designed?
- Has the project achieved predicted qualitative impact *(process means)* on a consistent basis in the project schools?
- Have the predicted quantitative impacts *(product ends)* occurred in the project schools?
- Can these process and product impacts be ascribed to the project?

The first question is answered through a comparison of the actual versus planned delivery of the project (*See* tables in Section I) using formal project reports, and through principal and teacher interviews. The last question is answered primarily through the application of a quasi-experimental design using covariance as its analytical tool, and also through an examination of the pre-project equivalence of project and control samples. The remaining two questions are addressed by the rest of the design.

## 3. Quantitative Research

This research is directed to the third question, that of product impact or what might be called the external validity of the READ model, i.e. does it result in its predicted product outputs? A quasi-experimental design, using pre-, mid- and post-project tests for project and control groups has been applied to the question of product impact. The testing tracks 2 four-year cohorts, starting in 2000 at grades 3 and 4, and 1 three-year cohort, starting at grade 5. Grade 7 is also tested each year to allow comparison of changes in 'school standards' at this level over the life of the project.

The project has been extended into a fifth year since the design was approved; discussions with READ have commenced about the inclusion of this year in the evaluation, but the rest of this report deals with the originally approved design.

### 3.1. Indicators

The indicators used for examining pupil impact are based on the concept that the single strategic purpose of any educational project or intervention is to improve the quality of the outcomes of schooling; to improve, in other words, the performance of the pupils for whose sake the system exists. Indeed, it is very difficult to think of any other sensible strategic reason why any educational project or intervention should be established, even if the tactical routes to this goal vary widely in both focus and method.

There are two specific indicators:

- There were will be a significantly higher gain between pre and post testing on a battery of literacy and numeracy instruments in project over control cohorts. (Analysis of covariance)
- The repeat rate in project schools will be significantly lower than that in control schools.

It is logically assumed that any increase in language performance should be reflected in a corresponding general improvement in performance in all subjects taught in that language and that, as a necessary consequence, overall pupil repeat rates should decline. Numeracy testing examines this assumption more directly; it is generally acknowledged that literacy and numeracy are the foundation of all formal education, and it is most unlikely that any general improvement in pupil performance would occur without it being reflected in either. The logical assumption, again, is that if a language programme results in improved performance in numeracy it is safe to assume that improvements are likely to occur in other subjects as well.

### 3.2. Instruments

A battery of three instruments is used at grade 3 level while four instruments are used at grade4 to 7 level. (Table overleaf)

**Table Six: Instruments Used for Impact Testing**

| | | |
|---|---|---|
| Grade 3 & 4 | Word Recognition | Sentence Completion: 12 items |
| | Reading | Modified Cloze*****: 15 items |
| | Free Writing****** | Visual Stimulus: 18 items |
| Grade 5 & 7 | Literacy | Modified Cloze*: 68 items |
| Grade 5 & 7 | Free Writing****** | Written Stimulus: 30 items |
| Grade 4 to 7 | Numeracy******* | 48 Items |

***As far as I am aware, the outcome of the long debate about cloze and literacy testing in the literature is that cloze is acknowledged to be at least as reliable as any of its competitors, singly or in combination. It is based on meta-cognitive and literacy skills in that respondents are required to 'decode' text sufficiently well to be able to comprehend and complete its intended meaning.*

****The free-writing test items are scored 0 (irrelevant to stimulus, no recognizable words) to 2 (relevant to stimulus, no errors) with a further possible point to 3; where the writing is relevant or extends the stimulus and the language has a degree of sophistication or originality that makes it stand out even if it may have an error.*

*****The numeracy test is a combination of straightforward arithmetic operations with visual and numeric fractions, shape recognition, symbolic logic, conversions, sequences, combined literacy/numeracy word sums and table interpretation - all fundamental to the concept of numeracy*.

All of the grade 3 & 4 literacy instruments, and the grade 5 & 7 writing instrument, were designed by Professor W. Elley of the University of Canterbury in New Zealand and piloted by ***ESA***. The grade 5 & 7 literacy and numeracy instruments were designed and piloted by ***ESA.***

It is important to note that the instruments are intended to yield data about relative change which will be measured through the analysis of covariance. Covariance measures the differential gain between pre and post testing of the project and control groups, considering the degree of higher gain achieved by the project sample, should such occur, as the measure of project impact. Put simply, since both project and control groups receive 'normal schooling' over the life of the programme, and since cohort mean scores in both groups will increase as a result, the subtraction of the gain measured for the control group from that of the project group eliminates the effect of normal schooling and maturation; we are left with the measure of the gain achieved by the programme.

The instruments are, therefore, *not* intended to measure absolute performance against national normative standards. (Which have not yet, in any case, been reliably established.) For one thing, such instruments will probably be designed in terms of the content of the prevailing curriculum whereas we are still applying the 'old' curriculum in some grades, C2005 in others and awaiting immanent changes in C2005 at all levels at the same time. The instruments used in this evaluation are, therefore, based upon the generic skills one would expect to find in *any* curriculum.

There are three primary requirements of instruments when used for the analysis of covariance over a considerable period:

- They should be capable of consistent discrimination between weaker and stronger respondents.
- They should yield an appropriate spread of scores across the grades for which they are intended, as well as within each such grade level.
- Baseline tests should yield mean scores low enough to allow observation of the distinction between project-caused change and the expected gain caused in both project and control that will, in any case, occur through growing maturity and the impact of routine schooling in the 3 cohorts. This avoids what is called the 'ceiling effect'.

Piloting, and the extensive prior use of these and very similar instruments for the same purpose in a variety of other projects, has consistently yielded data confirming that the instruments are reliable.

Indeed, they have also yielded data about relative learner performance for individuals very similar to that obtained by a far more extensive and intensive internal school system in another project.

The instruments, taken as a whole, do focus on a circumscribed set of skills and knowledge, but that is an inherent characteristic of all tests. The battery is intended to allow the investigation to focus on the most 'fertile' parameters that can logically be taken to indicate the existence or otherwise of genuine growth in performance in literacy and numeracy. The tests use core indicators - if children are not able to deal adequately with them it is very difficult to see how they could possibly be achieving high levels of literacy and numeracy.

The particular means and methods which READ proposes to achieve the universal improvement of literacy and numeracy in our schools are not at question in the testing of whether there has, or has not, been any such improvement. The instruments simply provide reliable answers to the question of whether such improvement has occurred, or not. The process question of exactly how improvements are achieved is the province of the rest of the evaluation design.

### 3.3. Sample

*Research Sites*

The project is national in scope and consequently the whole national sample, as against the provincial components of that sample, is the key focus of analysis in the interpretation of the data. In this sense, the pupil sample theoretically could have been chosen from the class lists of the 967 schools participating in the READ/BT project by using a table of random numbers.

The national education context is, however, clearly differentiated in terms of provincial interpretation and application of national policy, historical development, and socio-economic and educational conditions. Consequently, to arrive at a nationally representative sample it was necessary to select the pupils from a stratified sample of schools that reflected the variety inherent within the national context.

The most obvious staring point for the stratification was the 9 provinces. Gauteng is physically the smallest and the most densely populated, urbanized and industrialized of them; there are, consequently, fewer rural and farm schools, and many more resources historically available in a more developed socio-economic context than in the more rural provinces. The Eastern Cape, Mpumalanga and Northern Provinces, on the other hand, incorporate former ‘homelands’, accommodate most of their pupils in rural schools, many of them in remote areas, and have a much lower level of historical resource availability and socio-economic development. Further, the provincial contexts are themselves differentiated in similar terms and it was decided to select schools from at least two and sometimes three locations within each province to allow for some of these internal differentiations. Lastly, the project is targeted primarily on rural schools; the pupil sample needed, therefore, to be drawn primarily from rural schools though major urban areas are also represented.

The pupil sample was eventually drawn from 90 schools (54 project and 36 control) with 10 schools (6 project and 4 control) in each province. In all, research is taking place in 22 separate sites across the country.

*Selection of Project and Control Schools*

The project and control schools for the evaluation were jointly chosen by the respective Departments of Education and the local READ offices as being representative of the bulk of the

schools selected for the Business Trust Project in each province. The main criteria supplied by ***ESA*** were that project and control schools should come from similar socio-economic environments, be as physically close to each other as possible and should not be included in any other current major school development projects.

*Selection of Pupil Sample*

To arrive at a sample size, we first estimated the target pupil population in 1 000 schools conservatively at around 600 000. (READ now estimates that there are around one million children in the 967 participating schools.) The vast majority of project schools offer grades 1 to 7, while fewer are junior primaries offering grades 1 to 4 (most of the latter are in remote rural areas). Consequently we assumed, again conservatively, that there would be an average of five grades offered at each school. Secondly, many schools have three classes at each grade level while fewer have either two or four, and even less one or five - we assumed three classes at each level. Finally, much recent research suggests that class sizes are averaging around 40 in primary schools, though there are still schools with higher, sometimes much higher, numbers, as well as those with lower, sometimes much lower, numbers; we assumed 40 per class. Consequently, 1 000 schools x 5 grades x 3 classes per grade x 40 pupils per class = 600 000.

The programme encompasses both junior primary (Foundation Phase) and senior primary (Intermediate Phase) levels and pupils will pass from one level to the next within the life of the programme, and some will leave primary school altogether. If we are to track cohorts of pupils over the period of project implementation, the sample needed to encompass both levels. Further, though the primary focus of data collection is on the cohorts, the fact that some pupils would complete primary schooling suggested that it would be useful to track annual changes in the 'exit' levels at the schools as well. These considerations arrived at the following testing process.

**Table Seven: Testing Process for Cohorts**

| | **2000** | **2001** | **2002** | **2003*** |
|---|---|---|---|---|
| Cohort I | 3 | 4 | 5 | 6 |
| Cohort II | 4 | 5 | 6 | *7* |
| Cohort III | 5 | 6 | *7* | |
| *Exit* | *7* | *7* | | |

***The READ/BT project has subsequently been extended to 2004 and, consequently, post tests will take place in that year.*

Given an expected average of 120 pupils at each grade level at each school (see above) we opted to select one-sixth of that number at each level (20), allowing for the choice of a reasonable number of pupils from each class. The eventual result was that the sample was to consist of 7 200 pupils (90 schools x 4 grades x 20 pupils).

According to the literature (*Krejcie and Morgan, 1970*), a sample of this size for a target population of 600 000 yields data with a 95% confidence level at a precision close to 1%, more than adequate for the purposes of the study. Even if we consider only the project group of 4 320 within the overall sample in relation to the 600 000 population of the project schools, the data remains at a confidence level of 95% while the precision is closer to 2%, again more than adequate for our purposes.

Initial pupil selection was carried out by the field researchers on the day testing first took place at each school. Since we are following cohorts the process does not need to be repeated. The following process was followed:

- The number of pupils required at each grade level (20) was divided by the number of classes at each level. Thus, if there were 4 classes at a grade, 5 pupils would be selected from each of them.
- The number of pupils in any particular class was divided by the number of pupils to be selected and the result used to select individual pupils. Thus, to continue with the example, if there were 40 pupils in a particular class every $8^{th}$ pupil ($40 \div 5 = 8$) was selected.

The method encourages selection across the full range of ability and peer groupings within each grade level and class.

*Validity of Provincial Sub-samples*

The sample is intended to be representative of the READ/BT schools on a national level and it is much less likely that the provincial sub-samples are equally representative of each province. To make them so would have required very much bigger provincial sub-samples than were required for national validity - this was not required by the brief nor possible in terms of the cost parameters.

However, provincial mean scores for the pupil impact tests and for many of the more qualitative/contextual data items have been provided for interest, and to give some idea of the provincial differences within the whole sample. The pupil impact scores retain the project vs. control distinction but, to make the rest of the items more meaningful, means have been calculated from the combined scores of all 10 project and control schools – to limit the effect of one or two schools on the mean.

Also note that the overall urban/rural spread is nationally representative of the target group but not always at provincial level; no urban areas, for example, are included in the KZN, Mpumalanga or Northern Province sub-samples and no rural areas in the Gauteng sample. On the other hand, the other five provincial sub-samples all contain a mix of urban and rural areas. This is because the bulk of the target schools nationally are rural and the national sample, therefore, needed to have a majority of rural schools. The urban areas that have been chosen are very likely to be representative of the whole country.

*All in all, the focus of systematic analysis is the whole national sample and care must be exercised when comparing different provinces – in many cases they are not comparable on a systematic level.*

**4. Qualitative Research**

A wide range of data is collected within the qualitative design and systematically organized on the data base to reflect the contextual and process differences between project and control schools and classrooms.

This research is primarily focussed upon the second of the evaluation questions dealing with the effectiveness of the qualitative process changes (methodological, behavioural, attitudinal etc.) achieved in the schools by the READ/BT project. It is also concerned with the understanding of the socio-economic and contextual environments of project and control schools, and yields data which can confirm the measurement of project/control equivalence made possible by the quantitative research.

READ proposes to achieve impact upon pupils through a mediated process, it does not, in other words, work directly with pupils. The project aims rather at improving the quality of pupil

outcomes through enhancing the quality of the education which they receive through the supply of books and other teaching/learning material, and the training of teachers to use them effectively. Consequently, the second generic question upon which the design is based; *Has the project achieved predicted qualitative impact (process means) on a consistent basis in the schools?* becomes particularly important.

To achieve *consistent* impact upon pupil performance implies the need to achieve consistent qualitative impact upon the means through which the achievement of such pupil impact is proposed. Inconsistent qualitative impact, varying widely in quality and completeness, is likely to result in equally inconsistent and variable pupil impact.

Qualitative research is organized through the same quasi-experimental design and is carried out in the same 90 schools as the quantitative research, running parallel to it throughout the evaluation.

### 4.1. Indicators and Variables

The indicators used to develop the qualitative interview and observation schedules are a combination of biographical and contextual items with the observation of a number of standard behavioural and methodological factors in classrooms. Some of these provide indicators for observing the achievement of qualitative change (e.g. types of teaching and learning activities) and some provide information on variables that may affect project outcomes (e.g. socio-economic environment of schools). The list below simply nominates these items – they are all discussed at greater length in the following sections of the report.

- teacher/pupil ratio
- class sizes
- change in enrolment
- surplus/deficit of classrooms
- school facilities
- pupils living with pensioners
- level of community unemployment
- level of community illiteracy
- living conditions of pupils
- payment of school fees and school funds
- supply of learning materials from the doe
- INSET and materials received from NGO's
- attitudes to other participants (teachers, pupils & parents)
- years of experience (teachers)
- qualifications of teachers
- availability of 'core' learning and teaching materials in classrooms
- displays of pupils' work
- teacher-made materials
- lesson planning and process
- monitoring and assessment of pupils' work
- lesson preparation
- modes of pupil activity (i.e. % of pupil time spent listening, speaking, writing, reading, doing)
- focus of classroom activity (i.e. group, pair, individual)
- types of teaching and learning activities (e.g. applied activity, use of worksheets, discussion)
- use of questions by teachers (e.g. observe, estimate, summarize etc.)

- use of questions by pupils (e.g. content, method, permission etc.)

In addition, the READ delivery planning schedules provide detailed outcome statements for all courses. These statements constitute a second set of indicators for lesson observations in the second and later reports of the evaluation allowing us to understand the degree and nature of classroom level impacts achieved by the project in relation to what it has predicted at each point of its implementation.

Qualitative data is collected through:

- Formal READ reports
- Principal and teacher interviews using a combination of structured schedules and open-ended questioning.
- Lesson observations using a combination of structured schedules, unstructured observation and verbatim/situational recording.

A variety of triangulation techniques are used. Multiple researchers, and the use of a variety of data collection methods, enhance the reliability of qualitative data.

**6. Location of Schools**

As some of the more remote 'deep rural' areas may be unfamiliar to the general reader, the table below lists the 22 research sites in terms of the nearest town; the Tzaneen schools in the Northern Province, for example, are actually in the Bolebedu District to the North of the town. The list does, however, portray the variety of the 22 research sites.

| | | | |
|---|---|---|---|
| Northern Province | Tzaneen | KwaZulu/Natal | Bergville |
| | Soutpansberg | | iXopo |
| Mpumalanga | Badplaas | Eastern Cape | Motherwell |
| | Malelane | | Queenstown |
| Gauteng | Soweto | | Peddie |
| | Vosloorus | Western Cape | Piketberg |
| North West | Potchefstroom | | Worcester |
| | Orkney | | Khayelitsha/Athlone |
| | Mabopane | Northern Cape | Kimberley |
| Free State | Bloemfontein | | Koopmansfontein |
| | Welkom | | Postmasberg |

## APPENDIX THREE

### The Evaluation Of Educational Development Projects

Since early 1994, Eric Schollar and Associates ***(ESA)*** has been commissioned to carry out a number of evaluation studies of a wide range of development projects and other educational interventions in schools across South Africa.

These interventions have all developed particular core strategies to deal with the central problems they have identified; some focus, for instance, on the broader departmental or school-level issues of governance and management, while others have a narrower focus upon the classroom and teacher-level. Common to all of them, however, is a proposition that their strategic approach will prove effective in solving educational problems and an implication, even if implicit, that this strategic approach could or should be applied on a larger provincial or national scale. In these terms, the central task of evaluation research is to distinguish reliably between effective and ineffective strategic approaches. This provides policy makers and funding agencies alike with functional information about the most effective means of achieving educational change.

When designing an evaluation, ***ESA*** considers intervention projects to be experimental applications of particular strategic approaches. These strategies are operationalized through a process of increasing 'concreteness' and detail through the linked stages of tactical description and organizational design; to result in a delivery system that attempts to realize and deliver the practical implementation of the theoretical approach.

We can summarise the task of evaluation research in these terms in two questions:

- Does the strategy of an intervention result in the predicted change? This is the central question for evaluation as an applied Social Science - policy-makers are faced with competing strategies for achieving change, and they need to know which strategic approaches are likely to be most effective when employed on large scales, especially when resources for development are scarce.
- Does the delivery system embody the strategic approach of the intervention, and is it sufficiently coherent to achieve change? Both interventions and policy-makers need to know what are the critical management, quality and quantity variables in designing effective large scale delivery systems.

Evaluation research is guided by four generic questions:

- Has the intervention project been delivered as designed?
- Has the project achieved predicted qualitative impact *(process means)* on a consistent basis in the project schools?
- Has the predicted quantitative impact *(product ends)* occurred in the project schools?
- Can these process and product impacts be ascribed to the project?

The first question deals with the organizational coherence and alignment of the project delivery system with the strategic approach of the intervention. The effectiveness of strategy is not being investigated if the delivery system is incoherent or if it fails to adequately express the theoretical basis of the intervention. The second question is concerned with the issue of 'internal validity' and asks whether the delivery of the strategy has succeeded in achieving the predicted or expected qualitative changes that are intended to achieve impact upon pupils in schools. Has it, in other words, succeeded in achieving the internal process *means* through which the external product *ends* are sought? The third question, in turn, is concerned with 'external validity' and asks whether these qualitative changes have, in fact, resulted in the expected impact upon children. It is best

investigated through the use of a quasi-experimental design in which project and control cohorts are tracked from pre to post-testing and the resulting data analyzed in terms of covariance. The fourth question is answered through the use of a quasi-experimental design through which both qualitative and quantitative data collection is organized and analyzed.

This approach to designing evaluations does not consider qualitative internal or process change in schools to be evidence of summative external product impact. In short, achievement of the means through which impact is supposed to occur is not a sufficient proof that the impact has, in fact, occurred. As a consequence, *the internal process objectives of interventions cannot be used to derive the external product indicators of an evaluation.*

The investigation of impact is instead based upon the concept that the single strategic purpose of any educational project or intervention is to improve the quality of the strategic outcomes of schooling; to improve, in other words, the performance of the pupils for whose sake the system exists and who are, furthermore, compelled to attend. It is very difficult to think of any other sensible strategic reason why any educational project or intervention should be established, even if the tactical and organizational routes to this goal vary widely in both focus and method.

30 Nymphe Street
Kensington
Johannesburg 2094
South Africa

*Eric Schollar and Associates c.c.*
*reg. no. CK 94/36640/23* *VAT reg. no. 414 014 7069*

**Telephone** **+27 (0) 11 622 7075**
**Fax** **+27 (0) 11 622 7589**
**Cell** **+27 (0) 83 255 0109**
**e-mail** **schollar@global.co.za**

# FINAL REPORT

# THE EVALUATION OF THE LEARNING FOR LIVING PROJECT

# A PROJECT OF THE BUSINESS TRUST AND THE READ ORGANIZATION

# 2000 - 2004

## 24 February 2005

## CONTENTS




## ***ACKNOWLEDGEMENTS***

***This evaluation has been conducted in association with McCay Projects***

***Thanks to Professor W. Elley for the design of the Grade 3 and 4 test instruments***

***Thanks to Buyile and Lindiwe Mkhulise, Luisa Roscani, Kim Gregory, Jabulile and Sibongile Maratana for their contribution to the field work for this evaluation***