

https://dspace.org/

Strategies for assessing the impact on primary school pupils of a language development programme directed at teachers

Dawjee, Rabia Cassim

2017
http://hdl.handle.net/10539/21616

Downloaded from a DSpace repository. https://dspace.org/

Strategies for assessing the impact on primary school pupils of a language development programme directed at teachers

submitted by

Rabia Cassim Dawjee

Student no.9510351G

A research report submitted in partial fulfilment of requirements of the Master of Education degree at the University of the Witwatersrand.

Declaration

I declare that this research report is my own unaided work. The report is being submitted in partial fulfilment of the requirements for the degree of Master of Education at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other university.

RDawjee

Rabia Cassim Dawjee

31 day of December 1998

Acknowledgements

I wish to thank all the people who participated in this study, TELIP staff and my supervisor, Dr. Jar Castle without whom this research would not have been possible.

Abstract

Millions of rands have been invested in teacher development programmes provided by South African non-governmental organisations (NGOs) since the mid-1970s. This investment increased dramatically through the 1980s and levelled off in the first half of the 1990s. Money will continue to be invested in in-service teacher education and training (INSET) as the need is still huge. However funders are asking the challenging question of what impact these programmes have on pupils.

The Teacher's English Language Improvement Programme (TELIP), a University based educational NGO which provides English language improvement courses for primary school teachers must also address this question: what impact does it's programme have on pupils? TELIP does not prescribe how and what teachers should be teaching their pupils but it directly develops teachers' language skills and confidence. Therefore, assessing the impact on pupils is a challenge. This study aimed to find an appropriate model of impact assessment for TELIP specifically and for other NGOs with similar orientations.

The research design was based largely on a literature review, interviews with specialists in the field and workshops with TELIP staff. The four research methods studied were; experimental design, longitudinal model, action research, and ethnographic research. The research findings indicate that a longitudinal study would be appropriate for TELIP to assess the impact of its English language development programme on pupils. There are four arguments favouring this model. Firstly, it provides an opportunity to do an in-depth study tracing the progress of the pupils in relation to the skills learnt and put into practice by the teacher during and after the TELIP course; secondly, baseline data could be collected by testing pupils; thirdly, pupils' written and spoken English can be monitored according to the criteria set by the project; and fourthly, the need to find a matching control group is eliminated. The findings also emphasise the need to work out specific assessment criteria and to consider contextual factors which impact on learning, such as, the authoritarian, bureaucratic nature of the education system, poor facilities and equipment, ethos of learning and teaching, other INSET courses that teachers do, and pupil maturation.

Key words: assessment strategies
INSET
assessment models
language development
South Africa

CONTENTS

		Page
Title Page		I
Declaration		II
Acknowledgements		III
Abstract		IV
CHAPTER ONE: INTRODUCTION		
1.1.	AIM	1
1.2.	BACKGROUND TO THE STUDY	2
1.3.	CONTEXTUAL FACTORS WHICH IMPACT ON TEACHING AND LEARNING	9
1.4	BACKGROUND TO IN-SERVICE TEACHER TRAINING AND DEVELOPMENT	16
1.5.	IMPORTANCE OF THE STUDY	20
CHAPTER TWO: RESEARCH DESIGN		
2.1.	INTRODUCTION	21
2.2.	THE RESEARCHER'S EDUCATIONAL PHILOSOPHY	21
2.3.	WHY THE 'MULTI METHOD' APPROACH	27
2.4.	VALIDITY	28
2.5.	METHODS OF RESEARCH	31
2.6.	STRENGTHS AND WEAKNESSES OF THE OVERALL DESIGN	40
2.7.	CONCLUSION	40
CHAPTER THREE: THE RESULTS OF THE STUDY		
3.1.	INTRODUCTION	42
3.2.	EXPERIMENTAL DESIGN	47
3.3.	LONGITUDINAL DESIGN	56
3.4.	ACTION RESEARCH	61
3.5.	ETHNOGRAPHIC RESEARCH	72
3.6.	SUMMARISING THE FINDINGS	79

CHAPTER FOUR: DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

4.1.	AIM OF THE STUDY	81
4.2.	RATIONALE	81
4. 3.	THE RESEARCH DESIGN	82
4.4.	THE RESEARCH OUTCOMES: IMPLICATIONS FOR TELIP	83
4.5.	IMPLEMENTATION	90

REFERENCES 98

LIST OF TABLES PAGE NUMBER

Table One: Where and How Interviews were carried out 35

Table Two: Main arguments for and against the four methods of research 43

Implementation Schedule 95

LIST OF FIGURES PAGE NUMBER

Figure One: The cycles of action research 63

LIST OF APPENDICES

Appendix A: Interview Schedule

Appendix B: Interview Schedule used for NGOs

Appendix C: Interview Schedule for Teachers

Appendix D: Analysis of Interviews held with experts in the field of evaluation

Appendix E: Analysis of Interviews held with teachers

CHAPTER ONE
INTRODUCTION

1.1. AIM:

In-service teacher education and training (INSET) projects in South Africa are being asked the daunting question of what difference, if any, their interventions are making, not only in the lives of teachers, but also, more importantly, on the pupils that the teachers teach.

The Teachers English Language Improvement Programme (TELIP), a University based educational non-governmental organisation (NGO) that provides English language improvement courses for teachers is also faced with precisely this question.

Funders are increasingly insisting that the ultimate product of INSET should be pupil achievement and that the impact of teacher development interventions on pupils should be assessed. In the context of this pressure, TELIP needs to assess what difference its programme makes to pupils. Assessing the impact on pupils of a language course directed at teachers is complicated. TELIP does not prescribe how and what teachers should be teaching their pupi' . I stead the programme works on the assumption that if the teacher's English language skills are improved then the teacher will be better able to teach in the medium of English and that pupils will benefit. For a programme that directly develops teachers' language confidence, assessing the impac on pupils is a challenge. This, however, is a critical challenge posed by funders.

Since teacher education and training is a huge need in this country, INSET providers will include not only state education departments, colleges, universities and technikons, but also the private sector and NGOs. As more money is invested in INSET the need to assess impact on students will become greater. The debate about what and how assessment should be done will continue.

This study aimed to

(a) investigate the advantages and disadvantages of research methods used in educational evaluation

(b) analyse these methods to see how they could be applied to impact assessment for TELIP specifically and generally to other NGOs with a similar orientation to TELIP.

1.2. BACKGROUND TO THE STUDY:

The TELIP Research and Development Unit (R& D), which is based at the Centre for Continuing Education at the University of the Witwatersrand in Johannesburg, South Africa, develops English enrichment courses for teachers, orients and trains tutors, provides ongoing support and follow up for teachers and tutors; and evaluates project operations. TELIP's learning materials have been informed by research into teachers' interests and language needs, principles of adult education, a learner centred methodology and current theories of language learning and teaching.

The TELIP project began as a sub-project of SELRP (Schools English Language Research Project) in January 1981. The staff of the SELRP project realised that the programmes they were developing for schools in Soweto (a black township in the South West of Johannesburg) could not be fully implemented because the English language proficiency of the teachers was inadequate to cope with the demands being made by the SELRP programmes. In 1982 TELIP became a project in its own right, sponsored by the Anglo American and De Beers Chairman's Fund Educational Trust for an initial period of three years At the time of writing this report in 1998, TELIP had been in existence for 17 years. Over the years, the project's funding base has grown and funds are now obtained from the Joint Education Trust and Eskom.

The primary target group of TELIP's courses is Black primary school teachers. Over the last five years, approximately 5000 teachers have completed TELIP courses.

The courses are designed to raise teachers' level of competence in English by concentrating on the major linguistic problem areas which have been diagnosed by TELIP and which the teachers themselves have indicated. Some of the linguistic problems faced by the teachers are:

* lack of English oral communication skills related to the fact that English is an additive[1] language and not mother tongue;
* lack of English reading skills;
* lack of English writing skills.

TELIP offers five courses to raise the teacher's level of competence in English:

The first course, **INTRODUCTION TO LANGUAGE AND LEARNING,** aims to develop participants' abilities to learn actively in a group and to communicate freely and comfortably in English. The course covers basic reading and writing skills and also aims to raise awareness of varieties in English and how these can be used appropriately for different contexts.

The second course, **INTRODUCTORY READING,** motivates participants to read more regularly and develops reading skills which will enable them to read with more confidence and understanding and to access information more effectively.

Course three, **PERSONAL WRITING,** aims to extend the reading and writing skills of participants and to link the experience of reading and writing to their own lives. Participants are encouraged to draw on their life experiences as source materials for their creative writing. The course builds on the skills developed in the first two courses.

The **ORGANISATIONAL SKILLS** course aims to develop English language and communication skills for organisational purposes. The course begins by looking at language appropriacy in different social and organisational contexts, and provides opportunities to develop a range of skills with sensitivity to these contexts. The course has a major focus on formal writing skills such as reports, letters and minutes.

---

[1] Additional language to a person's mother tongue.

The **STUDY SKILLS FOR ACADEMIC PURPOSES** course deals with different aspects of academic writing and study skills. The course aims to support teachers and other learners who are upgrading their qualifications and are studying at tertiary level.

TELIP's methodology is broadly learner centred. This approach implies building on what participants already know, encouraging people to relate what they learn to their own lives and actively involving them in the learning process. When this methodology is used for language learning, the emphasis is on using the language meaningfully to improve English competence and to increase confidence in using English.

Tutors are trained to facilitate TELIP groups. They undergo a 140 hour certified training course which develops their facilitation skills and equips them to teach TELIP and other English language courses effectively to adult learners.

Tutors are selected on the basis of their competency in English, sensitivity to and experience of working with adults, knowledge of English language teaching and learning issues, and prior experience in the area.

A criterion-based approach is used to assess participants' work. Each course assignment has a number of criteria which the tutor uses to assess the participant's level of competence. There are certain specified core criteria which participants must meet in order to proceed to the next course. All assignments are moderated by TELIP to ensure that required standards are maintained. On successful completion of each course, a certificate of competence is issued by Wits University through the Centre for Continuing Education.

The actual running of courses was not seen as the task of TELIP R & D and so implementation needed to be delegated to other appropriate structures. The Independent Teachers Centre (ITC) in Soweto and the Institute of Education, University of Bophuthatswana, were identified as such, and later approached about the matter.

Soweto was also seen as an appropriate implementation area for practical reasons. Running

courses for a township-based target group in Johannesburg was not a suitable arrangement. Transportation was problematic for learners who came from Soweto to attend classes in Johannesburg. Thus the University's geographic location affected the potential number of teachers who otherwise might have taken TELIP courses.

The Soweto Field Unit (SFU) was established under the auspices of the ITC, which was headed by Dr. Franz Auerbach, and housed at the Funda Centre in Soweto. The SFU was thus TELIP's first implementing arm.

From 1986 to 1996 TELIP courses were implemented through NGO structures, University structures, Education Departments, and industries such as Johannesburg Consolidated Industries (JCI) when it was an amalgamated company of gold, coal and platinum mines. The Bureau for In-Service Teacher Education and Development (BITED), located at the Johannesburg College of Education, also uses TELIP courses as part of a teacher's diploma course.

The relationship between TELIP and the organisations which implement TELIP courses is one of partnership. TELIP provides tutor training, course material, certification, evaluation, capacity building and support to organisations which are independent of TELIP at Wits University. They raise their own funds, employ their own staff and tutors, recruit learners and are responsible for organising and running TELIP courses.

The TELIP R&D unit is accountable to a Steering Committee on which organisations implementing TELIP courses are represented. This committee helps to guide and steer the activities of TELIP. The staff at TELIP are comprised of the following people:

* a co-ordinator who sees to the daily running of the project and fund raising;
* a materials writer responsible for updating and editing materials;
* an evaluator who tracks the profiles of learners and monitors the impact that courses have on teachers and other adults;
* an administrator who looks after the finances and also fulfils the functions of a receptionist.

The staff at TELIP are also involved in training tutors to facilitate TELIP courses. Training includes workshops and follow-up support to ensure that tutors are comfortable with the TELIP methodology, course content and assessment procedures. TELIP also offers a 140 hour certificated tutor training course which develops the tutors' facilitation skills and equips them to teach TELIP and other English language courses effectively to adult learners.

In 1994 action research was introduced to tutors who facilitate TELIP courses. Through action research tutors acquired practical experience of research. The authority role between tutors and learners broke down as tutors also became learners. Tutors reflected on their practice, identified problems, brainstormed solutions, implemented, and then repeated the cycles again. In implementing this process tutors had the support and guidance of TELIP staff. For TELIP as a research and development project, the spin off was that the data helped to re-work the training programme for tutors according to their needs. This process came to an end in 1997 as a result of lack of human resources on the part of TELIP. This initial positive experience of action research influenced TELIP staff to consider action research as a model for impact assessment.

From TELIP's inception the intention was that pupils should benefit from the greater linguistic competence of their teachers. However, this was not clearly explained in project documents. Instead, the focus of TELIP's work has been on improving teachers' language skills. Internal evaluations have concentrated on the impact that the courses have on teachers' professional growth. (see for example, Dawjee (1994) - Phase 1: Impact evaluation of courses on teachers).

TELIP has undertaken both formative and summative evaluations of its courses. A formative evaluation of the programme was carried out in the very process of course construction. The emphasis in this phase was on the teaching process, and the information generated was used to improve the curriculum during implementation. Summative evaluation was carried out at the end of a lengthy curriculum development phase in 1994. These evaluations did not address the question of impact on pupils.

The questions facing TELIP are what aspects of its training lend themselves to pupil assessment and how best to do the assessment? The assessment of end users has been emphasised by funders recently. For example, the Anglo American and De Beers Chairman's Fund have asked TELIP to evaluate its English language improvement courses for teachers in terms of outcomes for pupils in the classroom:

> Wh. matters is that whatever teaching methods are used or how pupils respond to them, a programme should be evaluated on its ultimate results rather than on just whether the correct methodology is being applied. This means that the children's performance before and after the intervention needs to be looked at. In education, what the pupil has learnt and understands is the indicator of proper teaching. Whether this should be on a one-off basis or through continuous assessment of class work is another debate, but it should nevertheless be a priority in any evaluation.
>
> (George, 1997: 1)

The pressure to evaluate student outcomes is not unique to TELIP but applies to most INSET NGOs in the South African context. The Joint Education Trust (JET), an important funding agency that provides funds to ninety eight NGOs working in the field of teacher education, held an evaluation conference in February 1996 to assess how INSET NGOs conducted evaluations. At this conference several problems were highlighted but the point made was that INSET evaluation usually failed to measure outcomes on pupils.

Professor Johann Louw of the University of Cape Town stated that:

> Social programmes follow an intrinsic cause and effect logic. If we enrich the teaching practices of school teachers by providing an in-service training programme, then the educational performance of school children taught by them will improve. This means that outcomes have to be evaluated.
>
> (Louw, 1996:4)

Professor Jonathan Jansen stated his belief that:

> Teacher development delivered through NGOs in the form of in-service education (INSET) does not produce significant learning gains in the classroom. One evaluation after another has delivered the same finding: that while INSET provides important motivational benefits to practising teachers, and begins to influence the behaviours of participating teachers, such programmes simply do not translate into learning gains for students.
>
> (Jansen, 1996:14)

Jansen went on to say that although these programmes do not translate into learning gains for students, millions of Rands have been invested in teacher development programmes of South African NGOs since the mid 1970s, increasing dramatically through the 1980s and levelling off in the first half of the 1990s. Many funders were present at this influential conference and agreed that evaluations should show that the money invested in teacher INSET eventually benefited the students.

There is some evidence that teachers teach differently because of their participation in teacher development programmes. For example, TELIP evaluations (1996) have shown that teachers have moved away from a teacher dominated instruction to small group learning. But generally there is an unspoken assumption in INSET that the direct training benefits to teachers translate into learning gains for their students. Students are the ultimate beneficiaries of INSET. They are expected to learn more and better because of their teachers' exposure to INSET. But impact studies have focused on immediate training benefits to the teachers and have neglected to assess the benefit for students.

At the same conference Dr. Nick Taylor of the Joint Education Trust (JET) summarised the state of evaluation practice in South Africa in 1996. He declared that all evaluation studies of INSET produced fascinating and useful insights into teaching and learning practices, conditions in schools and the valuable work undertaken by the NGOs. However, in most cases, the methods employed in the studies lay below the quality threshold required to inspire confidence in the validity of their principal findings. For example, a number of evaluators concluded, on the strength of single class visits to a small number of project teachers, that the practices of these teachers had improved and that this improvement was due to the project intervention. What was missing in the conclusions, according to Taylor, is careful comparison and control.

JET's conclusion was that the whole spectrum of quantitative and qualitative research methods needed to be used where appropriate, and that the nature of the information required should be the determining factor in choosing evaluation tools rather than some predetermined

political position. There was also allusion to the fact that evaluation was relatively new in South Africa and that there was a need to develop more effective assessment strategies. These various observations began to push NGOs to evaluate learning outcomes on the pupils of teachers who had been through INSET programmes. The question which arose was how best to do this?

There is no simple answer to this question, especially for TELIP. One reason for this is that many variables impact on the transfer of skills from the teacher to pupils in the classroom. Some of these variables include:

* teachers and pupils assimilate and use information differently;
* some forms of learning don't translate easily into measurable/observable behaviour, for example self confidence;
* pupils may display improvement in writing and reading skills in ways not easily assessed, i.e. one might have to look at reading and writing skills used outside the classroom;
* pupils may improve or do poorly because of factors outside the school situation
* the school context may or may not lend itself to implementing the skills learnt in the TELIP programme;

1.3. CONTEXTUAL FACTORS WHICH IMPACT ON TEACHING AND LEARNING

This section first provides insights on contextual factors in South African schools and how these impact on learning and teaching. This is followed by a discussion of what these contextual factors mean for TELIP when deciding how to conduct an impact study.

Anderson and Burns (1989) state that since the primary purpose of instruction and teaching is learning, a major goal of classroom research is to connect instruction and teaching with the learning of students. They caution that this may seem easy but it is important to remember that teaching and learning take place in a complex chain of interrelated events and that the influence of teaching on learning tends to be context specific.

Implementation of new knowledge acquired from INSET depends not only on the teachers but also on the school organisation, and on the provision of support, encouragement, and assistance. (Veenman, S. et al, 1994) Yet there are many contextual factors which inhibited the success of INSET delivery in South African schools during the apartheid era. Writers such as Hartshorne (1984) and Millar, et al (1984) underlined the problems created by apartheid. At the institutional level, the authoritarian, bureaucratic nature of the education system did not encourage school-based innovation. The needs of the system triumphed over the needs of individual teachers. Poor systems management and inadequate leadership in schools presented another constraint. (Hartshorn, 1985; Levy 1987; Mehl, et al 1989). Schools were embattled by political tensions and inadequate resources (Bateson, 1985). Thus, the school ecology was not receptive to innovation. The government of the African National Congress (ANC) is committed to whole school development and as this policy is implemented innovative methods of teaching will receive the support to thrive and grow in South African schools.

According to Hofmeyer and Pavlick (1987), teacher characteristics also serve as constraints on the delivery of INSET. Most African teachers are underqualified, young and inexperienced, with a limited knowledge of their teaching subjects and English, the dominant medium of instruction. They suffer low salaries and professional status, and poor working conditions. Their morale has been seriously eroded by the turmoil surrounding black schools since 1976. Survival rather than innovation tends to be the main focus of teachers' attention.

Bateson (1995) adds that the already poor facilities and equipment provided for urban black schools in the apartheid era were allowed to deteriorate and often to be destroyed through a total lack of resupply, care and maintenance. In the 1990s under the ANC the situation has improved in some schools but in many it still remains the same. Most black schools are not places where any person would want to be. They are ugly, dilapidated, dirty and sometimes unsafe. Materials and equipment to assist teachers in the classroom are hardly visible. Those that are visible are usually broken or unusable. In most cases nobody seems to care. TELIP works with teachers who teach in schools with the same problems. Under such conditions,

implementation of INSET training is extremely difficult and wh... it is implemented it has both successes and failures.

Whatever the conditions of schools in black urban areas described by Bateson, Vinjevold (1996) in a study of the Northern Cape primary schools found school buildings and classrooms in good condition. There were no schools under trees or in tents, huts or rondavels as were said to be in other provinces. Only three of the sixty local schools in her sample were built of prefabricated materials but in all cases these were sturdy and comfortable. The windows and doors of all schools were in good repair and a number of schools had recently received a fresh coat of paint.

Setting aside the issue of the poor physical conditions in black schools, there is a widespread problem of low work rate. Taylor (1995) reports that "one of the most striking and widespread findings of evaluations of teacher education interventions is the low work rate in many South African schools". He observes that a number of evaluations mention this as a factor which severely inhibited both their own work, and the implementation of the programme under evaluation.

The magnitude of the problem is indicated again by Bateson (1995). His study showed that in Kwa Zulu Natal, 50 days of the school year were lost to effective education. Taylor goes on to say that this was corroborated by Peacock's work in the same province a year later which estimated that the average school week was reduced by 40%.

* an inordinate amount of time was given to time tabling (Schollar,1995) and testing (Peacock, 1995)
* early closing of the school day (Schollar, 1995) and the week, (Peacock,1995, Bateson,1994)
* no school on pay day (Bateson, 1996).
* lengthy preparations for athletics (Peacock, 1995:Schollar,1995:Khulisa,1995)
* violence and strikes (Jansen and Perold,1994, Bateson)
* doubling of classes (Schollar,1995, Peacock,1995)

Tessmer (1991) emphasises that the success or failure of instruction depends to a large extent on the context. The people who use the instruction, where they use it, and how they use it can impair the best designed instruction, or facilitate the work.

These enormous problems in the education of black South Africans make it very difficult for teachers to implement the skills learnt in TELIP. However the Impact Evaluation Phase I (Dawjee,1994) shows that despite the difficulties that teachers experience, they attempt to implement skills learnt in the TELIP courses in their classes. The two groups of teachers interviewed in this research also confirmed that they implement skills learnt in the TELIP courses. But what exactly is the nature and scope of this implementation? How much of it benefits pupils? These questions need to be assessed within the contextual factors described above.

Whilst the above outline provides a picture of the teaching and learning context as crucial to the impact on learning, it does not touch on the need to also consider the individual pupil's background and how this impacts on learning. A further problem in this literature is that it fails to make explicit distinctions between urban and rural schools. In South Africa urban schools generally have better teaching facilities and better qualified teachers. Pupils from urban schools are exposed to television and radio programmes which influence their language acquisition. In rural areas where electricity is not always available such exposure is limited for pupils.

An INSET programme like TELIP works with teachers who come from urban and rural schools. This means that an impact study will also need to look at how urban and rural factors impact on learning.

The above conditions in the majority of South African schools as highlighted in South African literature need to be kept in mind when conducting an impact assessment. In addition the literature also cautioned about other variables that impact on learning :

* **Ethos of learning and teaching**

In many South African schools pupils do not view learning as a serious matter but are happy to while away their time. This non-caring attitude is also prevalent amongst many teachers who would engage in any activity other than teaching during school hours. Many pupils arrive late and leave early and this causes a low work rate. Under circumstances in which very little teaching and learning is taking place, an in-depth impact assessment within a given period may be difficult to conduct.

* **Other INSET courses that teachers do**

As highlighted in the introduction of this study a large majority of teachers are under or unqualified. In an attempt to obtain qualifications, most teachers study with NGOs or teacher training institutions which provide distance education courses to upgrade qualifications. The majority of teachers doing TELIP courses are also engaged in other studies. It will be important to draw up a profile of teachers in the study to monitor whether these other courses influence learners in similar ways to TELIP. This is important in order to make valid claims that TELIP courses, and not other courses that teachers do, are responsible for changes.

* **Class teacher may not remain constant**

Often in the course of the year teachers are transferred to areas where they might be urgently needed. A teacher could also be made to change classes in her or his school if there is a shortage of teachers. As a result there is no guarantee that the teacher will remain with the same class for a year. If the teacher in the study is transferred and the replacement teacher is not doing TELIP courses the impact assessment will be disrupted.

* **Teachers own personality impacting on teaching and learning**

Some teachers seem to have a natural talent and the motivation to improve their teaching and learning on their own without depending on INSET courses.

This factor also needs to be looked for when observing teachers and drawing inferences about the impact of TELIP's courses.

* **Changes taking place in the school that impact on learning**

The arrival of a new principal or Head of Department can lead to changes in how the school operates. This means that both pupils and teachers have to adjust to new leadership styles. New principals or Heads of Department may not be supportive of innovative methodology in teaching and this could hamper how much the teacher is able to transfer to the class. They could also be too bureaucratic or too lenient which affects how pupils behave and learn in the class. All major research on innovation and school effectiveness shows that the principal strongly influences the likelihood of change, but it also indicates that most principals do not play instructional or change leadership roles. Berman et al (1977) found that "projects having the active support of the principal" were the most likely to fare well. Principals' actions serve to legitimate whether a change is taken seriously and to support teachers both psychologically and with resources. Berman, Mc Laughlin and associates (1979) note that one of the best indicators of active involvement is whether the principal attends workshop training sessions. Principals need to gain some understanding of the training sessions attended by teachers in order to provide support for implementation. The teachers who attend TELIP courses indicate that principals know nothing of the courses that they attend and have no interest in getting to know them. As a result there is no support for implementation. In the words of one interviewee "The context of learning is important for TELIP. Questions that TELIP needs to ask are - are the teachers encouraged to try new things? Are principals involved in motivating and supporting teachers?"

* **Pupil drop out rate**

Pupils leave school for financial reasons particularly because parents are unable to pay fees or buy books and uniforms. They also leave because of

lack of motivation during the course of the year. If too many pupils drop out of a sample the validity of the study becomes questionable.

* **Pupils maturation having an effect on performance**

  As pupils grow in age their performance also changes for better or worse between observations. This means that the changes that take place could be independent of the treatment given to them.

* **Transfer of skills**

  TELIP staff highlighted that language skills learnt may not be transferred immediately to the classroom as the natural approach to language learning (which TELIP uses) does not demand speech from learners before they are ready for it. In the natural approach there is an emphasis on exposure or input, rather than practice; optimising emotional preparedness for learning; a prolonged period of attention to what the language learners hear before they produce language; and a willingness to use written and other materials as a source of comprehensible input. There is always a "silent period" which is then followed by language usage (Richards and Rodgers, 1986:113). Any impact assessment needs to be informed by the language learning theory which underpins provision.

The above conditions in the majority of black schools in South Africa need to be kept in focus when an impact assessment of INSET courses is carried out as they have a bearing on teaching and learning.

1.4. BACKGROUND TO IN-SERVICE TEACHER TRAINING AND DEVELOPMENT

To understand fully why INSET plays and will continue to play a central role in education in South Africa, it is necessary to explore briefly the history of apartheid education. According to Kallaway (1984) Bantu[2] education policy in the 1950s and 1960s was seen as the mechanism for the reproduction of rigidly segregated occupational structures in which blacks were virtually excluded from all job categories except that of unskilled labourer. The Bantu Education Act of 1953 was the major instrument through which the regime attempted to shape education to perform this function. However by 1976 the government of the day was plunged into a crisis when students in Soweto took to the streets in protest against the use of Afrikaans as a medium of instruction in secondary schools (Unterhalter and Wolpe,1991: 5).

A resolution of the crisis took a long time to work out but limited improvements such as the scrapping of Afrikaans as a language of learning and the scrapping of the hated label Department of Bantu Education in favour of Department of Education and Training, took place. (Kallaway 1984:350). The Education and Training Act of 1979 paved the way for a decision by most schools in South Africa to use English as a medium of instruction from Standard three[3] upwards. Grave concern was expressed by officials of the Department of Education and Training about the problems involved in transferring to English after twenty years of the use of the vernacular. A particular concern at this point was the lack of English proficiency amongst black primary school teachers.

The University of the Witwatersrand's response to these problems was to set up, during 1979, the Schools' English Language Research Project (SELRP), financed by the Anglo American and De Beers Chairman's Fund Educational Trust. SELRP aimed to improve the quality of

---

[2] 'Bantu' refers to the indigenous African population of Southern Africa.

[3] This is a child's fifth year at school

English language teaching in primary schools for blacks together with the effective use of English as the medium of instruction across the curriculum, from Standard three to Standard five.

The Teacher's English Language Improvement Project (TELIP) was born out of necessity, when the staff of the SELRP project realised that the programmes they were developing for schools in Soweto could not be fully implemented because the English language proficiency of the teachers was inadequate to cope with the demands being made on them.

The inequalities in South African education, tied to the policy of apartheid, are documented in detail by Kallaway (1984), Nkomo (1990) and others. What is important to state here is that teacher education and development remains crucial to the broader improvement and transformation of education in South Africa. A recent national teacher education audit (1995:31) found that 36% of the total teaching force is underqualified and that 60% of African teachers fall into this category. The report indicates further that the majority of African teachers holding the senior posts of principal, deputy principal and head of a department in primary schools are also underqualified. There is, thus, an urgent need to develop competent and effective teachers and school managers who can bring about meaningful, qualitative change in education. INSET programmes play a crucial role in the provision of such education.

The number of NGOs providing INSET increased greatly since the Soweto uprising in 1976. The reason for this is related to private sector involvement in education.

During the heyday of Bantu education in the 1960s the state discouraged private sector involvement in education. In the 1970s state education policy for the first time permitted private sector investment in education and training facilities for blacks in the urban areas of South Africa. Since 1976 in particular there has been a massive increase in corporate investment in education and training, and by the 1980s the ethos of "social responsibility" was well established.

One important area of corporate involvement in education was through funding of teacher upgrading programmes. According to Millar (1985) the massive disparities in the quality of schooling provided to white and black South Africans were paralleled by sharp contrasts in teacher supply and qualifications. For equality in educational provision by the year 2020 at a pupil ratio of 30:1, it was estimated that 300,000 teachers would need to be trained. This situation prompted the development of intensive upgrading programmes for black teachers. Such programmes were administered by the employing education departments, sometimes with the assistance of tertiary institutions and private sector funding. Private sector funding was also channelled to NGOs providing In-Service Teacher Training. In the Department of Education and Training the number of underqualified teachers in upgrading programmes rose from 300 in 1974 to 7000 in 1984.

Private sector initiatives in education have been spearheaded by large firms, most notably the Anglo American Corporation (AAC) which by the 1970s had come to dominate the South African economy. The most significant private sector actors in education after 1976 were the Anglo American Corporation Chairman's Fund (AACCF) and the Urban Foundation (UF), which, while formally independent of each other, were closely related.

In general terms, corporate investment in both formal and non formal education has sought to promote:

* the improvement of black education;
* skills development;
* in-house training and education for the employees of individual companies.

According to Wolpe and Unterhalter (1991) the motivation behind corporate contributions to education and training has been twofold:

* to promote political and social stability, and
* to provide both for general and specific human resource needs for economic growth.

Between 1980 and 1985 the proportion of funds going to teacher development increased from 32% to 58% of the total expenditure on social development. The AACCF also supported new research initiatives in "white' universities. By the 1960s donations to universities for research purposes had been made tax deductible.

The AAC increased its expenditure on education as it grew to become the major conglomerate in South Africa (O' Dowd,1987). The Financial Mail estimated that in 1987 the top 100 companies in South Africa collectively spent a higher proportion (ten per cent) of their employees' payroll on social responsibility activities than the average Fortune 500 company in the US (0,4 per cent) (Financial Mail Supplement, 30.1.89). It is important to remember that this is only their estimate. Foreign firms in South Africa, like Mobil, also invested heavily in education and business development during the 1980s, mainly in order to appease shareholders in their home countries.

That mining conglomerates were the first to invest in education is not surprising given their pre-eminent position in the economy. According to O'Dowd, the AACCF always regarded education as key for it was "an area where you could influence the future" (1987: 98). As a consequence, the AACCF modified its objectives to encompass both reactive and proactive goals. It began to seek out "model" projects that would influence the direction of state education policies. The intention of the AACCF was to sponsor projects that would be exceptionally productive for providing an educated labour force and have a multiplier effect. The multiplier effect meant that if teachers were better qualified, then pupils would benefit and make a better contribution to society.

It was this availability of funding for education in the 1970s and 1980s that resulted in NGOs occupying a centre stage and providing INSET. But in the 1990s the edu ational NGO sector is in crisis. Foreign and local corporate funding is now allocated directly to government with the understanding that large scale intervention is needed to correct the imbalances created by the apartheid regime.

However, in-service teacher education and development, whether provided by Education Departments or by NGOs remains crucial. The national Department of Education sees teacher development as part of a larger process of reconstruction and development in South Africa, and as a central pillar of national human resource strategy in global change. The Department's vision is centred "on a community of committed, competent and reflective teaching professionals who can help to establish and sustain peaceful and purposeful learning environments" (Department of Education, 1997:8).

This commitment to teacher development also implies assessment of the impact of training on pupils or else its worth will become questionable. Impact assessment on pupils requires providers of INSET to have clear understanding of different models of assessment that they can use to carry out the study.

1.5. IMPORTANCE OF THE STUDY

This research is important as it provides an understanding of different research methods which could be used for impact assessment on pupils both for TELIP as well as other NGOs who are under pressure to provide evidence of the impact of INSET training on pupils.

1.6. OUTLINE OF THE REMAINDER OF THE STUDY

Chapter two gives an account of the research design, provides a rationale for the choice of research methods, explains how the study was conducted and states the limitations of the design.

Chapter three presents the results of the research findings drawing on the literature review, interviews and workshops.

Chapter four contains a discussion of the findings and suggestions for a model for impact assessment of TELIP courses on pupils.

CHAPTER TWO
RESEARCH DESIGN

2.1. INTRODUCTION

Like all research designs, this one has a philosophical underpinning. This philosophical underpinning influences how I teach and how I conduct research. This chapter first describes the educational philosophy that underpins the research design, then provides a description of the research method and a rationale for the choice of the multi-method approach.

2.2. THE RESEARCHER'S EDUCATIONAL PHILOSOPHY

I think of myself as eclectic about my beliefs regarding teaching and learning. As Elias and Merriam (1980:206) note, "in this approach one chooses certain elements from different theories and operates according to those principles."

The greatest impact on my philosophy and ethical beliefs has been from the humanist movement, particularly writers such as Carl Rogers and Knowles. What appeals to me most about humanism is the belief in the natural goodness of humankind, in freedom of choice, in the dignity and worth of all people, and in the value of establishing an environment in which the potential inherent in every person can be developed. Because I view learning as a highly personal endeavour, I act as a facilitator, a helper and most importantly as a partner in the learning process rather than as the expert. I see my role as that of facilitating learning related to identified needs.

In the teaching situation I encourage discussion, the raising of questions, small group activity and a positive attitude about learning. I also undertake ongoing informal needs assessment so that course content provided is informed by what learners need. In my experience it is not always possible to address every need of all the learners, but I try to provide what most learners need.

Outside the classroom learners are encouraged to approach me if they need to bounce off ideas or to work through problems. Because I believe that each person has potential and can be developed, I provide feedback on written work rather than merely passing judgement on the work. If learners identify topics they wish to explore, I secure new information for them and in this way serve as a resource locator.

My role as facilitator is not without problems. I work with adults who have been so deeply conditioned by their previous schooling to perceive their role as learners to be that of dependent, almost passive recipients of content, that when they enter any activity labelled education they fold their arms and expect to be taught. It takes time, patience and careful planning to move adults to accept their roles as participatory learners. Drawing the content of lessons from their experiences helps them to engage more in the learning process but I often struggle with the narrowness of limiting learning only to their experiences. Ideally learning should be a window to wider experiences. I also find that since the emphasis in humanistic education is on respect for the dignity and worth of every individual, views and opinions often are not challenged and this creates an obstacle to exploring issues in great depth in the learning situation.

The role of the facilitator as helper can also be misused when learners expect the facilitator to solve personal problems. Here I find myself having to be careful not to resolve personal problems for which I could be blamed if the resolution is not acceptable to the learner.

Although the humanist movement has influenced me greatly, I also find myself drawing from other philosophies. For example I place emphasis upon critical reflection of experiences instead of taking experiences for granted. I use learners' experiences to build on weaknesses and strengths and move them to different levels of learning. There are times when I find myself providing information, especially when I sense that learners lack knowledge or tools to help them to become self directed. One of my main criticisms of humanism is that it takes for granted that all learners are self directed. I believe that sometimes one has to lead and support people on the road to self direction, and so sometimes I act as the expert who knows more than the learner. I place emphasis on using teaching aids such as colourful

visuals. This could be seen as a stimulus response way of learning. I draw on other philosophies and my role as teacher is neither totally passive nor totally directive.

My educational beliefs translate not only into my teaching and learning situation, but also into how I conduct research. Two major research paradigms, the positivist and anti-positivist paradigms, are explained in detail in the next paragraph. I find myself leaning towards the anti-positivist paradigm rather than the positivist paradigm but like my teaching philosophy I seem to draw from both these paradigms when conducting research.

Two Contending Views of Science

Cohen and Manion (1994) identify two contending views of social science that inform emerging philosophies. These views are the established, traditional view and a more recently emerging view. The former commonly known as positivism holds that the social sciences are essentially the same as the natural sciences and are therefore concerned with discovering natural and universal laws regulating and determining individual and social behaviour. The latter view, commonly known as anti-positivism, while sharing the rigour of the natural sciences and the same concern of traditional social science to describe and explain human behaviour, emphasises that people differ from inanimate natural phenomena and from each other. These contending views , and their corresponding reflections in educational research, stem in the first instance from different conceptions of social reality and of individual and social behaviour.

Assumptions Underpinning the Two Views

Burrell and Morgan (1979) identified three sets of assumptions underpinning these two views of social science. The first set of assumptions are of an ontological kind - assumptions that concern the very nature or essence of the social phenomena being investigated. The researchers ask, is social reality external to individuals imposing itself on their consciousness? Is reality of an objective nature or the result of individual cognition? Is it a given out there in the world or is it created by one's own mind? These questions spring directly from what is known in philosophy as the nominalist-realist debate. The nominalist view holds that objects of thought are merely words and that there is no independently

accessible thing constituting the meaning of a word. The realist position, however, contends that objects have an independent existence and are not dependent for it on the knower.

The second set of assumptions identified by Burrell and Morgan (1979) are of an epistemological kind. These assumptions concern the very bases of knowledge - its nature and forms, how it can be acquired, and how communicated to other human beings. One view sees knowledge as hard, objective and tangible and the other sees knowledge as personal, subjective and unique. How one aligns oneself in this debate will affect how one conducts research. To subscribe to the former is to be positivist and to the latter anti- positivist.

The third set of assumptions concerns human nature and, in particular, the relationship between human beings and their environment. One image portrays human beings as responding mechanically to their environment and the other as initiators of their own action.

The three sets of assumptions have implications for the methodological concerns of researchers. If one adopts an objective approach to the social world and regards social reality as being hard, real and external to the individual, then the research methods will be traditional options such as surveys, experiments and the like. On the other hand if one favours the more subjective or anti- positivist approach and views the world as being of a much softer, personal and human-created kind, one will select from a comparable range of recent and emerging techniques such as personal accounts, participant observation and personal constructs.

When one subscribes to the view which treats the social world as if it were a hard, external objective reality then the methods of scientific investigation will be quantitative. If one favours the alternative view of social reality which stresses the importance of the subjective experience of individuals in the creation of the social world, then the search for understanding focuses upon different issues, and approaches them in different ways. The approach now takes on a qualitative rather than a quantitative aspect.

Each of the two perspectives on the study of human behaviour outlined above has profound implications for research in classrooms and schools. The choice of a problem, the formulation of questions to be answered, the characterisation of pupils and teachers, methodological concerns, the kind of data sought and their mode of treatment - all will be influenced by the viewpoint held.

The Debate Concerning Qualitative and Quantitative Approaches

In the 1970s, debates among social scientists concerning differences between quantitative and qualitative research gained ground. The pivotal point for much of the controversy was the appropriateness of a natural science model to the social sciences. Proponents of qualitative research argued that this method was inappropriate for studying people. Research methods were required which reflected and capitalised upon the special character of people as objects of inquiry. A qualitative research strategy, in which participant observation and unstructured interviewing were seen as the central data gathering planks, was proposed since its participants would be able to get closer to the people they were investigating and to be less inclined to impose inappropriate conceptual frameworks on them.

Increasingly, the terms quantitative and qualitative research came to signify much more than ways of gathering data, they came to denote divergent assumptions about the nature and purposes of research in the social sciences.

In some treatments, for example, Burrell and Morgan (1979) and Cohen and Manion (1994), qualitative and quantitative paradigms a e viewed as competing views about the ways in which social reality ought to be studied, and as such they are essentially divergent clusters of epistemological assumptions, that is, of what should pass as warrantable knowledge about the social world. For other writers, for example Bhola (1990) quantitative and qualitative researches are simply denotations of different ways of conducting social investigations which may be appropriate to different kinds of research questions and are even capable of being integrated. Keohane, and Verba (1994) state that the best research "often combines the features of each."

A number of writers have proposed alternative terms for quantitative and qualitative research. For example Guba and Lincoln (1982) propose a contrast between rationalistic (quantitative) and naturalistic (qualitative) paradigms, while Evered and Louis (1981) use a contrast between "inquiry from the outside" and "inquiry from the inside."

Lincoln and Guba (1989) assert that commitment to the scientific paradigm leads to an overdependence on formal quantitative measurement. The rigour that this paradigm appears to promise rests on the hardness of the data that are fed into the process. Hard data implies quantifiable data, data that can be measured with precision and analysed with powerful mathematical and statistical tools.

Since the methods of science promise to provide us with information about the way things really are, they claim a certain authority that is hard to resist. Hanna Arendt (1963) has noted this coerciveness of truth. Scientific truth is non negotiable. Anything being researched that is supported by positivism is locked in as the right thing to do. The use of the scientific method closes out alternative ways of thinking. Since science discloses the truth about things, any alternative must be in error. Because science is putatively value free, adherence to the scientific paradigm relieves one of any moral responsibility for his or her action. A researcher cannot be faulted for just telling the truth or for giving the facts.

Lincoln and Guba (1989) suggest an alternative approach which they call constructivism. This approach hardly resembles science at all, particularly in its basic assumptions which are virtually polar to those of science. It denies the existence of an objective reality, asserting instead that realities are social constructions of the mind, and that there exist as many such constructions as there are individuals. They argue that science itself is such a construction.

The constructivist paradigm denies the possibility of subject-object dualism, suggesting instead that the findings of a study exist precisely because there is an interaction between an observer and observed that literally creates what emerges from that inquiry. Methodologically the naturalistic paradigm rejects the controlling, manipulative approach that characterises science and substitutes a process that takes full advantage, and account, of

the observer/observed interaction to create a constructed reality that is as informed and sophisticated as it can be at a particular point in time.

According to Patti Lather (1986) "post-positivism" is characterised by the methodological and epistemological refutation of positivism and by increased visibility of research designs that are interactive, contextualised and humanly compelling because they invite joint participation in the exploration of research issues. Post-positivism is marked by approaches to inquiry which recognise that knowledge is "socially constituted, historically embedded, and valuationally based" (1986: 259). Theory serves an agentic function, and research illustrates rather than provides a truth test.

I view knowledge as personal, subjective and unique, and human beings as initiators of their own action, and therefore find myself located more in the anti-positivist paradigm. This is also in line with my leaning towards the humanist model of learning and teaching. Just as I have stated that I draw on other educational philosophies when the teaching situation demands it, I also draw on data collecting methods which might fall in the anti positivist paradigm.

The data for this investigation into strategies for impact assessment was gathered largely through literature review, supported by in-depth interviews with stakeholders and other experts in the field of evaluation. Two workshops were also conducted with TELIP staff. The interviews and workshops were used to gather data as I believe that knowledge is not only gained from books but also that people are able to construct knowledge as well. These methods which will be discussed below fall mostly in the anti-positivist paradigm and are consistent with Lincoln and Guba's (1989) concept of constructivism.

2.3. WHY THE 'MULTI-METHOD' APPROACH?

I chose to collect the data through three different methods and from different stakeholders in the research. The rationale for using a variety of methods and sources of information was to enrich the quality of the data.

The use of multiple methods, or the 'multi-method' approach as it is sometimes called, contrasts with the generally more vulnerable single method approach that characterises so much research in the social sciences. Triangular techniques in the social sciences attempt to map out, or explain more fully the richness and complexity of human behaviour by studying it from more than one standpoint. Cohen and Manion stress the importance of the multi-method approach in the following way:

> *Firstly* the researcher needs to be confident that the data generated are not simply artefacts of one specific method of collection. This confidence can only be achieved when different methods of data collection yield substantially the same results. Furthermore, the more the methods contrast with each other, the greater the researcher's confidence. If, for example, the outcomes of a questionnaire survey correspond to those of an observational study of the same phenomena, the more the researcher will be confident about the findings.
>
> (1994: 234)

The second advantage is that triangulation techniques will help to overcome the problem of 'method boundedness.' One of the earliest scientists to predict such a condition was Boring who wrote:

> as long as the new construct has only the single operational definition that it receives at birth, it is just a construct. When it gets two alternative operational definitions, it is beginning to be validated. When the defining operations, because of proven correlations, are many, then it becomes reified.
>
> (cited in Cohen and Manion, 1994:234)

A multi-method approach overcomes this limitation providing a full and detailed study. This method also helps with the issue of validity on which I will expand below.

2.4. VALIDITY

The issue of validity arises from a need to persuade researchers and others of the authenticity and trustworthiness of the methods of collecting and presenting information and the interpretations which are derived from it. Maxwell (1992) states that validity has long been a key issue in debates over the legitimacy of qualitative research. Proponents of quantitative and experimental approaches have frequently criticised the absence of "standard" means of assuring validity, such as quantitative measurement, explicit controls for various validity threats, and the formal testing of prior hypotheses. This criticism is supported by the fact that existing categories of validity for example, concurrent validity, predictive validity,

convergent validity, criterion related validity, internal validity, are based on positivist assumptions that underlie quantitative and experimental research designs. Qualitative researchers have generally responded either by denying the relevance of the quantitative or scientific paradigm for what they do (Guba and Lincoln, 1989) or by arguing that qualitative research has its own procedures for attaining validity that are simply different from those of quantitative approaches (Kirk and Miller, 1986).

Reason and Rowan (1981) and Guba and Lincoln (1981) offer important suggestions in this regard. Reason and Rowan advocate borrowing concepts of validity from traditional research, but caution researchers to revise and expand those concepts in ways appropriate to an interactive dialogic logic. Guba and Lincoln state that in order to fulfil the minimum requirement for assessing validity in new paradigm research, the techniques of triangulation, reflexivity and member checks should be enlisted. Lather (1984) offers four concepts of validity appropriate for qualitative research. The four concepts she describes apply to this research.

Firstly, she talks about triangulation. Triangulation is critical in establishing the trustworthiness of data. The researcher must consciously utilise designs that allow counter patterns as well as convergence if data are to be credible. In this research the use of literature review, interviews, and workshops with stakeholders generally provided a convergence of data. Sometimes the data from the three methods concurred and occasionally they provided different perspectives. These are openly stated and inferences are drawn from them in the results.

Secondly, construct validity must be dealt with in ways that recognise its roots in theory construction (Cronbach and Meehl, 1995). Our empirical work must operate within a conscious context of theory building. Where are the weak points of the theoretical tradition we are operating within? Are we extending theory? Revising it? Corroborating it? Determining that constructs are actually occurring, rather than that they are merely inventions of the researcher's perspective, requires a self critical attitude toward how one's own preconceptions affect the research. Building emancipatory social theory requires a

confrontation with, and respect for, the experiences of people in their daily lives to guard against theoretical imposition. A systematised reflection on the data helps to prevent imposition of pre conceived ideas and also throws light on any new and emerging ideas. In this research the researcher constantly reflected on the data emerging from the three methods and her own ideas and found that along the journey her model of assessment changed. For example, she started off with a strong belief that action research was an appropriate model but upon reflection on data this belief began to change.

Thirdly, face validity needs to be reconsidered. Kidder (1982) contends that although it has sometimes been treated lightly and dismissed, face validity is relatively complex and inextricably tied to construct validity. Face validity is operationalised by recycling description, emerging analysis and conclusions back through at least a subsample of respondents. This is best illustrated in Reason's account of a fellow researchers' activities,

> Madison went round the research cycle (many) times. Over and over again he interviewed, theorised, fed back his theories to the next lot of students, interviewed more (observed, read diaries, gave tests etc.) theorised, fed back, tried out, interviewed. Over and over again he checked his impressions, his tentative conclusions, his concepts, his categories, refining and clarifying and deepening and differentiating them. When he finally wrote his book, it was with a sense that he reached a reasonably stable point in the process
>
> (Reason, 1981:248)

In this research the researcher recycled the data emerging from the literature review with the staff at TELIP. The results were refined in the light of the subjects reaction.

Fourthly, Lather talks about catalytic validity. Catalytic validity represents the degree to which the research process reorients, focuses, and energises participants toward knowing reality in order to transform it. The argument for catalytic validity is premised not only within a recognition of the reality-altering impact of the research process, but also in the desire to consciously channel this impact so that respondents gain self-understanding and, ultimately, self-determination through research participation. My own participation in this research has led to a understanding of various research methods that could be used by TELIP. I am also clearer about the limitations and advantages of using the various methods.

TELIP staff who participated in the research also feel clear about the method that TELIP should use for impact evaluation. We have gained an understanding which was not there before this study. On a broader scale, through this study we will also be able to share our findings with other NGOs needing to do impact evaluation.

Maxwell (1990) states that understanding is a more fundamental concept for qualitative research than validity. Finally, Phillips (1980) reminds us that there are no procedures that will regularly (or always) yield either sound data or true conclusions. Bringberg and McGrath (1985:13) make the same point "validity is not a commodity that can be purchased with techniques..... Rather, validity is like integrity, character and quality, to be assessed relative to purposes and circumstances".

A more detailed description of the methods used in this research is provided below.

2.5. METHODS OF RESEARCH

Three methods of research were used in this study:

* literature review
* interviews
* workshops with TELIP staff

Literature Review

The literature review involved close reading of books, journals and impact assessment reports of INSET activities both international as well as South African. The literature was sourced through CD Roms of the Education Research Information Centre (Eric) system at the University of Witwatersrand. The library facilities at the Rand Afrikaans University as well as the University of South Africa were consulted.

I chose to do a literature review to gain an in-depth understanding of different research methods so that an appropriate method could be selected for TELIP.

At the time of starting this research, in 1997, I already had an understanding of how to assess TELIP's outcomes on teachers which I set out in an evaluation report (Dawjee:1994). How these outcomes could be assessed on pupils required deeper understanding. I needed to come to grips with what to assess, how to assess and what limitations to take account of.

The literature helped to identify serious conceptual, measurement, organisational and political problems likely to be encountered in the process of designing and implementing an assessment program. Identifying some of these pitfalls was to help TELIP to choose a method of assessment best suited to its needs.

As I read the literature I first took notes at random. I then interrogated the notes with questions such as 'what are the themes emerging from the notes?' This helped me to come up with the following thematic categories which I used to analyse the literature, to organise my thoughts and to make inferences.

* Different methods of research;
* Strengths and weaknesses of the different methods;
* Definitional issues: What is to be assessed ? Knowledge outcomes, skills outcomes, attitudes outcomes?
* Organisational and implementation issues; school environment and what makes it supportive or non-supportive to implement new ideas;
* Methodological issues; different assessment techniques for observations, tests, portfolios, oral tests, control groups;
* Implications for TELIP.

These categories were presented in a seminar to my peers in the M. Ed. group in September 1997 for comment. The M. Ed. is a course work plus research degree offered by the faculty of Education, Wits University. The responses of my peers helped to confirm and refine the categories.

Interviews

Altogether seven people plus two groups of teachers were interviewed. Interviewees were chosen because of their expertise and experience in evaluation.

* Professor D. Russell, Director of the Centre for Continuing Education in which TELIP is housed, was interviewed because he has a long standing connection with TELIP and also for his extensive experience of project management and evaluation;

* Ms Penny Vinjevold, Evaluation Manager at the Joint Education Trust (JET), a major funder of educational research, development and delivery in South Africa, is involved in developing the capacity of NGOs to conduct impact evaluation and as such had valuable contributions to the study. Ms Vinjevold had been involved in TELIP's impact evaluation on teachers and understands TELIP well;

* Professor J. Jansen, Dean of the Faculty of Education at the University of Durban Westville, has extensive experience in conducting evaluation of NGO projects. He has evaluated at least twelve programmes delivered by NGOs with the specified goal of improving the development of in-service teachers in science, mathematics and language education (mainly English). He has also presented papers at national and international conferences providing an overview of assessment in South Africa. His suggestions proved invaluable to this study;

* Mr.Paul Musker, an independent education consultant who works for the Gauteng Department of Education was interviewed as he has extensive experience in evaluation and knowledge of English language teaching. He also held the position of Director at the English Language Teachers Information Centre for Teachers (ELTIC). ELTIC is a non-governmental, education trust with a long history of providing a range of services for teachers. ELTIC is committed to effective language

learning, and the effective use of languages for learning as well as supporting educational development and the educational rights of all learners through teacher education. As Director of ELTIC Paul Musker has an understanding of INSET provision and evaluation in the South African context.

* Liz Johansen who was a co-ordinator of TELIP and also one of the writers of the TELIP courses. Liz understands TELIP and has also grappled with the issue of assessing impact on pupils.

The Director of ELTIC in Johannesburg and the Director of ELET (English Language Education Trust) which is an organisation dedicated to the development of enhanced English teaching and learning in Durban, were also interviewed to find out what kind of evaluations they had conducted and also to obtain suggestions for TELIP.

Two groups of teachers were also interviewed because I believe that teachers who go through the TELIP courses and who use the skills learnt in their classroom have "hands on" experience to make suggestions for impact assessment. The two groups of teachers were based in the rural areas of Mmabatho and Taung in the North West Province. The first group was made of six and the second group of ten teachers. The majority of these teachers have Std. Ten and a Primary Teacher's Certificate (PTC). Only one teacher had completed a BA degree, through the University of South Africa (UNISA). Most of them are however involved in further studies to improve their qualifications. None of them had any formal experience of research but responded from 'hands on' teaching perspective. This profile is typical of TELIP learners and the two groups of teachers form a representative sample.

How and When Interviews were Conducted

The table below indicates when and where the interviews were carried out as well as how data were recorded:

Table One

Person Interviewed	Date of Interview	Place of Interview	Method of Recording Interview
Penny Vinjevold	20:03:98	Johannesburg, at Joint Education Trust offices	Notes by hand and tape recording
Paul Musker	02:02:98	Johannesburg, University of the Witwatersrand, Centre for Continuing Education	Notes by hand and tape recording
Liz Johansen	03:02:98	Johannesburg, University of the Witwatersrand, Centre for Continuing Education	Notes by hand and tape recording
Professor Jansen	13:02:98	Durban, University of Durban Westville , Education Department	Notes by hand and tape recording
Director of ELTIC	16:01:98	Johannesburg, ELTIC offices De Korte Street Braamfontein	Notes by hand and tape recording

Director of ELET	13; 02: 98	Durban ELET offices Aliwal Street	Notes by hand and tape recording
Teachers	07:07:98	North West Province Mbatho High School Taung College of Education	Notes by hand

Rationale for Interviews

I chose to conduct interviews for this research because significant stakeholders within and outside TELIP have relevant ideas and experience around impact assessment and how it should be conducted. Through interviews, I was able to access their arguments in favour of, and opposed to, different methods of assessment. Interviews with teachers provided a perspective from their experiences of implementing TELIP in the classroom. This helped me to view the literature with a critical eye and also to work out how the ideas would apply to TELIP.

Another reason for choosing interviews is that they are adaptive. As I proceeded with each interview, I was able to modify the questions, to choose an area to probe, or to make changes to adapt the interview to the situation. In my situation this was an advantage as it helped me to probe and obtain a deeper and clearer understanding of the complexity of impact assessment strategies.

Bell, puts this succinctly;

> A skilful interviewer can follow up ideas, probe responses and investigate motives and feelings, which the questionnaire can never do. The way in which a response is made (the tone of voice, facial expression, hesitation, etc.) can provide information that a written response would conceal. Questionnaire responses have to be taken at face value, but a response in an interview can be developed and clarified.
>
> (1993 : 91)

Problems with Interviews

Interviews also have problems. They are time consuming and often costly. Interviewing is a highly subjective technique and therefore there is always the danger of bias. Many factors can influence responses, one way or another. Borg draws our attention to a few of the problems that may occur:

> Eagerness of the respondent to please the interviewer, a vague antagonism that sometimes arises between interviewer and respondent, or the tendency of the interviewer to seek out the answers that support his preconceived notions are but few of the factors that may contribute to biasing of data obtained from the interview. These factors are called response effect by survey researchers.
>
> (1981:87)

It is easier to acknowledge the fact that bias can creep in than to eliminate it altogether. Gavron, who carried out research into the position and opportunities of young mothers, was very conscious of the dangers inherent in research by solitary interviewers. She wrote, "it is difficult to see how this (i.e. bias) can be avoided completely, but awareness of the problem plus constant self - control can help" (1966:159). I attempted to keep my own biases in check by not imposing my own views while probing for interviewees' responses.

All the interviews were carried out in the months of February, March and July 1998 and ranged from 45 minutes to one hour in duration. At the beginning of each interview, its purpose was explained, background to TELIP was provided, anonymity of responses assured, and a copy of the report, once completed, was promised to each person participating in the interviews (as advocated by Nielsen and Reinharz ,1992). Interviewees were informed that they could refuse to answer any questions they wished, as well as to ask me for clarification. In this way I was able to establish trust and respect during interviews. People who agreed to be interviewed were given consideration and I fitted in with their plans however inconvenient it might have been for me. The venue was also negotiated with them and I provided as much prior information concerning the study as well as the TELIP programme, as the interviewee required.

I used an interview guide to make sure key areas were explored (see Appendix A). An interview guide is not a structured schedule or protocol. Rather, it is a list of general areas to cover with each informant. In the interview situation I was able to decide how to phrase questions and when to ask and probe them.

The key areas of the interview guide were:

* Which aspects of TELIP's INSET programme could be measured on pupils?
* Why these aspects?
* How to measure these aspects?
* Advantages of assessment strategies suggested?
* Disadvantages of strategies?

For teachers and NGOs, I used a different schedule (see Appendices B and C). The reason for this is that with NGOs I had to first establish whether they had carried out any evaluations of their own projects and what lessons could be learnt from these before asking for their suggestions about TELIP. With teachers I had to first establish whether they used skills learnt in TELIP in their classes, then probe into how they think pupils benefited or did not benefit before asking them to think about how TELIP could assess this.

Analysing Interviews

The interviews were recorded in two ways. I took notes and with the permission of interviewees tape recorded the interview. Tape recordings helped me to listen several times while doing content analysis to identify categories. It also helped when I needed to use direct quotations. Before using direct quotations I checked with respondents by faxing statements to them to verify the statements.

Each interview was transcribed verbatim. After all the interviews were completed I read the transcripts several times. I then compared responses across each of the key areas listed above. Certain themes emerged from reading the transcripts. These I noted on a chart. I shared my initial interpretations with a colleague so that I did not rely on my own interpretation only. This increased construct validity. (See Appendices D and E for an analysis of interviews)

The ideas emerging from the interviews and workshops were organised in the following manner:

* suggested methods of research
* problems and limitations with methods
* criteria for assessment
* contextual factors impacting on teaching and learning

Workshops with TELIP Staff

Rationale for Workshops

One reason for having workshops with TELIP staff was to present to them the research methods emerging from the literature review and interviews so as to weigh up the advantages and disadvantages of these in terms of an assessment model for TELIP. These discussions helped to gain a common understanding amongst staff and also to decide jointly on the best model for TELIP.

Another reason for the workshops was to discuss the criteria that might be used to assess impact on pupils. TELIP has clear outcomes of its programme for teachers but the project needed to think through which of theses outcomes could be used to assess the pupils that the teachers teach. Alongside these criteria the project also had to agree on instruments that would be used to obtain baseline data.

Workshops held

In the first workshop held with staff on the 31st of July 1997, the options emerging from the literature review were discussed. The second workshop was held on the 16 March 1998 to present ideas emerging from the interviews and to confirm the model of assessment. The main points discussed were:

* the model of assessment emerging from the interviews
* limitations/ problems with model
* criteria for assessment on pupils

* instrument to collect baseline data
* criteria for choosing the sample for study
* who should do the impa study and the cost implications.

2.6. STRENGTHS AND WEAKNESSES OF OVERALL DESIGN

The strength of the overall design lies in the use of the multi-method approach which provided scope for checking whether the data coming from the different sources was similar and consistent or not. My own analysis of interviews was compared with an analysis of the same interviews done by a colleague. The recycling of information from the literature review and interviews to participants in workshops helped to refine findings from the perspective of the participants. Most importantly the research design helped me and other TELIP staff to gain an understanding of different research methods.

One limitation of the design is that there were no language impact evaluation reports available which could be read as part of the literature review. General impact evaluation reports were read for inferences.

2.7. CONCLUSION

Torbert (1981, in Reason, 1988) suggests that the reason why neither current practice nor current research helps us to identify and move towards good educational practice is that both are based on a model of reality that emphasises unilateral control for gaining information from, or having effects on, others. Research in business, government and educational institutions shows that administrators in all fields choose, without question, behavioural strategies which seek to maximise their unilateral control over situations (Argyris 1969, 1971, 1974). Both in research and in organisational practice, the effort at unilateral control presumes that the initial actor (whether researcher or practitioner) knows what is significant from the outset and that this knowledge is to be put to the service of controlling the situation outside the actor in order to implement the pre-defined design as efficiently as possible. If students, subordinates, or research subjects seek to question whether there isn't something more significant at stake in the first place, the initial actor tends to redouble the effort to control the situation unilaterally. If she or he fails to do so, she or he tends to regard the

effort as a failure and the situation as "out of control." If everyone in a given situation acts in accord with the model of unilateral control, then no-one is open to learning new strategies or to examining their own assumptions.

I feel that there cannot be a contradiction between what the researcher is researching and how research is conducted. I believe that if learner centred learning environments are encouraged and valued in adult classes, then research must mirror these principles. This research was empowering (to a greater or lesser extent) for all participants, including the researcher. It is not just educational because the subject matter is education. It incorporates the humanist principles of respect, sharing, growth in learning, providing a foundation for a very transparent, fairly participatory, very focused research design which was practically feasible and effective.

The next chapter describes the results obtained from the research design.

CHAPTER THREE

THE RESULTS OF THE STUDY

3.1. INTRODUCTION.

The previous section spelt out the research design, provided a rationale for the choice of research methods and explained how the study was conducted. This chapter presents the results of the literature review, interviews, and workshops held with TELIP staff.

The results of the literature review are foregrounded in this chapter, as an examination of literature was the major research tool. The strengths and weaknesses of each method as they emerged from the literature review are supported or challenged with data from the interviews and workshops. The literature review and interviews proposed the following four research methods which could be used for impact assessment.

* Experimental Model
* Longitudinal Study
* Action Research
* Ethnographic Research

The table below summarises the main arguments for and against and against each research method as noted from the literature review, interviews and workshops. It also provides an indication of the support each method received as a model of impact assessment for TELIP. The table is followed by a detailed critical account of each method and a final weighing up of each method's feasibility for TELIP.

Table Two

Research Method	Supported by authors/interviewees	Main arguments in favour	Main arguments against	Potential usefulness for TELIP
1. Experimental model	*all authors reviewed supported this model * two out of the seven interviewees supported this model	*the researcher is able to decide the nature of the treatment, and who should be subjected to it * after the treatment is applied for an appropriate length of time, researchers measure the groups receiving different treatments to see if they differ. If the average scores of the two groups differ, and researchers cannot find any sensible explanation for the difference, they can conclude that the treatment had an effect and is likely the cause of the difference. For this difference, they may conclude that the treatment had an effect and is the cause of the difference.	*control is achieved only in restrictive and artificial settings *humans react to artificially restricted, manipulated conditions differently from the way they react to naturally occurring conditions. If the research is conducted under artificial conditions then the generalisibity of the results is severely limited. *control and experimental groups have to be selected carefully to ensure equality. That means the groups must have equal facilities and resources, receive equal care and attention, be taught by equally good teachers and so forth.	*TELIP will have difficulty finding a matching control group * The ethical issue of not providing treatment to the control group even though they may request it.
2. Longitudinal Study	*all authors reviewed supported this model. *five out of the seven	*the focus is on development of understanding and competence over a period of time.	*it is time consuming. By the time results are obtained initial participants may no longer be associated with the project.	*this model will provide an opportunity to do an in depth study tracing the progress of the pupils in relation to skills learnt

Research Method	Supported by authors/interviewees	Main arguments in favour	Main arguments against	Potential usefulness for TELIP
	interviewees supported this model	*pupils' progress can be measured against their previous achievement rather than against a norm. *within a longitudinal study, skills, processes, as well as knowledge can be assessed by means of student projects, portfolios and performance assessment.	*individuals change addresses and occupations and become difficult to locate. *costs of collecting data can be high *the researcher cannot be sure that behaviour change occurs because of the treatment. It is possible that some other variable brought about the change or even that the change would have occurred naturally without any treatment.	and put into practice. *it provides an opportunity for monitoring pupils' written and spoken English according to the criteria set by the project. *it eliminates the need to find a matching control group *it helps to eliminate many of the problems of contextual variables in experimental design
3. Action research	*all authors reviewed supported this model with caution. *none of the interviewees supported this model	*focus is on practical problems or areas that need attention and through a systematic cycle of critical reflection works towards improving the situation. *it aims to draw theory from practice. *teaching becomes an inquiry in which the teacher constantly endeavours to evaluate and improve the process of education for	*teachers need an atmosphere in which they are free to identify problems for inquiry, experiment with solutions and express and share ideas with colleagues and administrators. *the above context is difficult to find in most black schools in South Africa. *it is time consuming *there is more action than research	*TELIP could use it to help teachers improve their own practice rather than provide evidence of impact on pupils. *teachers will have to be trained and provided with on going support which places financial and human resource burdens on the project.

Research Method	Supported by authors/interviewees	Main arguments in favour	Main arguments against	Potential usefulness for TELIP
		herself and the people in her care.	*difficult to find an end	
4. Ethnographic Research	*all authors reviewed supported this model *one interviewee and two groups of teachers supported this model	*it involves extensive field work. The researcher engages with a site, a group of people, an institution, or a field to collect data. *fieldwork involves becoming familiar, through observation and interviews, with the phenomenon under study.		
		*the researcher studies as much of the problem as possible for as long as possible, interpreting, explaining, hypothesising or theorising about the phenomenon. *it makes provision for natural settings to be investigated. *it does not entail extensive pre field work design *the strategy and even direction of the research can be changed	*participant observation is a time consuming and demanding technique. *the researcher must establish rapport and trust with a group and become familiar enough to gain insight into the meaning of their lives. *the researcher must be an observer, remaining as objective as possible while collecting information. *its nature focuses on single	*the problem for TELIP is that this model does not have a pre-test and post-test and therefore will not be convincing for funders who want concrete evidence of impact on pupils

Research Method	Supported by authors/interviewees	Main arguments in favour	Main arguments against	Potential usefulness for TELIP
		easily, in line with changing assessment of what is requires by the process of theory construction.	cases or at most on a limited setting of action so there are searching questions about using ethnographic data for scientific generalisation. as well as for policy formation and decision making	

The four research methods, Experimental Design, Longitudinal Model, Action research, and Ethnographic research will now be discussed in detail.

3.2. EXPERIMENTAL DESIGN

The 'experimental design' appears frequently in the literature on impact assessment in education. Authors described the method, gave examples of how it could be used and then highlighted problems with using this model in educational settings. This section first provides a critical account of the method drawing on the literature reviewed, highlights factors that threaten validity, provides examples of its use, and finally discusses observations made by interviewees and decisions taken in the workshops with TELIP staff.

Terenzini (1989) cautions us that research design is a series of compromises. Designs that increase the power of study in one area come almost invariably at the expense of some other aspect of the study. Whenever something is gained, something else is given away. The key to useful and psychometrically sound inquiry is to know what is being gained and what is being given away.

A common approach to assessment of change in pupils is the use of a successive cross section design, typically involving a control group and a treatment group. The control group are compared with the treatment group on some measure of the variable/s on which change is being studied. Observed differences are then taken as an indication of the effects of the programme on pupils.

Fraenkel (1990) talks about the uniqueness of experimental designs. Experimental research is unique in two very important respects. It is the only type of research that directly attempts to influence a particular variable, and it is the only type that can really test hypotheses about cause and effect relationships. In an experimental study, researchers look at the effects of at least one independent variable on one or more dependent variables. The independent variable in experimental research is also frequently referred to as the experimental or treatment variable. The dependent variable, also known as the criterion or outcome variable, refers to the results or outcomes of the study.

Fraenkel highlights a major characteristic of experimental research which distinguishes it from all other types of research. According to him, in this research the researcher is able to manipulate the independent variable and decide the nature of the treatment, that is, what is going to happen to the subject of the study, to whom it is to be applied, and to what extent. After the treatment has been administered for an appropriate length of time, researchers observe or measure the groups receiving different treatments (by means of a post-test of some sort) to see if they differ. Another way of saying this is that researchers want to see if the treatment made a difference. If the average scores of the two groups on the post-test differ, and researchers cannot find any sensible alternative explanations for this difference, they can conclude that the treatment had an effect and is likely the cause of the difference. Experimental research, therefore, enables researchers to go beyond description and prediction, beyond the identification of relationships, to at least a partial determination of what causes them. Correlational studies may demonstrate a strong relationship between socio-economic level and academic achievement, for instance, but they cannot demonstrate that improving socio-economic level will necessarily improve achievement. Only experimental research has this capability. Such designs have a number of limitations. One of the limitations is that they take for granted that pupils are similar and that the variables that influence learning are also the same.

McMillan and Schumacher (1993) believe that experiments are a simple way of learning something by varying some condition and observing the effect on something else. As humans we use natural experiments constantly to learn. Young children experiment with a host of tactics to see which one will affect mom or dad most; teachers try a new approach to discipline to see if it works; and students vary study techniques to see which ones seem to result in the best grades. This simple trial and error behaviour is an attempt to show causation, which is the primary purpose of an experiment. The difference between these experiments and highly sophisticated experiments conducted by laboratory scientists is the extent to which the experimenter can be certain that the varied conditions caused the observed effect. It is the interpretation of causation, then, that is a key element in experimental research.

A true experiment according to McMillan and Schumacher is a procedure for investigating cause and effect relationships by randomly assigning subjects to groups in which one or more independent variable is manipulated. They also believe that a key element in experimental research is that the investigator deliberately sets up conditions in which different groups of subjects have different experiences. Taylor cautions us about the use of experimental designs when he says:

> once a suitable measure of success has been identified, assessing whether change has occurred requires a comparison. A surprising number of evaluators not only managed to discern change on the strength of one off observations of lessons with a limited number of teac' s, but also considered their method rigorous enough to attribute this change to the programme under evaluation. Such findings cannot be accepted as valid. Discerning change and attributing it to a specific intervention are the most problematic aspects of evaluation, requiring particular care in design, instrumentation and data testing. Controlled comparison is a sine qua non.
>
> (1995:8)

McMillan and Schumacher (1993) bring to our attention some of the problems in experimental research. According to them control is most easily achieved with research on humans only in restrictive and artificial settings. This is a weakness in education for two reasons. Humans react to artificially restricted, manipulated conditions differently from the way they react to naturally occurring conditions, and if the research is conducted under artificial conditions, then the generalisability of the results (external validity) is severely limited. The researcher would need to select the variables most likely to affect achievement, such as aptitude, time of day, and composition of groups, and control these as well as possible. This approach makes it more difficult to show that one or the other method of teaching is more effective, but the results are more generalisable to classrooms. The real challenge is in designing the procedures so that the results obtained can be reasonably generalised to other people and environments - that is balancing internal and external validity in a design.

Critics of experimental research for example Reason and Rowan, Carr and Kemmis, Guba and Lincoln believe that it is not appropriate for all educational research: experimental

research is appropriate only for some investigations seeking knowledge about e and effect relationships.

For many educational problems the experimental method would be inappropriate, such as descriptive studies (for example, "What is the attitude or level of achievement?") or studies of relationship (for example, "Is there a relationship between age and self-concept?"). In some situations an ethnographic approach would be more valid for explaining events, and in evaluation studies, experiments are frequently used with other approaches to investigate questions about a single practice. They represent a number of different epistemological positions which are loosely described as "post positivist" or "post structuralist" research favouring qualitative methods.

Cohen and Manion (1994) caution us about the number of factors that provide threats to internal validity in experimental research. Some factors that they highlight are :

* ***History.*** Frequently in educational research, events other than the experimental treatments occur during the time between pre-test and post-test observations. Such events produce effects that can mistakenly be attributed to differences in treatment.

* ***Maturation.*** Between any two observations subjects change in a variety of ways. Such changes can produce differences that are independent of the experimental treatments. The problem of maturation is more acute in protracted educational studies than in brief laboratory experiments.

* ***Statistical Regression.*** Regression means, simply, that subjects scoring highest on a pre-test are likely to score relatively lower on a post-test ; conversely, those scoring lowest on a pre-test are likely to score relatively higher on a post-test. In short, in pre- test, post-test situations there is regression to the mean. Regression effects can lead the educational researcher mistakenly to attribute post-test gains and losses to low scoring and high scoring respectively.

* ***Testing.*** Pre-tests at the beginning of experiments can produce effects other than those due to the experimental treatments. Such effects can include sensitising subjects to the true purposes of the experiment and practice effects which produce higher scores on post-test measures.

* ***Instrumentation.*** Unreliable tests or instruments can introduce serious errors into experiments. With human observers or judges, error can result from changes in their skills and levels of concentration over the course of the experiment.

* ***Selection.*** Bias may be introduced as a result of differences in the selection of subjects for the comparison groups or when intact classes are employed as experimental or control groups. Selection bias, moreover, may interact with other factors (for example history, maturation,) to cloud even further the effects of the comparative treatments.

* ***Experimental mortality.*** The loss of subjects through dropout often occurs in long running experiments and may result in confounding the effects of experimental variables. For, whereas initially the groups may have been randomly selected, the residue that stays the course is likely to be different from the unbiased sample that began it.

Neuman (1991) says that in an experimental study the researcher compares two groups of students to determine the impact of completing a course. In order to be compared, the two groups must be similar in most respects except for taking the course. If the group that completed the course is also older than the group that did not, for example, the researcher cannot determine whether completing the course or being older accounts for differences between the groups.

Charles (1988) also emphasises that the control and experimental groups have to be selected carefully to ensure equality, but they must be treated in exactly the same way, except for the experimental variable introduced into the experimental group. That means the groups must

have equal facilities and resources, receive equal care and attention, be taught by equally good teachers, and so forth. Without equal treatment of the two groups (except for the experimental manipulation), it is not possible to say what caused any differences that might have resulted in the experimental group.

Bateson (1995) carried out an evaluation of the Science Education Project (SEP) and although he did not use an experimental design he found difficulty in finding matching non-SEP schools to carry out his survey. I looked at the Bateson report in detail because SEP, like TELIP, s a non-government organisation which works with teachers with the assumption that if the teachers skills are improved then pupils will benefit.

SEP's mission statement sums up what it does;

> SEP is a system for promoting innovation and change in science education in South Africa by encouraging an activity-based approach to science education in order to advance education in South Africa. It works with and serves the teachers and learners of science and does so by co-operating with communities, authorities, and other concerned groups.
>
> (Bateson,1995:11)

It carries out its mission through five major activities:

- developing and providing print curriculum materials/worksheets;
- developing simple, self-contained kits of apparatus, equipment, and materials for conducting practical work;
- providing in-service training to teachers;
- providing follow up direct assistance in classrooms; and
- involving itself in evaluation, research, and development projects.

Although the mission statement of SEP refers directly to effects on teachers, the obvious reason for desiring change in teachers and teaching is a corresponding impact on students Most funders are naturally concerned with SEP's effects on students. The evaluation conducted by Bateson in 1995 therefore included a component that looked at the effects of SEP on pupils. The key to looking at the effects of SEP on students was to find reliable, valid indicators of its impact. The indicators that were agreed on were: enrolment of students in

senior, elective science courses, pass rates for Std. six, seven and eight science courses, affective outcomes, and cognitive achievement. In order to obtain measures of these indicator variables, two data collection instruments were designed and utilised. Science Education Questionnaire which was designed to obtain information on enrolments and pass rates, and a Science Education Project Survey which was designed to measure affective and cognitive achievement. The latter instrument also contained three items that were designed to confirm the implementation or non - implementation of SEP in a particular classroom. For each SEP school on which data was collected a non- SEP school had to be selected to obtain the same data. The non-SEP school had to match the SEP school in every respect. Bateson experienced difficulty in finding samples for the controlled group to match the experimental group. As a result data were collected only on 196 SEP schools and 144 non-SEP schools. The results of the cognitive testing showed that students in the tested SEP schools significantly out performed their counterparts in non SEP schools although achievement was not as good as might be desired. Based on the results of the cognitive survey there was no doubt that the SEP program had significant impact on science education in South Africa.

Vinjevold, (1996) also found it difficult to match groups in her evaluation of the Northern Cape Primary School Workbook Pilot Project (NCWP). I looked at this evaluation in detail to determine whether the experimental post-test and control groups used to make comparisons would be a suitable model for TELIP. Below is a description of this evaluation together with its problems.

The stated aim of the Northern Cape Primary Workbook Pilot Project (NCWP) is to contribute to the transformation of education through the development and provision of high quality, low cost educational materials. To this end all Standard two, three and four pupils in the Northern Cape were provided with two workbooks in February 1996. The purpose of the study was to,

- provide profiles of a representative sample of primary schools and Std.2 teachers and pupils in the Northern Cape; and
- assess the impact of the introduction of 'The Learning Adventure' workbooks into Std. two classes.

To assess the impact of the Learning Adventure workbooks, a comparison of the results of the experimental post-test and control group scores were used to make conclusions.

It was difficult to find similar Std. two groups in the Northern Cape, especially because pupils come from varied socio economic backgrounds and speak so many different home languages.

Ultimately it was decided to use Std.three classes from 24 of the 60 sample schools. The rationale for this was that pupils in the same school provide a very close match in terms of socio-economic and educational background. The assumption made was that a minimum of school-based educational activity takes place between October and February of the following year and that therefore the control pupils could be considered to be at much the same stage at the beginning of their Std.three year as the sample Std. two pupils at the end of the year, that is, at the point of intervention. However, there is extensive literature which asserts that maturation alone leads to learning gains and therefore, four months could result in higher scores on academic tests. On the other hand, there are well documented studies which show a decay in the academic results of students after the summer break. That is, if you give a test at the end of the school term and then several months later after the summer vacation, the scores are lower the second time around. Despite these limitations it was felt that comparing the February results of the Std. three pupils with the October results of Std. two pupils of the same school (as opposed to the entire experimental and control population) would provide the best available indication of the impact of the intervention.

This design therefore intended to establish if there occurred

* significant change in the intervention or experimental group from the pre-test in February to the post-test in October,
* significantly better performance on October tests for the intervention group compared to February performance of the control group (Std. three classes).

Lipsey asserts that if the above pattern of results was established then the data would be consistent with the hypothesis that the intervention is effective. However, the evidence would be far from conclusive. There are many possible rival hypotheses and no effective way to exclude them. Thus the results would be 'weak evidence' of the effectiveness of the

intervention. However, Lipsey suggested that if one could rate the schools on how fully they implement the intervention and made use of the new books then one could supplement the pre-test post-test analysis by looking at the correlation between exposure and pre post-change.

The results of Vinjevold's (1996) study showed that the introduction of The Learning Adventure appears to have;

* impacted on pupil learning activities and habits such as working on their own, working with peers and increased educational activities at home.
* had a positive effect on pupil learning. In particular, the degree to which the books were used affected the post-test results.
* assisted teachers in that they provided teachers with accessible supplementary material, influenced their teaching practices, helped motivate pupils and engage their interest in educational material, and assisted teachers with planning and preparation.

In the interview with the researcher Vinjevold stated that " it was politically untenable that some classes should not receive the programme for the sake of an evaluation." This alerts us to the need to consider ethical issues when setting up an experimental study.

Cohen and Manion (1994) also caution us that in planning the design of the experiment, the researcher must take account of the population to which she wishes to generalise her results. This involves her in decisions over sample sizes and sampling methods. Sampling decisions are bound up with questions of funds, manpower and the amount of time available for experimentation. With problems of validity in mind, the researcher must select instruments, choose tests and decide upon appropriate methods of analysis. Before embarking upon the actual experiment, the researcher must pilot test the experimental procedures to identify possible snags in connection with any aspects of the investigation. During the experiment itself, the researcher must endeavour to follow tested and agreed on procedures to the letter. The standardisation of instructions, the exact timing of experimental sequences, the meticulous recording and checking of observations.

The above literature review provides an understanding of what is involved in experimental research together with its advantages and limitations, but what does this mean for TELIP and how did participants in the study view experimental research?

The experimental design was suggested for TELIP by two interviewees with the precaution of finding a matching control group as well as consideration of the ethical issue of not providing treatment to the control group even though they may request it. In the words of one interviewee "it is extremely important to have a closely matched control group in terms of profile as well as exposure to English. This is where TELIP could run into problems if experimental design is used."

Workshops run with TELIP staff did not favour this model for the same two reasons-that it would be difficult to find a matching control group which has equal facilities and resources, receive equal care and attention, and be taught by equally good teachers. TELIP staff also felt uncomfortable with not providing treatment to the control group for the sake of evaluation.

3. 3. LONGITUDINAL DESIGN

The Longitudinal design appeared most frequently in the literature on impact assessment in education and training and was strongly supported by interviewees as well as TELIP staff. This section first provides a critical account of the method, highlights problems with the method and then discusses observations made by interviewees as well as TELIP staff in terms of its applicability for TELIP.

According to Terenzini (1995) longitudinal designs are a frequently recommended alternative to the experimental design. In a longitudinal study the researcher measures the characteristics of a class in a variety of aspects and then after a period of time studies the group again and compares students with themselves, controlling for entering characteristics, that is variables that impact on learning in the context that is being studied. The focus is on development of understanding and competence over a period of time. A range of techniques can be used to build up a comprehensive picture of performance as well as a profile of

strengths and weaknesses. Pupils' progress can also be measured against their previous achievement rather than a norm. Progress records of this type can accompany pupils throughout their schooling.

Bhola (1979) believes that this is one of the simplest and most effective designs available for impact assessment. It can be used effectively where tests or records are kept on a regular basis so that students are not aware that any special data collecting is taking place.

A longitudinal study requires administering a pre-test and then a post-test to gather evidence of change that might occur. This requires careful construction of tests or using available standardised tests (Terenzini, 1989). Locally developed tests may be more carefully tailored to local purposes and educational objectives, but they are also less likely to have been tested rigorously (at least in the short run) and, consequently, are of unknown validity.

A disadvantage of testing is that respondents know that they are being studied and this may influence their responses in varying ways. Cohen and Manion (1994) mention the Hawthorn effect[4] which contaminate experimental treatments in educationnl research when subjects realise their role as guinea pigs.

For these reasons unobtrusive measures - ones that do not require a conscious response from the subject, can be highly useful as well as efficient. For example, one alternative to the intrusive testing of students is an analysis of their written work to make inferences about their writing.

Catherine Cross (1996) suggests that within a longitudinal study, skills, processes as well as knowledge could be assessed by means of student projects, portfolios and performance assessment. A project is an activity performed and assessed over time, through which the student demonstrates his or her exploration of a topic in a variety of ways, including

[4] When research subjects know that they are being observed, they perform well.

displays, oral presentation and detailed research. A portfolio is a collection of student's written work, also over time, which reveals student development and growth for example, as a writer, a mathematics problem solver, or a scientist. The portfolio allows the student to self-assess, and the teacher to assess writing developmentally; often teachers and students have conferences to discuss the development of these new skills and processes. The performance task, like the project, allows students to show what they know through hands-on activities, but performance tasks usually take place in a more formal setting. Students often work in groups: in science, for example, a group of students might be asked to observe and analyse similarities and differences in phenomena and then develop an hypothesis. As this is a group activity, it is difficult to trace individual performance to make conclusions. Within these broad categories of assessment we can place dozens of other assessment possibilities: role play and dramatic presentations, debates, integrated activities across subjects, and many different types of group assignments and oral presentations. The key is that the assessment process involves a way of accurately judging the skills under consideration. For example, an oral presentation as part of a project on geological features of the Gauteng region needs to be assessed in at least two ways: in light of the quality of the oral presentation as a speech, and the quality of the visual aids and accompanying written report. An assessment is not authentic unless the desired skills and processes are actually captured in some accurate way.

The main problem with longitudinal study is that it is time consuming. Johnson, (1977) says that longitudinal studies are far less common than surveys and experiments. The extended time period needed discourages investigators from using the approach. By the time results are obtained, the person who initiates the study may no longer be associated with the project. Longitudinal studies introduce problems in collecting information as the studies often involve a follow up of individuals. Individuals change addresses and occupations and become difficult to locate. Even when they are contacted, some may decline to supply further information. Costs of collecting data can be high. In addition, if changes take place in teaching approach or materials, the validity of the study becomes questionable.

A longitudinal study also requires careful planning. The investigator defines the purposes of the study and identifies test scores and other data to be collected. Schedules must be drawn

up for administering the tests and collecting other types of data. Responsibilities for administering and scoring tests and for updating student records must be assigned. Collecting, manipulating and storing longitudinal data, especially where many students are involved, can lead to management problems. While advance planning is needed, it is unlikely that a longitudinal study can be planned to the last detail. Flexibility may be needed in selecting tests and scheduling events. Conditions change and plans require modification.

Fraenkel, (1990) advocates that the basic approach of researchers using a single subject design is to expose the same subject to two conditions or phases. The first condition or period is the pre-treatment condition, typically called the baseline period. During the baseline period, the subject is observed for several sessions until it appears that his or her typical behaviour has been reliably determined. Then a treatment of some sort is introduced. During or following each administration of the treatment, the individual is again observed until the researcher can determine the effects of the treatment.

Typically, though not necessarily, a highly specific behaviour is taught during the intervention condition, with the instructor also serving as the data collector - usually by recording the number of correct responses (for example comments) or behaviours (for example looking at the teacher) given by the subject during a fixed number of trials. In this module the baseline measurements or observations are made repeatedly until the researcher feels stability has been established. The treatment is then introduced again and a series of measurement or observations are made during or after each administration of the treatment. If the behaviour of the subject improves during the treatment period, the effectiveness of the treatment is presumed.

The problem with this design is that the researcher does not know if any behaviour change occurs because of the treatment. It is possible that some other variable actually caused the change, or even that the change would have occurred naturally, without any treatment at all. For example, a pupil could speak English well because of exposure to the language outside of the classroom and not necessarily because the teacher is speaking the language with skill and confidence.

The other notable limitation of this technique is the likelihood of data collector bias (the individual who is giving the treatment also usually collects the data and may want to report that the treatment is producing positive effects).

There is also the possibility of instrumentation effect (the need for an extensive number of data collection periods can lead to changes in the administration conditions). To some extent, the welfare of the subject is placed second to obtaining clear research results.

The size of the sample can also be problematic. If the study is done on a small scale the question of generalisibility will arise. Charles, (1988) asserts that very small samples are likely to be biased - that is they fail to reflect the population accurately. If a sample of only five students out of a population of 1000 is selected, even one highly atypical student can give a distorted picture of the total population. The more students are included, the better the chances that typical and atypical students will be represented in the same proportions that exist in the population. For some research, samples of 1,000 or more are used. But samples in educational research are typically much smaller. Often, a well-selected sample of 30 individuals is adequate, since samples of that size can resemble the population. With less than 30, the chances of obtaining a biased sample increases rapidly.

What implications does the longitudinal study have for TELIP? The longitudinal study despite its limitations was strongly recommended by all interviewees in the study. In the words of one interviewee " a longitudinal study will help TELIP to make claims about pupils progress made over a period of at least two - three years."

Participants cautioned TELIP about the length of time that this study involves. Although the period of study may be long , according to them, interim reports will help to satisfy donors about the progress made or not made by pupils. Respondents suggested that TELIP work out exactly what will be assessed before embarking on this study. For baseline data, TELIP needs to look at what language testing instruments are available or to construct its own instrument. If TELIP constructs its own instrument then the instrument needs to be validated.

A longitudinal study also requires financial commitment for a period of at least three to five years. In the present climate in which funding is hardly available, such a commitment from funders is highly unlikely and therefore does place additional pressure on a study of this nature.

Participants also alerted TELIP to the problem of pupil drop out and the fact that the same teacher in most instances would not be teaching the same class.

TELIP staff felt comfortable with this model for the following reasons:

* it provides an opportunity to do an in-depth study tracing the progress of the pupils in relation to the skills learnt and put into practice by the teacher during and after the TELIP course.
* it provides an opportunity of monitoring pupils' written and spoken English according to the criteria set by the project.
* it eliminates the need to find a matching control group.

The longitudinal study would seem to satisfy donors' need for teacher and learner impact assessment with a greater focus on learners. It helps to eliminate many of the problems of c[illegible]tual variables presented in experimental design. It could make teachers aware of their own practice before and after the TELIP course. It would also push TELIP to support teachers in the classroom to use new skills. The longitudinal study would also give the opportunity to assess written, as well as oral English skills in a variety of ways.

3.4. ACTION RESEARCH

This section starts with a description of action research, discusses its philosophical position in INSET, highlights the conditions that make this research possible, and then discusses observations of the interviewees and TELIP staff.

The literature reviewed stressed the value of action research as enabling teachers to reflect and improve their practice. Kemmis (1992) states that "action research is participatory: it is research through which people work towards the improvement of their own practices (and

only secondarily on other people's practices)."

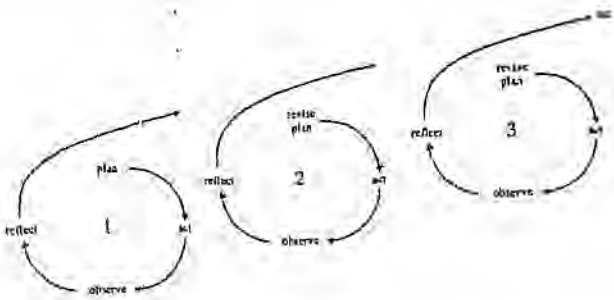
Action research, a term first used in the 1940s by Kurt Lewin, implies the application of tools and methods of social science to immediate, practical problems, with goals of contributing to theory and knowledge in the field of education and improving practice in schools (Kemmis, 1980). Collaborative action research suggests that each group represented in the process shares in the planning, implementation, and analysis of the research and that each contributes different expertise and a unique perspective. Collaborators often include school district personnel, university faculty or educational research and development centre staff, and national education agencies which provide financial support and guidance.

Action research projects have three general aims: staff development, improved sc' l practice and the modification and elaboration of theories of teaching and learning. Staff development through action research may take a number of forms, including increased teacher understanding of the classroom and school (Carr and Kemmis, 1986; Grundy and Kemmis, 1982; Nixon, 1981); increased self es'eem resulting from active involvement in research, professional conferences and perhaps publication (Elliot, 1985; Mc Cutheon, 1981; Sheard, 1981) and greater feelings of competence in solving problems and making decisions related to teaching and learning. Improved practice results from practitioner participation in the investigation of actions and issues of immediate importance. Contributions to educational theory include the discovery and elaboration of theoretical frameworks unde'lying teacher practice (Carr and Kemmis, 1986) and the development of theory grounded in the realities of the school and generalisable to other educational contexts. Although not every project aims at or meets all of these goals, most include elements of all three.

In terms of method, a self reflective spiral of cycles of planning, acting, observing, and reflecting is central to the action research approach. Kurt Lewin described the process in terms of planning, fact finding and execution. The basic cycle of activities is *Identifying a General Idea, Reconnaissance, General Planning, Developing the First Action Step, Implementing the First Action Step, Evaluation, Revising the General Plan.* From this basic cycle the researchers then spiral into *Developing the Second Step, Implementing, Evaluation,*

Revising General Plan, Developing the Third Action Step, Implementation, Evaluation and so on. The figure below illustrates these cycles.

Figure One



McNiff (1993) points out that action research regards the individual practioner as the centre of the research study. In philosophical terms, it is the development of knowledge of self, the integrity of the living "I" as the focus of educational enquiries. In educational terms, it is the concern by a practitioner to focus critically on areas that need attention and, through a systematic cycle of critical reflection in action, to work towards improving the situation. In the traditional pattern of INSET research, someone observes and describes teachers' classroom actions and gives advice on how they might be improved. McNiff refers to this as the E - (externalised) enquiry. According to McNiff an I - (internalised) enquiry is that conducted by the individual into her own practice. She or he reflects critically on her work, either privately or through discussion with others, and aims to think of original ways that will help her to improve.

In E-enquiries, the focus of the research is the practices of others. In I-enquiries the focus is the practice of the self. In E-enquiries the purpose of research is to observe, describe and explain what other people are doing. Its status is derivative - that is, the accounts given of the research are those of the recorder, but not always of the practitioner. The accounts themselves aim to offer explanations to others through an "objective" study of the data, to see if those data (facts about the study) fit the recorder's theory. In this conventional INSET research pattern, the observer has reasonab' - clear ideas about how a pedagogical situation ought to be; she or he watches the teacher, and advises the teacher on her or his action plans.

In the I-enquiry the purpose of the research is to explain what the practitioner is doing. Its status is personal. The accounts rendered are those of the practitioner, and aim to offer an externalisation of the practitioner's mental processes as she or he tries to bring about change; that is, it shows how a teacher was dissatisfied with personal practice, and why, and the steps taken to improve. The teacher's practice is an outcome of her or his thought, and improved practice is an outcome of improved understanding. Educational research aims to encourage the development of personal understanding that will lead to an improved form of practice. It becomes an enquiry by the self of the self; and, rather than aim to fit personal practice into another person's theory, it concerns itself with enabling individuals to develop their own personal theories. Seen from this perspective, the process of an enquiry in action aims to draw a theory out of practice. Contrary to the traditional form of INSET research, where theory acts as the basis for others' practices, this approach centres on an individual's understanding and sees practice as the ground for the development of the process of theorising.

McNiff (1993) observes that the act of teaching, in a traditional view, involves the passing on of skills and concepts to learners. This view is pedagogical rather than educational. Education is to do with the development of the process of individual rationality. In the "shifting centres" approach to the processes in classrooms and other workplaces, it is possible to see the teacher's job as facilitating the same critical awareness of personal practice to her or his pupils. It is a sort of chain reaction within a network. In this process the teacher is seeking to involve herself or himself in helping others to improve themselves.

This shifting centres model is based on the following assumptions:

* Teachers are regarded as experts, who empower themselves to offer accounts of their own practice. These accounts are to be legitimated through the validation of peers and clients. Teachers are encouraged actively to create out theories, and develop these personal theories through their accounts.

* Teachers, teacher-supporters and clients are awarded equal status and responsibility for helping the other person's process of understanding to evolve. In this collaborative view, all practitioners at all levels (learners, teachers, supporters) are

involved in the process of the development of their own, and each other's rationality: they are improving the quality of their own learning. Teaching and learning are interchangeable terms, existing as processes that regulate the interrelationships within a network of thinking practitioners.

- The model is personalised. The focus is the understanding by the individual of her own life. (The understanding of the self by the self). The intent is to improve the process of education within a particular present situation.

- The model is based on a 'process' view of learning. There is no end product in sight, other than an 'end product' of 'no product'; a final answer that there are only new questions; an end state that is the beginning of a host of new states.

- Research is seen as a form of teaching which explores new ways of life that promise to be beneficial to the community of which the researcher is a part. The act of teaching involves the concept of bringing about improvement. Seen in this light, self- reflective research not only provides a proper base for teaching (Rudduck and Hopkins, 1985), but also is teaching. Teaching becomes an "enquiry in action" in which the teacher constantly endeavours critically to evaluate and improve the process of education for herself and for the people in her care.

Successful collaborative action research depends on a project structure conducive to effective action research. A project structure conducive to effective action research consists of at least four elements: frequent and open communication among participants; democratic project leadership spiralling cycles of planning, acting, observing, and reflecting; and positive relationships with the school context within which the project occurs.

Communication

Hord (1981) and Wallat et al (1981) stress the importance of negotiating and articulating clear and specific goals from the outset of an action research project. Clear goals provide all participants with a sense of the project's value and what they will gain from it and help

establish a shared frame of reference from which hypotheses and future plans can be generated. Although defining mutual goals may consume a large part of the group's initial meeting time, this process provides the research group with a shared sense of commitment, mutual understanding and a framework for future tasks. Frequent interaction among participants in the research project, through team meetings and more informal discussions, is a requirement of action research which helps to overcome communication difficulties and contributes to mutual understanding of goals, techniques, and perspectives (Elliot,1985; Hord, 1981).

Leadership

Many of those studying action research also call for strong leadership in a collaborative action research project, by someone who can set a positive example as a collaborator (Ebbutt, 1985; Grundy and Kemmis, 1982). This often means that the leader must disperse his or her power, sharing control and allowing others to delegate and assume responsibility.

Spiralling Cycles

Lewin (1948) explains that action research proceeds through spiralling cycles of planning, execution, and reconnaissance (or fact - finding) in order to evaluate and perhaps modify the plan. Elliot (1981) and Kemmis and Mc Taggart (1982) adapt Lewin's description, each providing a model of the process of action research which emphasises recurring cycles of planning, acting, observing, reflecting, and revising. They, too, use a spiral pattern to indicate that initial ideas shift over time and that recurring reflection leads to modification of plans throughout the process. Grundy and Kemmis (1982) explain that spiralling cycles are necessary "to bring action research under the control of understanding, in order to develop an effective critique of the situation". Ebbutt (1985) sees the process of action research as a series of successive cycles, each incorporating the possibility for feedback of information within and between cycles. The emphasis remains the same, however; an action research project must provide participants with the opportunity to work through several cycles in order to be effective. This recursive rather than linear, research process allows practitioners to use their own reflections, understandings and developing theories to inform both practice and research. Action research projects must therefore be structured to allow this cyclical process

to occur.

School Context

Certain elements of the school environment contribute to the effectiveness of action research projects. Projects are most successful when the school climate encourages communication and experimentation and when the administration supports the project (with technical support and/or assurances of further implementation or continuation). Studies by Hyte (1986) and Cohn and Finch (1987) all indicate that the school context affects teacher's willingness and ability to participate in the process of action research. Crey (1952) suggests that teachers need an atmosphere in which they are free to identify problems for inquiry, experiment with solutions and express and share ideas with colleagues and administrators. Ideally, the administration not only provides teachers with the freedom to experiment, but also gives them the recognition needed to legitimise their project and ensure its continuation in the future (Cohn and Finch,1987).

Action research can exist without administrative or school support: teachers may work individually or in small groups to carry out projects in their classrooms (James and Ebbutt, 1981) or outsiders may provide the initiative, materials, and professional support needed to implement a project (Cohn and Finch, 1987; Whyte,1986). Most projects which engage in these more independent forms of action research later reflect that greater school/administration involvement might have helped expand the impact, longevity, and legitimacy of the project. As Carr and Kemmis (1986) note when describing a project in which they had worked:

> The situation did not change as radically as the teachers involved had hoped. But they learned something about the change process itself. That they needed to involve others in the learning process they had gone through, and to involve them early.
> (1986:170)

Another important element of Action Research is the role of the outsider. According to McNiff (1993) recent reports and analyses of action research have begun to examine the role of the 'outsider' in an action research project. Many of these studies note the problems inherent in having an outside researcher work with teachers.

First, teachers and researchers may use different languages, focus on different problems, and may therefore have trouble communicating. (Cummings and Hustler, 1986; Threadgold, 1985). Second, outsiders (particularly those from the university) tend to have higher status which can lead to intimidation or resentment. This may limit the group's ability to address necessary interpersonal and task demands (Cassidy,1986; Threadgold, 1985). Finally, because of their greater knowledge of research, their status, and the research framework they may bring to the project, outside researchers may have too much influence on what issues the group addresses, how they collect, analyse, and reflect on their data, and how their findings are used and reported (Carr and Kemmis, 1986). An outside researcher may, therefore, work against the democratic processes which lead to the exploration of teacher's theory and practice.

Despite these problems, many reports of action research see an outsider as an essential facilitator of the process and method, someone who serves many of the tasks and maintenance functions of leadership outlined in the literature of group dynamics. Hustler, Cassidy and Cuff write that "some form of dialogue with an outsider is not only desirable for action research, but almost one of its defining characteristics." (1986:15). They, and others, believe that the outsider researcher or an in-school person who plays a similar role, activates the process. Ebbutt (1985) notes that educational institutions often lack a 'collective dynamic' or an imperative for change. Someone must take the responsibility to initiate change, to activate the cycles of planning, acting and reflecting. Activation may require persuading teachers to become involved (Cassidy, 1986); asking questions which challenge the existing frameworks or theories that shape teachers' views and actions (Desforges, Cockburn and Bennet, 1986); and convening meetings to begin the group process (Ebbutt, 1985). Teachers are busy with their daily tasks; they are also fairly comfortable within established patterns of thinking and behaving in their schools and classrooms. An outsider can provide the impetus, energy and initial framework to question what is.

Second, the outside researcher may bring a variety of resources to the action research project that would not otherwise be available to the participants. These resources include time, specialised knowledge of research methods, and theoretical knowledge, which, if well-used,

can support teachers' developing understanding of their own practice (Elliot, 1985; Grundy and Kemmis, 1982; Nixon, 1981). Elliot explains:

> One of the facilitator's roles is to mediate theoretical resources in a way which enhances rather than constrains teachers' capacities to develop their own theoretical understandings. So long as the introduction of external ideas constitutes a support rather than a substitute for teachers' own thinking, it can speed up the process of aims clarification and consequently the process of problems identification, analysis, and the formulation of strategic action. Given this supportive rather than controlling context, the introduction of external ideas constitutes a support rather than a substitute for teachers' own thinking, it can speed up the process of aims clarification and consequently the process of problem identification, analysis, and the formulation of strategic action. The introduction of external ideas also gives teachers greater opportunities for making original contributions to the development of pedagogical theory.
>
> (1985:253)

Elliot found that when teachers were provided with theoretical literature before they had begun to examine their own concerns, they dismissed the theory as irrelevant. But when theory was provided after teachers had themselves begun to question, articulate, and experiment with their own understandings, the teachers found theory useful in helping to explain and analyse their own actions.

Third, the outside researcher may provide what Carr and Kemmis (1986) describe as a "sounding board against which practitioners may try out ideas and then learn more about the reasons for their own action". They also learn about process leadership that helps teachers define their concerns, plan strategies for change, observe the effects of change and reflect on their results. Carr and Kemmis believe that a model of shared leadership which requires no single leader is most congruent with the requirements of action research.

A fourth role for the outsider is that of organiser. He or she can co-ordinate the work of individual teachers, keep records of plans and meetings, negotiate relations between team members, other teachers, and the administration, and arrange for the dissemination of action research reports (Ebutt, 1985; Grundy and Kemmis, 1982; Nixon,1981). Although this may be a fairly technical role, it is one that teachers are often unable, because of time and scheduling constraints, and unwilling, because of school structures and hierarchies, to take

on. Co-ordination may also include helping the group develop a common language that reflects practitioners' concerns and assisting in the development of a group process in which power is distributed. "Since facilitators are usually regarded as 'expert' they must consciously encourage the group to take power for themselves," (Grundy and Kemmis, 1982:71).

An additional role that the outside researcher can play is one of developmental leader. An outside researcher can provide the opportunities and support necessary to stimulate teachers' development of new ways of inquiring, perceiving, and reflecting on their experiences. By recognising the developmental stages of participants, providing opportunities for them to take on more complex roles, and encouraging reflection on their actions, a developmental leader cultivates in participants an increased capacity for learning and understanding new knowledge and skills. Through these processes, the outside researcher contributes to an effective group process, establishes a norm of support for risk-taking and role changing, and broadens individual perspectives. A leader-as-developer approach requires that the leader have the ability to see and value other viewpoints and to be flexible in his or her responses in order to meet the developmental, professional, and context specific needs of the teachers involved (Grimmett, 1983; Joyce, 1980; Oja and Ham, 1987; Thies-Sprinthass,1984).

It is important for facilitators to recognise that their perceptions are filtered through their own developmental stages, exhibited in their interactions and relationships with teachers on the team. The university researcher working as a facilitator in collaborative action research must be aware of his or her own stage of development and its implications for dealing with the needs, cognitive styles, and concerns of teachers at various stages of development.

Kemmis (1992) argues that improving education means improving educational discourses, improving educational practices and improving forms of educational organisation. In concrete terms, this means changing people, their ideas, their activities and their social relationships. But "changing people" is extremely difficult to achieve, especially when the "people" are treated as "others", the objects of someone's plans for change, rather than as knowing subjects, willing and able to determine their own roles in the improvement process. Collaborative action research in education aims to establish groups of knowing subjects

committed to changing themselves and, by so doing, changing their educational work. It is about helping people to become more conscious and critical of their agency in processes of historical change - acting collaboratively as knowing subjects directing their efforts towards educational improvement. Action research is not only about 'doing', however; it is about learning by doing. It is about making changes, observing their consequences, improvement in the light of what has been learned through observation. Action research helps to monitor changes systematically, to promote continuing critical and self critical reflection which can inform our unfolding efforts towards improvement.

Cohen and Manion (1995) highlight some of the criticisms levelled against action research. According to them the points usually made are : that its objective is situational and specific (unlike the scientific method which goes beyond the solution of practical problems); its sample is restricted and unrepresentative; it has little or no control over independent variables; and its findings are not generalisable but generally restricted to the environment in which the research was carried out. While these criticisms hold in most cases, we need to remember that as action research programmes become more standardised, less personalised and more open, some of these strictures will become less valid.

What does action research offer as a model of impact assessment for TELIP? As mentioned in Chapter One, TELIP has had a long and beneficial experience of action research. This experience, however, was outside of the school context. The atmosphere in which tutors facilitated was relaxed and free from the bureaucratic constraints of school administration. Tutors were engaged in facilitation of TELIP courses without having to handle other tasks typical of a teacher in the classroom. All of this made it easy to conduct action research. In the workshop with TELIP staff on 16 March 1997 the assumptions and factors of action research in the school context were presented as they were understood from the literature review. Two important problems identified with the use of action research as an assessment model in schools were that:

* the teachers would have to be trained and provided with on-going support;
* the schools in which TELIP learners teach do not encourage communication and experimentation and their administrations do not support innovative projects. This

could affect teacher's willingness and capacity to participate in action research;

* the present climate in South African schools as explained in Chapter One places many burdens on teachers, and it does not seem feasible to provide further training for them.

Furthermore, pressure from funders to show concrete evidence of impact on pupils places emphasis on the product rather than the process which is the key in action research.

Interviewees in the study empha..ed their belief that action research helps the teacher to improve his or her own practice rather than provide data for impact assessment. One participant stated "action research is great, but it's primary purpose is self improvement - what I have achieved etc."

In view of the pressure from funders to provide concrete evidence of impact on pupils, TELIP needs to consider other methods of research.

3.5. ETHNOGRAPHIC RESEARCH

This section starts with a description of ethnographic research, the techniques associated with the method, then raises questions of validity and reliability, and finally presents observations made by interviewees and TELIP staff.

Jaeger (1988) points out that ethnographic research has been acknowledged and welcomed as a research strategy for inquiring into education in recent years. One often hears educators discussing "ethnography" or "ethnographic approach" today. The fact that educators use these terms does not of course assure that they have a clear sense of how ethnographers conduct their research or what ethnographic research shares in common with related approaches like participant observation studies, field studies, or case studies.

Over the last few years there has been growing interest in ethnography among researchers in many different fields especially for illuminating the practice of educating and training adults. According to Hamersley (1983) ethnography is simply one social research method, albeit a

somewhat unusual one, drawing as it does on a wide range of sources of information. The ethnographer participates overtly or covertly in people's daily lives for an extended period of time, asking questions, and collecting available data to throw light on the issues with which he or she is concerned.

Briefly the three characteristics of ethnography are: the development of, rather than the testing of hypotheses and theory; fieldwork (in more instances) as the usual means of collecting data; and reliance upon the investigator as the primary instrument for data collection and analysis.

According to Merriam and Simpson (1984), one important characteristic of this kind of research is that the researcher is the primary conduit for data collection and analysis. While surveys, tests or inventories might be used in support, the investigator assumes the major burden for collecting and processing data.

A second characteristic is that it involves extensive field work. The researcher physically goes to the site, the group of people, the institution, the "field", to collect data. Field work involves becoming intimately familiar, through observation and interviews with the phenomenon under study. The researcher goes into the field and studies as much of the problem as possible with the intent of interpreting, explaining, hypothesising, or theorising about the phenomenon.

Spradley (1980) views ethnographic inquiry as cyclical instead of linear. It begins with the selection of an ethnographic project and proceeds to asking ethnographic questions, followed by data collection and analysis. This cycle is repeated as research progresses, but the research questions are modified as the inquiry progresses from "descriptive" to "structural" questions. Descriptive questions are general and broad, appropriate particularly in the orientation phase, while structural questions are formulated in the phase of focused observation to concentrate on more articulately defined domains of inquiry. Unlike linear, quantitative research methods in social science, ethnography has no standardised procedures of investigation. Ethnographers often combine different methodological techniques, some of which are

devised personally, in such a way as to facilitate their data collection in particular field situations. The procedures commonly used in this type of investigation are participant observation, in-depth interviewing, life history, documentary analysis, and investigator diaries (records of the researcher's experiences and impressions). Employing any one of these procedures involves going into the field "immersing oneself in a collective way of life for the purpose of gaining first hand knowledge about some facet of it "(Shaffier et al., 1980). Fieldwork involves entering the chosen setting, establishing rapport with the residents of that setting, maintaining some type of relationship with the subjects, and finally, leaving the setting. Participant observation is the cornerstone technique of ethnography, and a researcher might assume any of several variations of this technique. Junker (1960) describes four variations:

* Complete participant. The researcher becomes a member of the group being studied, concealing the fact that he or she is observing as well as participating.
* Participant as observer. The observer's activities are not concealed but are clearly subordinated to activities as participant
* server as participant. The role of observer is publicly known, and participation becomes a secondary activity.
* Complete observer. The observer is invisible to the activity (as in the case of a one-way mirror or hidden camera) or tries to become unnoticed.

The role assumed depends upon the type of information being sought and the idiosyncrasies of the group being investigated. The researchers role can also shift during the process of the investigation.

The first observations that researchers do may be called "descriptive observation" (Spradley, 1980:32). The observation is governed by a general question, "What is going on here?" The intent of the observation is to gain insight into researchable problems when researchers have little knowledge of the culture to be studied. Researchers make field notes that provide a detailed, narrative description of what has been observed with particular attention to activities, actors, space, physical objects and the sequence of activities and events.

Subsequently, by reviewing their field notes, researchers identify and analyse domains of the phenomena observed. Spradley (1980) calls this process "domain analysis." The primary purpose of domain analysis is to identify categories - for example, activities and actors - and their relationships, leading to a taxonomy of domains. Categories should be flexible, so that they can be modified as the research proceeds. Domain analysis is helpful in outlining the general boundary of a problem and its internal dimensions. Guiding questions/ hypotheses develop continuously; they are especially useful in guidin,, research in its early stages. They are framed in general terms to allow other questions to evolve later.

The intensity and duration of participant observation varies with research projects. It usually involves ethnographic interviews in order to complement data gathered through observation (Spradley,1978). In most ethnographic studies, interviews are open-ended or loosely structured so that the respondents' views of the topic can be obtained. By using an open-ended format, investigators hope to avoid predetermining subjects' responses and, hence, their views of reality.

There are several types of interviews that can be employed in an investigation: team and panel interviewing, covert or overt interviewing, oral history interviewing, and structured and unstructured interviewing (Guba and Lincoln, 1981). In most ethnographic studies, interviews are open-ended or loosely structured so that the respondents' views of the topic can be obtained. By using an open-ended format, investigators hope to avoid predetermining subjects' responses and, hence, their "views" of reality.

Occasionally, interviewing key informants (people who have a great amount of knowledge and can conceptualise their group's norms and beliefs) leads to the collection of life histories. Life histories provide vivid integration of cultural information and are of great value for understanding particular ways of life. In addition ethnographers may also want to evaluate all available documents on the phenomenon being studied. Field workers are often encouraged to keep a diary of each day's happenings and record personal feelings, ideas, impressions, or insights with regard to those events. This diary becomes a source of data and

allows researchers to trace their own development and biases throughout the course of the investigation.

According to Hamersley and Atkinson (1983), the basic strength of ethnography is that people make sense of the world in everyday life. Some see this as a weakness and dismiss ethnographic research on the grounds that the data and findings it produces are merely subjective, idiosyncratic impressions that cannot provide a solid foundation for rigorous scientific analysis. Others argue that only through ethnography can the meanings that give form and content to social processes be understood. Experiments are rejected on the grounds that these are incapable of capturing the meaning of everyday human activities. Peoples' behaviour can only be understood in context. For this reason natural settings must be investigated. We cannot understand the social world by studying artificial simulations of it whether in experiments or interviews. To restrict the investigations of social behaviour to such settings is to discover only how people behave in experimental and interview situations.

Spindler (1982) raises questions around the generalisibility of ethnographic research. He says that although ethnographic research calls for a discipline that is as demanding as that required by any experimental design or correlational research strategy, its nature focuses on single cases, or at most on a limited setting of action, so there are searching questions about using ethnographic data for scientific generalisation, as well as for policy formation and decision making. Hamersley (1983) also points out that owing to the influence that a participant observer may have on the setting studied, and or the effects of temporal cycles within the setting, the conclusions he or she draws from the data are by no means necessarily valid for that setting at other times. For similar reasons, findings produced by participant observation in one setting may not be true of other settings of the same type.

Ethnographers, however, feel that an in-depth study that gives accurate knowledge of one setting not markedly dissimilar from other relevant settings is likely to be generalisable in substantial degree to these other settings. According to Spindler (1982) qualitative data obtained from ethnography can tell us what may be significant on a wider scale. A final answer to questions about generalisability is that correlational and experimental research

designs are also subject to severe strictures. They also sample universes. And if these designs fail to take context and meaning into adequate consideration, their results may be significantly less generalisable than the results from a good ethnographic study.

Sherman and Webb (1990) also raise the question of validity and reliability of ethnographic research. According to them, ethnographic validity refers to the degree to which participant observation achieves what it purports to discover, i.e. the authentic representation of what is happening in a social situation. Ethnographic reliability refers to the repeatability of a given study by researchers other than the original participant observer; the extent to which independent researchers discover the same phenomena in comparable situations (Le Compte and Goetz,1982). It is the reliability of ethnographic research and validity to a degree, that more often than not is criticised.

In participant observation, validity indicates the ethnographer's understanding of the meanings of the observed socio-cultural experience. Therefore, to attain a high degree of validity, the observer must repeat observation through prolonged involvement in the culture studied. It is because of this that establishing residence among the participants is an essential factor in research. The participant observer must maintain neutrality and have a multi-dimensional vision for both observing and interacting with research participants. This suggests that the researcher has to focus on the entire group and a wide range of informants (a stratified sample, if possible) representing the group. Also, the observer must guard against ethnocentrism and perceptual biases. Still another strategy is to involve participants in reconstructing the events recorded by the observer. This contributes to reducing the researcher's perceptual biases.

According to Hammersley (1983), a considerable advantage of ethnographic research is its flexibility. Since it does not entail extensive pre-fieldwork design as social surveys and experiments generally do, the strategy and even direction of the research can be changed relatively easily in line with changing assessment of what is required by the process of theory construction. As a result, ideas can be tried out quickly and if promising, followed up. In this way ethnography allows theory development to be pursued in a highly effective and

economical manner. However the contribution of ethnography is not limited to the phase of theory development. It can often be used to test theory. For example, cases that are crucial for a theory - those where it seems most likely to be proved false - may be examined through ethnography (though this is not always feasible for macro-social theory where the scale of the object under investigation often necessitates survey research). While the fact that, unlike in the experiment, variables cannot be physically manipulated hampers the evaluation of competing hypotheses, it does not rule it out. Moreover, what is lost in terms of the control of variables may be compensated by reduced risk of ecological invalidity.

Since ethnography investigates social processes in everyday settings rather than in those set up for the purposes of research, the danger that the findings will apply only to a particular artificial situation is lessened. In addition, ethnography's use of multiple data sources is a great advantage here. This avoids the risks that stem from reliance on a single kind of data: the possibility that one's findings are method dependent. The multi-stranded character of ethnography provides the basis for triangulation in which data of different kinds can be systematically compared. This is an effective manner in which reactivity and other threats to validity can be handled.

Merriam and Simpson (1982) point out that participant observation is a time consuming and demanding technique. The researcher must establish rapport and trust with a group and become familiar enough to gain insights into the meaning of their lives. At the same time, the researcher must be an observer remaining as objective as possible while collecting information.

Finally the techniques of ethnography - participant observation, and interviewing in particular have been adopted by people in applied fields of study where research problems or questions have warranted an exploratory, rather than hypothesis-testing approach. However, anthropologists take issue with educators and others who use ethnographic techniques and then think they are doing ethnography. For ethnography is more than techniques; it is also an account of the data, an account that interprets the data within a socio cultural framework. Ethnography has as its intent the interpretation of a situation that incorporates the

participants' symbolic meanings and ongoing patterns of social interaction. Concern with the cultural context is what distinguishes ethnography from grounded theory, which builds theory, and case study, which describes and interprets a situation or social unit from the perspective of the researcher. Wolcott makes this distinction between technique and account:

> Specific ethnographic techniques are freely available to any researcher who wants to approach a problem or setting descriptively. It is the essential anthropological concern for cultural context that distinguishes ethnographic method from fieldwork techniques and makes genuine ethnography distinct from other "on-site observer" approaches. And when cultural interpretation is the goal, the ethnographer must be thinking like an anthropologist, not just looking like one.
>
> (1980:50)

How can TELIP use ethnographic research for impact assessment on pupils? It is interesting that only one participant in the study strongly believed that TELIP should use ethnographic research for impact assessment. This participant felt that language should not be measured quantitatively. There should be no baseline test but the way that learners use language (both written and spoken English) should be observed in detail and conclusions about impact made from this in-depth observation. The two groups of teachers interviewed also seemed to recommend this approach to impact evaluation when they said "come and observe us in the class, interview the students and the Heads of Department."

Other participants in the study felt strongly that if no base line data were gathered then conclusions made about impact would not hold water. TELIP staff have a strong leaning towards ethnographic research. This technique was used to observe how teachers use English language in their classrooms in an internal Impact Evaluation carried out in 1994. Yet funders who read the evaluation felt that baseline data was needed to convince them of the effects of the TELIP courses.

3.6. SUMMARISING THE FINDINGS

The study started with the aim of finding a suitable model for impact assessment of TELIP courses specifically and of INSET courses generally. As noted in this chapter the literature review suggested four methods for consideration; experimental design, longitudinal study,

action research, and ethnographic research for consideration.

Method is only one of the considerations in a research design. Other factors that need to be considered in research design are;

* the philosophy of the researcher which influences the design
* the philosophy of the organisation implementing the research
* outside pressure placed by funders
* philosophy of funders
* availability of financial and human resources.

The next chapter provides a discussion of the findings and comes to a conclusion with regards to the model of assessment best suited to TELIP.

CHAPTER 4

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

This final chapter reflects on the aim of the study, the rationale for conducting the study, and the appropriateness and efficacy of the research design. It also discusses the research findings in terms of their implications for TELIP, and provides a plan for an impact assessment of TELIP courses.

4.1. AIM OF THE STUDY

This study aimed to;

(a) investigate the advantages and disadvantages of different research methods

(b) analyse these methods to see how these can be applied to impact assessment for TELIP specifically and generally to other NGOs with similar orientation.

These aims were met.

4.2. RATIONALE

TELIP courses are designed to raise teacher's level of competence in English by concentrating on the major linguistic problem areas which have been diagnosed by TELIP and which the teachers themselves have indicated. Some of the linguistic problems faced by the teachers were:

* lack of English oral communication skills.
* lack of English reading skills.
* lack of English writing skills.

From TELIP's inception, the intention was that pupils should benefit from the greater linguistic competence of their teachers. However, this was not clearly explained in project documents. Instead, the focus of TELIP's work has been on improving teacher's language skills. Internal evaluations have concentrated on the impact that the courses have on teacher's professional growth. (see for example, Dawjee (1994) - Phase1: Impact evaluation of courses on teachers).

TELIP has undertaken both formative and summative evaluations of its courses. A formative evaluation of the programme was carried out in the very process of course construction. The emphasis in this phase was on the teaching process, and the information generated was used to improve the curriculum during implementation. Summative evaluation was carried out at the end of a lengthy curriculum development phase in 1994. These evaluations did not address the question of impact on pupils. The question facing TELIP is what aspects of its training lend themselves to pupil assessment, and how best to do it? The issue of assessment of end users has been emphasised by funders recently. For example, the Anglo American and De Beers Chairman's fund have asked TELIP to evaluate its English language improvement courses for teachers in terms of outcomes for pupils in the classroom.

The pressure to evaluate student outcomes is not unique to TELIP but applies to most INSET NGOs in South Africa. The Joint Education Trust (JET), an important funding agency that provides funds to 98 NGOs working in the field of teacher education, held an evaluation conference in February 1996 to assess how INSET NGOs conducted evaluations. At this conference several problems were highlighted, but the point was made that INSET evaluation usually failed to measure outcomes on pupils. This study provides an understanding of the advantages and disadvantages of different research methods which could be used to assess impact on pupils.

4.3. THE RESEARCH DESIGN

The main aim of the study was to gain an in-depth understanding of different research methods so that TELIP could decide on a model best suited to assess the impact of its English language development programme on pupils. The research design, which involved a review of literature, interviews and workshops was set up to achieve this aim.

The research design was based on a conviction that knowledge is not only gained from books but that people also construct knowledge. This resulted in my carrying out interviews with teachers who have experience in the classroom, with other individuals who have experience in the field of evaluation, and holding workshops with TELIP staff. These interviews and workshops helped to decide on a model for TELIP, and confirmed my belief that people also

construct knowledge. In my task as researcher, I see the value of working in an interactive way with the research subjects contributing to the creative process of the inquiry. By working in an inclusive, participative way the research becomes more meaningful to both the researcher and the researched. An important feature of the design the multi-method approach which consisted of gathering data from books, from experts in the field, and from teachers who are the learners and the staff at TELIP. The constant recycling of data eventually led to a common understanding of a model of assessment. The recycling of data is a common feature of action research, a strategy which TELIP has used effectively for several years. A noteworthy point that emerged from the above process is that the research methods that were noted from the literature review and those that were suggested by participants were similar. Participants, however were very clear about the model TELIP should use in the light of pressure from funders who are interested in the end product.

4.4. THE RESEARCH OUTCOMES: IMPLICATIONS FOR TELIP

Four research methods were studied in some depth in this research:

* experimental design
* longitudinal study
* action research
* ethnographic research

Each of these research methods has it's advantages and disadvantages. The challenge in selecting an appropriate methodology for impact assessment is to minimise weakness and maximise strengths.

4.4.1 Reflection on the Methods of Research

Some of the arguments in favour of the **experimental design** were that the researcher is able to manipulate the independent variable and decide the nature of the treatment, that is, what is going to happen to the subjects of the study. The researcher can also decide to whom the treatment is to be applied, and to what extent. After the treatment is applied for an appropriate length of time, researchers observe or measure the groups receiving different treatments (by means of a post-test of some sort) to see if they differ. If the average scores of

the two groups differ, and researchers cannot find any sensible explanation for this difference, they can conclude that the treatment had an effect and is likely the cause of the difference.

The main arguments against this model were that control is most likely achieved with research on humans only in restrictive and artificial settings; and that humans react to artificially restricted, manipulated conditions differently from the way they react to naturally occurring conditions. If the research is conducted under artificial conditions then the generalisibility of the results is severely limited; the control and experimental groups have to be selected carefully to ensure equality.

The difficulty for TELIP in adopting this method would be to find a matching control group. The ethical question of not providing treatment to the experimental group although they may request it goes against the educational philosophy of the project. The need to set up an artificially restrictive setting to rule out other variables that impact on learning also seems unrealistic when working with human beings in a rapidly changing social environment. TELIP staff felt uncomfortable with not providing treatment to the control group for the sake of evaluation. The researcher felt uncomfortable using this model as it adopts an objective approach to the social world and regards social reality as being hard, real and external to the individual. When one subscribes to this view then the methods of scientific investigation will be quantitative. The researcher favours an alternative view of social reality which stresses the importance of the subjective experience of individuals in the creation of the social world. In this view the research focuses upon different issues and approaches them in different ways. The approach takes on a qualitative aspect.

The staff at TELIP also favour an alternative view of social reality and so the experimental model does not fit well with existing epistemologies and practices of the project.

The main arguments in favour of the **longitudinal study** are that both qualitative and quantitative methods can be used to gather data. The focus in this method is on development of understanding and competence over a period of time; pupils' progress can be measured against their previous achievement rather than a against a norm; skills, processes

and knowledge could be assessed by means of student projects, portfolios and performance assessment. This model is acceptable to TELIP as it enables the project to build up a comprehensive picture of individual and group performance as well as a profile of strengths and weaknesses. Furthermore it allows for administering a pre-test and then a post-test to gather evidence of change that might occur. Within this model, elements of ethnographic research could also be used by drawing up a profile of pupils and teachers and also doing observations in the classrooms to make inferences of what is learnt by both pupils and teachers from the TELIP courses.

The main argument against this model were that it is time consuming - by the time results are obtained the person who initiates the study may no longer be associated with the project; the study involves follow up of individuals and individuals change addresses and occupations and become difficult to locate; the cost of collecting data can be high; the researcher does not know if any behaviour change occurs because of the treatment. It is always possible that some other variable actually caused the change.

Yet for TELIP, this model provides an opportunity to do an in-depth study tracing the progress of the pupils in relation to skills learnt and put into practice by their teacher during and after the TELIP course. It provides an opportunity for monitoring pupils' written and spoken English according to the criteria set by the project. It eliminates the need to find a matching control group and it will also help to eliminate many of the problems of contextual variables in experimental design.

As far as **action research** is concerned, the main arguments in favour of this method are that it aims to draw theory out of practice and that teaching becomes an inquiry in action in which the teacher constantly endeavours to critically evaluate and improve the process of education for herself and the people in her care.

For action research to be implemented effectively teachers need an atmosphere in which they are free to identify problems for inquiry, experiment with solutions and express and share ideas with colleagues and administrators. This context is difficult to find in most black

schools in South Africa. TELIP works in such schools and has used action research with tutors who run TELIP courses. These teachers were able to reflect on and improve their practice over time and with assistance. This was possible because TELIP courses were run in the school holidays when the teacher was not burdened with many problems. TELIP staff felt that for action research to be successful teachers in the study need to be trained and provided with ongoing support. This places a substantial human and financial burden on the project. Participants in the study as well as funders felt that this model could be used to improve teachers practice but would not be sufficient to provide evidence of impact.

The main arguments in favour of **ethnographic research** were that it involves extensive field work as the researcher physically goes to the site, the group of people, the institution, the field to collect data. Fieldwork involves becoming intimately familiar through observation and interviews with the phenomenon under study. The researcher studies as much of the problem as possible for as long as possible, interpreting, explaining, hypothesising or theorising about the phenomenon. It makes provisions for natural settings to be investigated. It does not entail extensive pre-field work design as social surveys and experiments generally do, the strategy and even direction of the research can be changed easily, in line with changing assessment of what is required by the process of theory construction.

The main arguments against this model were that participant observation is a time consuming and demanding technique; the researcher must establish rapport and trust with a group and become familiar enough to gain insights into the meaning of their lives; the researcher must be an observer, remaining as open and neutral as possible while collecting information; its nature focuses on single cases or at most on a limited setting of action limiting the validity of using ethnographic data for scientific generalisation, as well as for policy formation and decision making .

The problem for TELIP is that this model does not support a pre-test and post-test and therefore will not be convincing for funders who want concrete evidence of impact on pupils. However as mentioned above elements of this model could be used in a longitudinal study.

4.4.2. Conclusion

The experimental method is not suitable for TELIP as the control and experimental groups have to be selected carefully to ensure equality. That means the groups must have equal facilities and resources, receive equal care and attention, be taught by equally good teachers and so forth. TELIP will have difficulty finding a matching control group. The ethical issue of not providing treatment to the control group even though they may request it is also problematic for TELIP. Furthermore the researcher and TELIP staff favour an alternative view of social reality which stresses the importance of the subjective experience of individuals in the creation of the social world. In this " the research focuses upon different issues and approaches them in different ways. The approach takes on a qualitative aspect in which the experimental model does not fit.

Action research is a model TELIP could continue to use with tutors teaching TELIP courses. This will help them to improve their practice. As a method for impact assessment it does not provide scope for pre-test and post-test which is required by funders as evidence of impact on pupils.

Considering the arguments for and against the four models of assessment it is clear that the most appropriate method for TELIP is the longitudinal study. TELIP staff felt comfortable with this model as it provides scope to use both quantitative and qualitative methods in assessment. The two groups of teachers felt strongly that assessment should be done through observations and interviews over a period of time. Although their suggestions fall largely under ethnographic research the longitudinal study provides scope to observe as well as to interview participants. This model will also meet funders demands that there should be baseline data to make conclusions about the impact of the TELIP programme.

4.4.3. Reflections on the Research Process

The discussions with TELIP staff which was part of this research impacted on their thinking. At the commencement of the study, TELIP leaned heavily on action research as an appropriate model for research and improving practice in the project. As the study commenced each research method with its strengths and weaknesses as noted in the literature review and interviews was discussed by staff. These discussions led to a shift in thinking. In the light of the discussions and the overriding questions asked by donors, staff members gained a greater appreciation of the range of available research methods and agreed as a team on the most appropriate model.

The research process also threw light on the importance of working out clear criteria and considering the impact of contextual factors on learning and teaching. Most importantly, the research process provided staff members with the confidence to state that TELIP can begin to measure impact on pupils without directly telling the teacher what or how to teach.

4.4.4. Indicators of Success in Impact Assessment

Participants in the study emphasised the need for TELIP to work out clear criteria for impact assessment before setting up a study. The literature reviewed emphasised the need to focus on indicators of success so that the evaluation can determine whether these are being met. Weingart (Muller and Vinjeveld. 1995) notes a move amongst evaluators internationally towards focusing on outcome indicators as the most reliable measure of programme impact. This trend is beginning to emerge in South Africa, as shown by Bateson (1995) who combined an examination of classroom practices with quantitative measures of learning outcomes.

Tessmer (1993) points out that the simplest indicator of success is to assess what learners will be able to do when they have completed instruction, often called the 'terminal objective' of instruction. But it is essential to identify the sub-objectives which help to acquire final objectives - in short, the knowledge skills and attitudes that form part of the instruction. Annexes (1987) on the other hand has suggested "progress assessment" rather than "outcomes assessment". The switch can be important symbolically as well as conceptually.

Assessing outcomes implies a certain finality: that a summative evaluation and judgement are to be made, that the bottom line is about to be drawn. Assessing progress, by contrast, implies an on- going, formative process, which in turn, suggests that time remains to make any necessary improvements.

TELIP is well placed to work out formative and summative indicators. The five courses provided by TELIP (see chapter one) build on language skills progressively and this helps to monitor progress made by pupils. For example formative indicators would be that pupils can construct a complete English sentence, then construct a correct paragraph. The summative indicator would be that pupils write a proper essay. This form of assessment is also in line with the principles of Outcomes Based Education the dominant paradigm in education and training in South Africa at the end of the 1990s.

The implication for TELIP is to work out clear criteria for the knowledge and skills teachers are expected to use in the classroom after completing TELIP courses and then measure the extent to which these criteria also impact on pupils.

4.4.5 *Contextual Factors*

The contextual factors within which teachers implement TELIP courses have a bearing on their success or failure. At the institutional level, the authoritarian, bureaucratic nature of the education system during the years of apartheid did not encourage school-based innovation. INSET providers had to struggle with the needs of the system which triumphed over the needs of the individual. This meant that the innovative teaching skills which teachers desired to implement were often hampered by the needs of the system. Under the new government of the African National Congress (ANC) there is a commitment to whole school development and hopefully the environment will become more supportive of innovative ideas.

Participants in the study also highlighted some of the above conditions in the majority of South African schools which need to be kept in focus when conducting an impact assessment.

Poor management and inadequate leadership present major constraint. Schools are embattled by political tensions and inadequate resources. The school ecology is therefore not receptive to innovation.

Teacher characteristics also serve as constraints. Most African teachers are underqualified, young and inexperienced, with a limited knowledge of their teaching subjects and English, the dominant medium of instruction. They suffer low salaries and professional status, and poor working conditions. Their morale has been seriously eroded by the turmoil surrounding black schools since 1976. Their energy is concentrated on survival rather than on innovation.

Other variables that TELIP needs to keep in focus are: ethos of learning and teaching culture; other INSET courses that teachers do; class teacher may not remain constant; language input coming from sources other than TELIP; teachers' own personalities impacting on teaching and learning; pupil drop out rate and pupil maturation having an effect on performance; the role of the principal in supporting change (see chapter one).

Together with the above factors TELIP needs also to consider what Berman and McLaughlin (1977) say about the role of principals in supporting change. They found that projects having the active support of the principal were the most likely to fare well. When setting up an impact assessment study TELIP will need to ask questions such as - is the sample school progressive? Are the teachers encouraged to try new things? Are principals involved in motivating and supporting teachers?

4.5. IMPLEMENTATION

For the first time TELIP will produce a definitive study of the impact of its intervention and its effects over time on a set of pupils. The critical question that this study will ask is: do pupils who are taught by TELIP trained teachers show significant learning gains over time in terms of their knowledge of English?

4.5.1 Working out Criteria for Assessment

TELIP courses have clear outcomes for teachers. What has not been made clear is what the teacher should use or transfer to his or her classroom. TELIP will need to make explicit to teachers which particular outcomes he or she is expected to use in the classroom. This must be followed by supporting the teachers to implement these outcomes in the class before testing pupils for assessment. Below is an example of how criteria can be worked:

Outcomes for the teacher in Course 1

The learner is able to

* do a close reading of a text.
* share responses and opinions on texts with other learners
* identify main ideas and supporting ideas in a paragraph
* transfer main ideas and supporting ideas from a text onto a tree diagram
* tell a story in sequence
* preview a passage and make predictions about the content of the passage before reading it
* use simple present and past tense
* prepare a thorough plan of an essay

From the above criteria the teacher could be asked to teach pupils:

* to write a proper paragraph with main ideas and supporting ideas
* to preview a passage and make predictions
* to use a tree diagram
* use simple present and past tense

As the teacher teaches the above skills, pupils could be tested. This process is repeated as the teacher progresses to course two, three, four and five.

4.5.2. Choosing a Sample

TELIP needs to choose a sample and do an in-depth study over a period of at least one to two years. The choice of the sample will depend on four factors: the number of field units operating in the year 1999 (the year of implementing the assessment study), where they are

operating, the number of learners, the willingness of school principals to allow schools to be in the study. If courses are running in both rural and urban areas then the project will need to decide on whether the study should be rural or urban or both. If both an urban and rural sample is chosen then this will help to make interesting comparisons.

4.5.3. Baseline Data

A detailed profile of pupils and teachers in the sample needs to be compiled. This information will assist when analysing whether it is TELIP courses or other factors that impact on pupils' and teachers' progress or both? If no progress has been made then it will also help to determine whether it is the school environment or the out of school factors that play a role. The other advantage of compiling detailed profiles is that the Project will have a database of users for future reference or further research. In compiling details profiles of teachers and pupils elements of ethnographic research will be used.

The profile of teachers will collect data such as;

* age
* qualification
* other studies
* other English language influences at home and at school.
* number of years teaching

The profile of pupils will collect data such as;

* age
* other responsibilities that affect school work
* languages spoken in the home
* supervision of home work
* access to books
* access to radio
* access to television

Together with this ethnographic data a pre-test will have to be given to both teachers and pupils so that baseline information is collected. It is important to monitor the progress of the teacher and the pupils because if the teacher makes no progress in the TELIP course it would

be difficult to claim that the progress made by pupils had to do with TELIP courses.

TELIP could construct its own pre-test instrument. The Project already has a pre-test for teachers. This could be adapted for pupils. Once the test is constructed it will need to be validated before using it. The test results will be analysed in detail to determine where teachers and pupils fare well and badly. The pre-test will then be followed by regular testing during the period of the study to determine whether progress is being made.

Data must also be collected through classroom observation and interviews. It is important to observe the teacher teach in the classroom before the TELIP course commences and to take notes of the methodology and English language usage both by pupils and teacher. This should be followed by regular observations as the teacher progresses through TELIP courses and uses the skills in class. These observations will provide an opportunity to compare the pupils' and teachers' progress before and after TELIP courses. Such data will also enrich and contextualise the results of language tests. Quantitative data will thus be supported by qualitative data. Interviews with teachers and pupils about their own language confidence before and after the TELIP course should also be used to enhance the data.

4.5.4. Contextual Factors

It will be important for the evaluator to observe carefully the ecology of the sample school. Some questions that will help to describe the context could be: are the teachers encouraged to try new things? Are principals involved in motivating and supporting teachers? Does the school organisation make provision for materials and equipment to assist teachers? Is there a constant change in teachers or are teachers able to remain in place for at least one year? Is the morale of teachers high or low ? What is the work rate in the school? Do pupils attend school regularly? All of these factors impact on learning and teaching and will help to make conclusions concerning the impact of TELIP courses.

4.5.5. Who Should do the Evaluation

Participants in the study and funders such as the Joint Education Trust and Anglo and De Beers Chairman's Fund feel that the evaluation should be done by a person outside the

Project as staff in the project will have vested interests which could make it difficult for them to be objective. However TELIP will work closely with the external evaluator who will be accountable to the Project. TELIP staff will be involved in some of the fieldwork such as observation and interviews so that different perspectives may enrich the data. The external evaluator will be expected to cycle information and analysis continuously to and with the project.

There are two implications of using an external evaluator for impact assessment. The first is funding for the evaluation. The second is identifying and contracting a suitable person to set up the study. One way of getting around the funding issue is to build in the cost of evaluation in funding proposals. In other words, funders who are insisting on impact assessment should be asked to fund the study.

In order to set up a longitudinal study, pre- impact assessment work should start in November 1998 and the study should begin in January 1999. The following schedule has been set up .

IMPLEMENTATION SCHEDULE

Nov/Dec 1998	Feb 1999	March 1999	April 1999	May 1999	June 1999
The following pre-implementation work needs to be completed: *confirm funding for evaluation * confirm pre-test instrument * identify external evaluator	*confirm sample for the study * test pupils and teachers for baseline data *capture contextual factors in school * confirm different ways of testing and monitoring pupils' progress	*complete profile of teachers and pupils *teacher does course one *confirm outcomes to be implemented in class	*teacher implements outcomes in class * support for teacher in class *observe teacher	*test pupils *interview teachers, pupils, principals, heads of department	*write progress report
July 1999	**Aug 1999**	**Sept 1999**	**Oct 1999**	**Nov/Dec 1999**	**Feb 2000**
* teacher does course two *confirm outcomes to be implemented in class	*teacher implements outcomes of course two *support teacher in class *observe teacher in class *interview teachers, pupils, principal, heads of department	*test pupils *monitor contextual factors *teacher does course three *confirm outcomes to be implemented in class	*teacher implements outcomes of course three *support teacher in class *observe teacher in class	*test pupils *interview teachers, pupils, principals, heads of department	*teacher does course four *confirm outcomes to be implemented in class
March 2000	**April 2000**	**May 2000**	**June 2000**	**July 2000**	**Aug 2000**
*teacher implements outcomes of course four *support teacher in class *monitor contextual factors	*test pupils *interview pupils, teachers, principals, heads of department	*write progress report	*teacher does course five *confirm outcomes to be implemented in class	*teacher implements outcomes of course five *observe teacher in class *monitor contextual factors	* interview teachers, pupils, principal, heads of department

			*support teacher in class		*test pupils
Sept 2000	Oct 2000	Nov 2000	Dec 2000		
* progress report	*analyse data collected	*analyse data collected	*final report		

Finally the pressure from funders to assess the impact of TELIP's programme on pupils seemed daunting and almost impossible at the outset of this study. TELIP staff were comfortable with the existing method of assessing impact of its courses on teachers arguing that as the teacher's language confidence improved then the teacher was better able to teachin the medium of English and that pupils would benfit.

This study has helped TELIP staff to understand that impact assessment on pupils can be done. In my own portfolio as evaluator I have gained tremendous knowledge about the implications of different methods of research and feel confident that a longitudinal study will help to assess impact on pupils.

Other NGOs with similar orientation to TELIP and who are under pressure to do impact assessment will benefit from TELIP's experience. More importantly the study will help to make policy decisions about the future of TELIP and will signal to potential funders how best to allocate resources to address the problems of English language usage.

REFERENCES

Allais, C. and McKay, V. (1995) A Sociology of Education. Johannesburg: Lexicon.

Bateson, D.J. (1995) Teaching and Learning Science: The SEP Approach: An Evaluation of the Science Education Project. Cape Town: Clyson Printers.

Bhola, H.S.(1990) Evaluating 'Literacy for Development' Projects, Programs and Campaigns. Germany: Robert Seemann.

Blake, J. et al. (1995) 'A portfolio based assessment model for teachers: encouraging professional growth.' NASSP Bulletin. Vol.79 : 26 - 272.

Borg, W.K. et al. (1995) Applying Educational Research. New York: Longman.

Cohen, L. and Manion, L. (1995) Fourth Edition. Research Methods in Education. London: Routledge.

Dalglee, R.(1994) Impact Evaluation: Phase 1. TELIP, Centre for Continuing Education, University of the Witwatersrand, Johannesburg.

Elliot, J. (1991) Action Research for Educational Change. Great Britain: Biddles.

Fraenkel, J.R. and Wallen, N.E. (1990) How to Design and Evaluate Research in Education. U.S.A: R.R. Donnelly.

George, T. (1997) 'Impact Evaluation.' Paper presented at a JET Evaluation Conference. Johannesburg, South Africa.

Harris and Bell (1986) Evaluating and Assessing for Learning London: Hogan Press.

Hofmeyr, J. and Hall, G. (1995) The National Teacher Education Audit: Synthesis Report sponsored by Danida.

Hammersley, M.(1990) Reading Ethnographic Research: A Critical Guide. London: Longman.

Hartshorne, K. (1994) Evaluation of the English Language Education Trust Durban.

Henerson, M.E., Morris, L.L. and Fitz Gibbon, C.T. (1987) How to Measure Attitudes. U.S.A: Sage Publications.

Improving Educational Quality Project (IEQ).(1996) Evaluation of the English Language Educational Trust. (ELET). Inservice Education and Training (INSET) Programme. South Africa. IEQ.

Joint Education Trust (JET) Conference (1996) Quality and Validity in INSET Evaluations. Johannesburg: Joint Education Trust.

Jaeger, R.M. (1988) Complementary Methods for Research in Education.Washington: American Educational Research Association.

Jansen, J.D. (1996) 'Does Teacher Development Work? True Confessions of a Hardened Evaluator' in JET Conference Proceedings Quality and Validity in Inset Evaluations. Johannesburg: Joint Education Trust:13-19

Johnson, M. C. (1977) A Review of Research Methods. Chicago: Rand Mc Nally College Publishing Company.

Kallaway, P. (1984) Apartheid and Education. Pretoria: Sigma Press.

Kemmis, S. and Mc Taggart, R. (1988) The Action Research Planner. Australia: Deakin University.

King, J. A., Morris, L. L. and Fitz Gibbon, C.T.(1989) How to Assess Program Implementation. U.S.A: Sage Publication.

Lather, P. (1986) 'Research as Praxis'. Harvard Educational Review, Vol. 56, No. 3.

Louw, J. (1996) 'Academic Research, Internal Validity and Programme Evaluation' in JET Conference Proceedings Quality and Validity in Inset Evaluations. Johannesburg:Joint Education Trust: 4-8.

Merriam, S.B. and Cunningham, P.M. (1991) Handbook of Adult and Continuing Education. San Francisco: Jossey Bass.

Merriam, S.B. and Simpson, L. (1984) A Guide to Research for Educators and Trainers of Adults. Malabar: Robert. E. Kreiger.

Mc.Niff, J. (1993) Teaching as Learning: An Action Research Approach. London: Routledge.

Mc Taggart, R. (1991) Action Research, A Short Modern History. Australia: Deakin University.

Mc Millan, J.H. and Schumacher, S. (1993) Research in Education : A Conceptual Introduction. U.S.A: R.R. Donnelley and Sons Company.

Millar, C. Raynham, S.A. Schaffer, A. (1991) Breaking the Formal Frame. U.K: Oxford University Press.

Mouly, G.J.(1970) The Science of Educational Research. New York: Van Nostrand Reinhold Company

Morris, L.L., Fitz Gibbon, C.T. and Lindheim, E. (1987) How to Design a Program

Evaluation. U.S.A: Sage Publications.

Morris, L.L., Fitz Gibbon, C.T. and Lindheim, E. (1987) How to Measure Performance and Use Tests. U.S.A: Sage Publications.

Neuman, W.L. (1991) Social Research Methods. U.S.A: Viacom.

O'Hanlon, C. (1996) Professional Development through Action Research in Educational Settings. London: Falmer Press.

Oda, S.N. and Simulyan, L. (1989) Collaborative Action Research: A Developmental Approach. Great Britain: Taylor and Francis.

Parlett, M., Dearden, G. (1981) Introduction to Illuminative Evaluation: Studies in Higher Education. California: Pacific Soundings Press.

Palmer, G. M. (1992) 'The Practical Feasibility of Diary Studies for INSET.' European Journal of Teacher Education. Vol.15 : 239 - 251.

Powell, E. and Moss, G.D. (1995) 'The Evaluation of an INSET Model for Gwent Primary School.' School Organization. Vol.10 : 26 - 272.

Romiszowki, A .J. (1981) Designing Instructional Systems. London: Kogan Paul.

Reason, P. and Rowan, J. (1991) Human Inquiry: A Source Book of New Paradigm Research. New York: John Wiley and Sons.

Rosenwald, G.C. and Ochberg, R.L. (1992) Storied Lives: The Cultural Politics of Self Understanding London: Yale University Press.

Richards, J.C. and Rodgers, T.S. (1986) Approaches and Methods in Language Teaching. Cambridge: Press Syndicate of the University of Cambridge.

Spindler, G. (1982) Doing the Ethnography of Schooling. New York: CBS College Publisher.

Sherman, R.R. and Webb, R.B. (1990) Qualitative Research in Education: Focus and Methods. Great Britain: Taylor and Francis Printers Ltd.

Skerit, O.Z. (1996) New directions in Action Research. Great Britain: Falmer Press.

Sherman, R.R., Rodman, B.W. (1990) Qualitative Research in Education: Focus and Methods. U.S.A: Falmer press.

Sholen, Lynn and others (1995) 'Developing Student Self' Assessment Strategies.' Tesol Journal. Vol 5: 22-27.

Stecher, B.M. and Davis, W.A. (1987) How to Focus an Evaluation. U.S.A: Sage

Publications.

Taylor, N. (1995) 'Inset, NGOs and Evaluation: A Review'. Paper presented at the Kenton-at-Settlers Conference, Grahamstown.

Tuchten, G. (1994) 'Assessment in Adult Literacy and Basic Education.' Unpublished paper. Johannesburg:DEAL Trust.

Unterhalter, E. Wolpe, H. Botha, T. Badat, S. Dlamini, T. Khotseng, B. (1991) Apartheid Education and Popular Struggles. Johannesburg: Ravan Press.

Vinjevold, P. (1996) Evaluation Report of the Northern Cape Primary School Workbook Pilot Project (NCWP). Johannesburg: Joint Education Trust.

Winter, R. (1987) Action Research and the Nature of Social Inquiry. England: Gower Publishing Co. Ltd.

APPENDIX A

INTERVIEW SCHEDULE NO. 1.

Provide background to TELIP as well as explain purpose of the research.

1. How do you suggest that TELIP evaluate the impact that courses have on pupils? Why?

2. Can you foresee any problems/limitations in the suggested evaluation model? Why?

3. What time frame should the evaluation cover? Why?

4. Who should do the evaluation? Why?

5. How should results of such an evaluation be reported? Why?

APPENDIX B

INTERVIEW SCHEDULE **NO. 2. (Used for language NGOs)**

Provide background to TELIP as well as explain purpose of research

1. Please describe the courses that you offer? (ask for brochure)

2. Do you evaluate the courses? If so how?

3. Do you evaluate the impact of your courses? If so how?

4. Please describe the success/problems you may have had?

5. How would you suggest that TELIP evaluate the impact of the courses on pupils rather than on teachers?

6. What problems do you think TELIP will have in carrying out the evaluation in terms of:

evaluation process
design
credibility
validity
interpretation

APPENDIX C

Interview Schedule for Teachers

Explain purpose of the interview

1. Have the TELIP courses helped you in any way? Please explain.

2. Have you used any of the skills from the TELIP courses in your classroom?
   If yes give examples of how you used these.

3. Have the pupils benefitted from the use of these skills? If yes, give examples.

4. Did you experience any problems when using the skills in the classroom? How did you deal with the problems?

5. How can TELIP assess whether pupils have benefitted?

6. Do you think TELIP will have problems assessing your pupils work? If yes what kind of problems.

7. Are you doing any other studies or short courses besides TELIP?

APPENDIX D

ANALYSIS OF INTERVIEWS

The following table provides an overview of the responses obtained during interviews.

Question 1	How should TELIP assess impact on pupils?
respondent 1	Test teachers and pupils before they start TELIP. Test again after they have done TELIP courses. Testing should be continuous - three monthly fro a period of 2 years. Many ways of testing - for example pen and paper, interview pupils , observe in class
respondent 2	should not be quantitative because it is too complicated to measure language proficiency. Study should be qualitative - use interviews, interview pupils and parents, principal and staff, about changes that might be taking place
respondent 3	test for pupils is crucial - site for research should be schools where teachers have done TELIP courses and then compare with schools which have done no TELIP courses. Matching schools might be difficult but work out a criteria. Use post ad hoc experimental design
respondent 4	choose a random sample of teachers and pupils. Do a longitudinal study - .
respondent 5	longitudinal study on its own or longitudinal study with a quasi experimental study which will help to make comparisons and give substance to data. Strong need to do pre and post test for both pupils and teachers - this will help to establish whether change has taken place.
respondent 6	Do a pre and post test on pupils and teachers. Pre and post observation in the class should also be done. A profile of both teachers and pupils also needs to be done so that we can make some conclusions about where else English input is coming from
respondent 7	the study will be longitudinal. Baseline data is important as a starting point . An instrument needs to be developed to assess language proficiency of teachers and pupils before and after TELIP courses. TELIP might want to develop own instrument or look around for one. To make the study more sophisticated have a control group with similar profile - but in establishing control group there are ethical questions to keep in mind - for example - what happens if the teacher in the control group wants to do TELIP courses.

The main points emerging from question 1 is that TELIP has two choices of an evaluation model.

* longitudinal study on its own
* longitudinal study with a control group

Longitudinal Study with or without the control group should take place over a period of 2 to 3 years.

Respondents feel strongly that to make valid claims about the impact of TELIP courses an in depth study is needed. Some of the respondents suggest that the longitudinal study could be set up with a control group while others felt that the longitudinal study need not have a control group. Setting up a control group is an ethical decision which TELIP needs to consider. There is a strong suggestion that baseline data be collected by carrying out pre and post tests. Only one respondent was opposed to a quantitative study because it is too complicated to measure language proficiency.

Question 2	What Problems/Limitations do you foresee in the suggested model
respondent 1	Classroom observation is a difficult task. Be vigilant of other courses that teachers may be doing. Workshops run by publishers could also influence teachers. Primary school children are not able to reflect Think of ways of asking them . Need a skilled person to interview children - somebody who can get information in a creative way.
respondent 2	time constraints - it takes 2 - 3 years of a longitudinal study before positive claims can be made. Observe factors and changes in the classroom that impact on learning. Watch what has happened in the class, school as a whole and parents to see how these also impact on learning. Constructing instruments to measure teacher's practice and pupil's practice can be problematic. Also look at negative factors such as retrenchment and the new curriculum and how these are impacting on teacher performance. Study is intensive requires gathering evidence in various ways.
respondent 3	In setting up experimental group - ethical question - what if teacher in experimental group requests TELIP courses. TELIP needs to find way around this. Difficult to control other influences that impact on learning - one needs to be vigilant of all other factors impacting on the group's learning. School's learning culture could also be a problem. The class teacher not remaining constant could also be a problem.
respondent 4	Variables that impact on learning. No standard acceptable theoretical model for language learning. Bear in mind the cognitive, emotional growth of the pupils during the time of the study. Extremely important that the control group matches the experimental group. Draw up specific criteria and then measure the success of these.

respondent 5	Confidence takes a long time to build and might not happen in 2 years. Requires skilled personnel to interview pupils to get required data. Principals of schools are anti - experimental group. Practical problems that TELIP might experience are the school's ethos, and learning culture People often are not reliable when providing information etc. Many other factors impact on learning such as other INSET programmes, the environment, as well as the teachers own personality. Before TELIP embarks on this research TELIP needs to work out performance indicators.
respondent 6	Time factor - it takes a long time to set up study . If the research is external it will be expensive. If TELIP constructs test it needs to be validated . There is a need to test the test and report technically on this.
respondent 7	keep in mind variables that impact on learning . TELIP needs to have clear criteria for evaluation and make this explicit before evaluation. Bureaucratic problems may prevent access to the classroom. Improvement is a slow process. Change is not just superficial. Learning takes a long time.

The main points emerging from question 2 is that

TELIP needs to be vigilant of variables that impact on learning and that the study takes time. Some respondents also commented on the difficulty of testing, interviewing and observing children.

Some examples of variables given are:

ethos of learning and teaching culture
other Inset courses that teachers do
class teacher may not remain constant
language input coming from sources other than TELIP
teachers own personality impacting on teaching and learning
changes taking place in the school that impact on learning
negative factors such as retrenchment having an effect on teachers performance

pupils drop out rate
pupils maturation having an effect on performance

* Longitudinal Study takes a long time before conclusions can be made
* Testing, interviewing and observing children are difficult issues.

Respondents also suggest that TELIP work out clear criteria for assessment before embarking on the study. In setting up the study TELIP also needs to keep in mind bureaucratic problems such as access to schools and at times resistance to classroom observation. To gather baseline data an instrument needs to be constructed and this needs to be validated. Be aware that the study is long and in depth as learning and transference of skills takes time.

Question 3	What Time Frame should the evaluation cover?
respondent 1	follow teacher through a course - one and half year of the course. Select a situation and look at it in depth. Observe teachers after each course but be aware of drop outs.
respondent 2	a year or 6 months can give interesting findings but to make claims 2 - 3 years are required. The study is intensive and requires gathering evidence in various ways
respondent 3	study should be done over 2- 3 years
respondent 4	study should be done over 2 to 5 years - not less than 2 years as this will not be meaningful. Receptive skills which are more than spoken and written skills are difficult to observe. Regular testing is important - at least every 3 months during the 2 odd years. The longer the study the steeper will be the learning curve.
respondent 5	It should be over a period of at least 2 years - this intensive study will help to make some claims that TELIP is making a difference.
respondent 6	TELIP should start with a rapid appraisal and then conduct a more in depth study
respondent 7	study should be over 2 to 3 years

The main point arising out of this question is that the longitudinal study must happen over 2 to 3 years . This in - depth study is needed to make any valid claims about what effect TELIP has on pupils. Some respondents suggested an initial shorter appraisal so that TELIP could provide information to fenders.

Question 4	Who should do the evaluation?
respondent 1	Outside persons would be more objective. Fenders tend to put more value on the report if it is not internal
respondent 2	An internal study will be biased but internal resources can be used to supplement the work of the outside person. If the evaluation is for fenders an external evaluation is necessary.
respondent 3	projects should do both kinds of evaluation - internal and external. An external evaluation has the advantage of becoming a powerful document for selling the project. If the primary audience are fenders then an external evaluation is important.
respondent 4	internal evaluation is not a problem as long as the research design is sound and we state that the internal evaluation is in itself a limitation. If the evaluation is for fenders especially then an external evaluation would be valuable.
respondent 5	in house evaluation is not a problem if it is set up in proper research mode and limitations noted. But for fenders it could be valuable to have an outside evaluator.
respondent 6	TELIP evaluates first and then commission an external evaluation
respondent 7	If evaluation is for fenders then it should be external with TELIP involved in the process

The issue emerging in this question is that if the evaluation is for fenders then it should be an external evaluation as this will eliminate bias. TELIP however should be involved in the process and not totally excluded. An external evaluation will help to validate TELIP's impact on pupils as it will take away subjectivity of internal evaluation. Some of the respondents however feel that an internal evaluation if done in the proper research mode with its limitations spelt out could also be valid.

Question 5	How should the report be written?
respondent 1	it should be written in proper research mode and disseminated widely. Discuss findings at conferences. Bring to the notice of NGOs as well as education Departments especially if the evaluation shows good classroom practice
respondent 2	it should be written in proper research mode. If the research is rigorous then it could also be reported in international journals.
respondent 3	write in research mode. Write interim reports and then a final report. Present findings as impressions and be careful of what is been measured.
respondent 4	write the report in research mode but also include an executive summary
respondent 5	write the report in research mode and explicitly state the limitations and constraints experienced
respondent 6	write the report in research mode. Also write interim reports for stakeholders
respondent 7	report could be write in evaluation format or research format keeping in mind the audience for whom it is intended

The main point emerging from this question is that the report should be written in research mode stating all the limitations and constraints and providing an executive summary. Since a longitudinal study takes time interim reports should also be written. The research should be widely disseminated through journals and conferences.

Question 6	Any other suggestions for TELIP
respondent 1	baseline test could be worked out by TELIP but make the test available to stakeholders who do their own assessment of test. Test the test and report technically on this
respondent 2	important that TELIP has clear criteria for evaluation before embarking on the study. Use triangulation for collecting evidence for example classroom observation, interviews, questionnaires for teachers and pupils.
respondent 3	TELIP first needs to analyse and decide what improvements in the class are likely to happen before trying to measure impact. For example do we want to see greater interaction between teacher and pupil? Has teacher moved away from chalk and talk, is there a variation in methodology? Are teachers asking real questions and getting real answers? Are pupils really involved cognitively?
respondent 4	ways in which pupils can be tested: pen and paper, interview pupils, observe pupils in the class. Observe the use of stimulated English usage. Look at written work and frequency of the use of English. Are they using more English.?
respondent 5	TELIP needs to decid on a model that is most feasible for itself. TELIP's own sense of ethics will help to come up with a model. Contact Maths Centre For Primary Teachers - they have set up a longitudinal study for impact assessment.
respondent 6	control group could provide strong data to make claims about TELIP.
respondent 7	Before writing own language test - try and find out what tests exists and work out whether this is suitable. Use a small sample so that an in depth study can be done properly.

Some of the suggestions made for TELIP are that:

* TELIP needs to decide on a model that is most feasible for the project.
* Before writing up own test - look around for language proficiency tests that exist and if these are suitable use them.
* make the criteria for evaluation explicit before setting up study.
* use various data collection techniques for example classroom observations, interviews, monitor written work as well as spoken English of pupils and teachers.

The most important issues that emerge from the interviews can be summarised under the headings :

* models of assessment
* limitations/ problems
* variables that impact on learning
* criteria for assessment

It is interesting that the issues emerging from the interviews are almost the same as those that emerged from the literature review.

APPENDIX E

ANALYSIS OF INTERVIEWS HELD WITH TEACHERS

The following table provides an overview of the responses obtained during interviews with two groups of teachers.

Group 1

Question	Response
1. Have the TELIP courses helped you in any way? Please explain	- improved essay writing skills - able to write better paragraphs with main ideas and supporting ideas - able to write CVs - reading skills have improved- have learnt to skim and scan - able to speak English with confidence and teach through the medium of English - learnt new methodology eg. Group work, pair work and role play - can write a summary - able to use a tree diagram
2. Have you used any of the skills from the TELIP courses in your classroom? If yes give examples of how you used these.	All the teachers claim to use some of the skills in their classrooms. examples of skills used: - taught pupils to skim and scan - in composition writing get pupils to brainstorm ideas for writing essays - encourage pupils to write poems - taught pupils to use tree diagram for summary writing
3. Have the pupils benefitted from the use of these skills? If yes, give examples.	- pupils write better essays - they write poems - they can skim, and scan - they use tree diagrams - they write complete sentences and get tenses correct
4. Did you experience any problems when using these skills in the classroom? How did you deal with the problems?	- the problem is that Heads of Departments (HOD) do not support new ideas, want us to follow the syllabus and do not understand how the new method helps. HODs not very supportive of the courses we do.

5. How can TELIP assess whether pupils have benefitted?	Choose a few schools in an area observe the teacher in the class interview the teacher look at the pupils books speak to the pupils
6. Do you think TELIP will have problems assessing your pupils work? If yes what kind of problems?	the problem will be to arrange with the Principals and then you can come
7. Are you doing any other studies or short courses besides TELIP?	one teacher is doing an advanced diploma in adult education but finds that the reading and writing skills learnt in TELIP help with his studying

The main points emerging from this interview are :

* that teachers use skills learnt in the TELIP courses in their classrooms
* specific examples of what they use are:
- skimming and scanning skills
- teach pupils to brainstorm ideas before writing essays
- teach pupils to use a tree diagram to summarise ideas
-teach pupils to write complete sentences and use correct tenses

*the ways in which TELIP can assess pupils is by
-observation in the classroom
- interviews with teachers and pupils
- looking at pupils written work

Responses of group 2

Question	Responses
1. Have the TELIP courses helped you in any way? Please explain.	- able to use logical connectors - tenses have become clearer - learnt to use reading skills such as skimming and scanning -able to use methodology such as role play and group work - learnt creative writing skills - able to do theme teaching - use tree diagram to summarise - plan before writing an essay - write better paragraphs
2. Have you used any of the skills from the TELIP courses in your classroom? If yes, give examples.	-all the teachers have used some of the skills in their classrooms - examples: - teach pupils creative writing - grammar - able to teach tenses - teach skimming and scanning skills - brainstorm ideas for writing an essay - write properly structured composition with introduction, body and conclusion
3. Have the pupils benefitted from the use of these skills? If yes give examples.	-pupils write complete sentences - grammar is improved, especially the use of tenses - pupils can skim and scan - pupils can summarise a piece of text - they write poems for creative writing
4. Did you experience any problems when using these skills in the classroom? How did you deal with the problems?	- no problems really - it will be nice if TELIP can observe us in the class and then tell us if we are on the right lines. Principals and HODs not interested.

5. How can TELIP assess whether pupils have benefitted?	- choose some schools in the area where teachers have done TELIP courses - look at pupils performance especially their books - comprehension and composition books - observe lessons - speak to the pupils - observe how teacher and pupils speak English
6. Do you think TELIP will have problems assessing your pupils work? If yes what kind of problems?	- we will be happy to have you in the class but you need to get permission from District Officials, Principals and HODs - principals and HODs give no support and no encouragement - in fact they are not even aware of the course we do
7. Are you doing any other studies or short courses besides TELIP?	2 teachers are doing a teacher upgrade diploma with Taung College of Education. The skills of reading and writing learnt in TELIP courses helps with the studies

The main points emerging from this interview are:

* that teachers use skills learnt in the TELIP courses in their classrooms
* specific examples of skills used are the following ;
- teach pupils creative writing
- teach pupils use of correct tenses
- teach skimming and scanning skills
- teach pupils brainstorming and clustering techniques
- teach pupils to write properly structured compositions with introduction, body and conclusion.

* TELIP can assess pupils work in the following ways:
- choose some schools where teachers have done TELIP courses
- look at pupils performance - especially their books - composition, and comprehension books,
- observe lessons
- speak to pupils
- observe how teachers and pupils speak English

Author: Dawjee R.C
Name of thesis: Strategies for assessing the impact on primary school pupils of a language development programme directed at teachers

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2015

LEGALNOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page)for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.