

**THE DEVELOPMENT OF AN EDUCATION MANAGEMENT INFORMATION
SYSTEM FROM A SENSEMAKING PERSPECTIVE AND THE APPLICATION OF
QUANTITATIVE METHODS TO ANALYSE EDUCATION DATA SETS**

CHRISTOFFEL VAN WYK



Dissertation presented for the degree of Doctor of Philosophy at Stellenbosch University

Supervisor: Prof. S. van der Berg

Co-supervisor: Prof. S. Berkhout

December 2006

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:

Date: 24 November 2006

ABSTRACT

Information is a necessary resource, produced by information systems and is a key building block to the management and decision-making in any organisation. The National Department of Education's guidelines to establish Education Management Information Systems (EMIS) in provincial departments is a recognition that proper management, planning and evaluation are contingent on quality data, data that is complete, relevant, accurate, timely and accessible. The lack of quality data and the lack of integration with other information systems hamstringing the effective use of EMIS. This study addresses these limitations in three basic objectives: a) developing an information systems development model, b) applying the model in a real-life context of the development of the Western Cape EMIS, and c) applying quantitative methods on integrated data sets derived from the EMIS in the Western Cape and other information systems.

The study culminates in the development of a four-phase process model for developing and using EMIS in an integrative manner that would provide a more comprehensive picture for policy and decision-making. It outlines the establishment of an information systems development (ISD) model that integrates innovative emerging trends, such as improvisation, bricolage and sensemaking, in designing and implementing information systems. These approaches postulate that beyond the numbers and quantifiable world there is a complex reality that traditional approaches do not always capture. These include, amongst other things, the atmosphere, culture and structure of an organization, together with the behaviour, emotions, knowledge and experiences of all the people who in one way or another interact with the information system.

The research presents an empirical application of this developed ISD model in education management information system (EMIS) and underscores the role of information systems in everyday practice. This work practice (Practice-in-Action) approach is used to describe how the day-to-day actions and practical experiences of role players contribute to the design, development, implementation, testing, maintenance and improvement of the EMIS and is used as a lens for understanding ISD.

The study further uses quantitative methods, namely education production function and learner flow-through models, to illustrate how the process of knowledge discovery in large data sets in EMIS could be facilitated. The education production function aims to identify those variables that could have a significant influence on the achievement of students in the matriculation examination. The learner flow-through models attempt to measure the effect of learner dropout and repetition on internal efficiency of the education system. Data analysis was facilitated through integration of data sets from various sources, and in turn illustrates the important role of bricolage in ISD. Through this analysis, the role of information systems of this nature to make sense of reality was highlighted. Policy making then can build on the findings from such data analyses to investigate in greater depth any trends or emerging problems, going beyond only the quantitative and macro level analysis by studies at the qualitative and micro levels.

OPSOMMING

Inligting is 'n noodsaaklike bron wat uit inligtingstelsels voortkom en is 'n sleutel bousteen vir besluitneming en bestuur in enige organisasie. Die Nasionale Departement van Onderwys het riglyne neergelê vir die totstandkoming van Onderwys Bestuursinligtingstelsels (OBIS) as erkenning dat behoorlike beplanning en evaluering afhanklik is van kwaliteit data. Kwaliteit data is volledig, relevant, akkuraat, tydig en toeganklik. Die gebrek aan kwaliteit data en die gebrek aan integrasie met ander inligtingstelsels kortwiek die doeltreffende gebruik van OBIS. Hierdie studie spreek hierdie gebreke deur drie doelwitte aan: a) die ontwikkeling van 'n inligtingstelselsmodel b) die toepassing van die model op 'n werklike situasie, byvoorbeeld die OBIS in Wes-Kaap en c) die toepassing van kwantitatiewe metodes op geïntegreerde datastelle soos verkry uit OBIS van die Wes-Kaap en ander inligtingstelsels.

Die studie kulmineer in die ontwerp van 'n vier-fase proses-model vir die ontwikkeling en gebruik van OBIS op 'n geïntegreerde manier, wat tot voordeel van beleid en besluitneming kan dien. In die totstandkoming en vestiging van die model vir die ontwikkeling van inligtingstelsels, word vernuwende en opkomende idees soos improvisasie, bricolage en betekenisgewing ingesluit. Hierdie benaderings veronderstel dat afgesien van die meetbare wêreld, daar 'n ander werklikheid is wat die atmosfeer, kultuur en struktuur van die organisasie insluit. Gepaardgaande hiermee word die gedrag, emosies, kennis en ervaringe van die mense wat op een of ander manier interaksie met die stelstel het, aangespreek.

Die navorsing beklemtoon die rol van inligtingstelsels in alledaagse praktyke en stel die empiriese toepassing van die inligtingstelselsmodel in OBIS voor. Hierdie werksgerigte praktyke (Praktyk-in-Aksie) word gebruik om die bydrae van alledaagse aktiwiteite en ervaringe van sleutelrolspelers tot die ontwerp, implementering, toetsing en instandhouding van inligtingstelsels, te verstaan.

Die studie gebruik verder kwantitatiewe metodes, soos die opvoedingsproduksie-funksie en leerder-deurvloei modelle om die proses van kennis-ontdekking in groot OBIS datastelle te illustreer. Die opvoedingsproduksie-funksie poog om veranderlikes wat 'n beduidende invloed op die matriek-eksamen het, te identifiseer. Die leerder-deurvloei modelle poog om die uitwerking van leerder-uitvalle en -herhalings op die interne effektiwiteit van die onderwysstelsel te bepaal. Data-analise is moontlik gemaak deur die integrasie van datastelle van verskillende bronne om sodoende die belangrike rol van bricolage in ontwikkeling van die inligtingstelsels te illustreer. Hierdie soort analise beklemtoon die belangrike rol van inligtingstelsels in die betekenisgewing van die werklikheid. Beleidsmaking kan dan voortbou op die bevindinge van sulke data-analises om indiepte die uitgeligte probleme te ondersoek. Sodoende vorder die analise verder as net die kwantitatiewe en makro vlakke, maar word tot op die kwalitatiewe en mikro vlakke geneem.

ACKNOWLEDGEMENTS

I would like to extend thanks to my supervisor, Prof. S. van der Berg for providing guidance and support throughout the study. Likewise, I am also thankful to my co-supervisor Prof. S. Berkhout for her valuable inputs.

I gratefully acknowledge the valuable inputs and contributions of my friend and mentor Dr. Luis Crouch and his willingness to share ideas.

Appreciation is also extended to the Western Cape Education Department for allowing me to utilise the EMIS data sets for the analysis in this study and I specifically wish to thank the EMIS staff and acknowledge their contributions to this study.

Sincere thanks to Dr. Mario Smith for his insightful contributions and who on short notice was prepared to do the editing.

I am appreciative to my wife, Sherine and children, Liesl, Robin and Adrian who were generous with their time and support. Special thanks to Sherine who was a pillar of strength throughout my study. She fulfilled the role of editor, proofreader and helped with the final preparation for printing. Without her dedicated assistance and continuous encouragement, it would not have been possible to complete the thesis.

I express my sincere thanks to my Heavenly Father for the inspiration and the strength received. Without His divine help, this thesis would not have been possible.

TABLE OF CONTENTS

CHAPTER ONE	1
THESIS OVERVIEW AND RESEARCH METHODOLOGY	1
1.1. Thesis objectives	1
1.2. Rationale for the study	2
1.3. Research methodology	8
1.3.1. Model development	8
1.3.2. Model application	10
1.3.3. Analysis	14
1.4. Overview and structure of the thesis	15
CHAPTER TWO	18
INFORMATION SYSTEMS DEVELOPMENT: LITERATURE REVIEW	18
2.1. Traditional view of information systems development	19
2.1.1. Linear design approach: Information systems development life cycle (ISDLC)	20
2.1.2. Management framework: Information Orientation (IO)	22
2.1.3. Functionalism	26
2.2. The application of the systems theory approach in information systems development	27
2.3. Information systems development in the knowledge economy	31
2.3.1. Information systems development is more than a rational activity	32
2.3.2. Revisiting methods and methodology in growing and emerging organisations	36
2.3.3. A sensemaking perspective and the important role of improvisation and bricolage as emerging concepts in information systems development	38
2.3.3.1. Improvisation	40
2.3.3.2. Bricolage Design	42
2.3.3.3. Sensemaking	42
2.4. Comparative summary of the subject literature	46
2.5. Components and activities of the fundamental conceptual framework or model	50

2.6. Conclusion	54
CHAPTER THREE	57
PRACTICAL APPLICATION OF THE ISD MODEL: AN EMPIRICAL STUDY OF EMIS	57
3.1 Introduction	58
3.2 Practice-in-Action approach method	58
3.3 Key Actors in developing EMIS	61
3.4 PHASE ONE: Information systems analysis and design (survey questionnaire)	65
3.4.1 Introduction	65
3.4.2 Input as an information systems constituent	66
3.4.2.1 A Sensemaking perspective of the survey design	67
3.4.2.2 Survey design as improvisation	69
3.4.2.3 Survey construction as bricolage design	70
3.4.2.4 Key actors in the questionnaire content development process	71
3.4.3 Transformation processing as a necessary characteristic of ISD	71
3.4.3.1 The transformation as a sensemaking process	72
3.4.3.2 Bricolage design	73
3.4.3.3 Improvisational action in the transformation processing	73
3.4.3.4 Key Actors in the transformation process	74
3.4.4 Output of the information systems design phase	76
3.4.4.1 Output from a sensemaking perspective	76
3.4.4.2 Key Actors	77
3.4.5 Questionnaire feedback and control	78
3.4.6 Concluding remarks on survey design phase	78
3.5 PHASE TWO: Application development	79
3.5.1 Introduction: Interactive web-based information system	79
3.5.2 Input of the application development phase	81
3.5.3 Transformation processing: Steps in building a web-application	82
3.5.4 Output	86
3.5.5 Key Actors	87
3.5.6 Concluding remarks on the application development phase	87

3.6	PHASE THREE: Information system data processing	88
3.6.1	Introduction	88
3.6.2	Input : Data Entry	89
3.6.3	Transformation processing	89
3.6.3.1	Sensemaking	90
3.6.4	Output of the phase	90
3.6.4.1	Quality of Data	91
3.6.5	Key actors	92
3.7	Concluding remarks on data processing phase	92
	CHAPTER FOUR	96
	USING QUANTITATIVE METHODS TO MEASURE THE IMPACT OF SELECTED VARIABLES ON STUDENT ACHIEVEMENT	96
4.1	Introduction	96
4.2	Different data sets as input for the application of statistical techniques	97
4.2.1	Sensemaking as an ongoing process	97
4.2.2	Bricolage , sensemaking and information systems development	98
4.2.3	The variables used as input for quantitative analysis: An exploration from a sensemaking perspective	102
4.3	Transformation of data by using statistical methods to analyse the resource variables that influence student achievement	104
4.3.1	Scatterplots to explore the bivariate relationship between dependent and independent variables	105
4.3.2	The Correlation Coefficient (r)	111
4.3.3	Multiple Regression Analysis	113
4.3.3.1	Introduction	113
4.3.3.2	Literature review of education production function analysis in the USA	115
4.3.3.3	Education production function analysis: A South African perspective	115
4.3.3.3.1	Empirical strategy and theoretical framework	115
4.3.3.3.2	Comparative summary of literature review	128

4.3.3.4	Research findings: Analysis of education production function results based on input variables of schools in the Western Cape	140
4.3.3.4.1	Analysis of estimates	140
4.3.3.4.2	Interpretation of the interaction between the dummy variable CED and the average teacher qualification (equation 4)	144
4.3.3.5	Application of production function results: a sensemaking perspective	150
4.4	Output and results : Discussion and concluding remarks	154
	CHAPTER FIVE	155
	AN ANALYSIS OF LEARNER ENROLMENT AND THE FLOW OF LEARNERS THROUGH THE EDUCATION SYSTEM: ASSESSMENT OF INTERNAL EFFICIENCY IN EDUCATION	155
5.1	Introduction	155
5.2	Different data sets as input for the application of flow through models	156
5.3	Transformation processing : Flow models to analyse data	157
5.3.1	The effects of dropout and repeaters on internal efficiency of the education system	158
5.3.1.1	Cohort Analysis	158
5.3.1.2	Educational attainment around the world	172
5.3.1.2.1	Data and Methodology	172
5.3.1.2.2	Enrolment and dropout patterns among the poor	174
5.3.1.2.3	“Wealth gaps” across countries.	175
5.3.1.2.4	Attainment profiles as diagnostic tools.	175
5.3.2	The effect of socio-economic status, age and gender on the educational enrolment and attainment in the Western Cape	177
5.3.2.1	Data and methodology	177
5.3.2.1.1	Constructing a poverty index of the community	177
5.3.2.1.2	Attainment profiles for the age cohort 15 to 19 of the Western Cape	179
5.3.2.1.3	Wealth gaps	182

5.3.2.1.4	Attainment profile for the age cohort 20 to 30 by gender of the Western Cape	182
5.3.2.1.5	Comparison of the attainment profile of the Western Cape with other provinces	184
5.3.2.1.6	Attainment profile for different age cohorts	186
5.3.2.2	Attainment profile as a diagnostic tool in the education system of the Western Cape	189
5.3.3	The Reconstructed Cohort Method in the education system of the Western Cape	190
5.3.3.1	Introduction	190
5.3.3.2	The reconstructed cohort method: Data requirements	191
5.3.3.2.1	Computation of flow rates using data on repeaters and enrolment.	192
5.3.3.2.2	The calculation of the transition rates (repetition, promotion and drop out rates)	207
5.3.3.3	The reconstructed cohort method: analysing the flow of learners through the education system using transition rates	214
5.3.3.3.1	Reconstructing a cohort of learners using the transition rates of a particular year	215
5.3.3.3.2	Indicators of internal efficiency	217
5.3.3.3.3	Reconstructing a cohort of learners using the transition rates of a particular year and applying the transfers of learners from other systems.	225
5.3.3.3.4	Reconstructing a cohort of learners using the average transition rates for the years 1995 to 2001	228
5.3.3.3.5	Reconstructing a cohort of learners using the actual transition rates for the years 1995 to 2001	231
5.3.4	Concluding remarks	234
		235

CHAPTER SIX

SUMMARY, FINDINGS AND IMPLICATIONS FOR FURTHER

RESEARCH 235

6.1 The development of an ISD-model 235

6.2 Using the ISD model in a practical and real situation based on a Practice-in-Action methodology 236

6.3 An illustration of sensemaking through the application of quantitative methods to data obtained through the ISD model 237

6.4 The EMIS Process Model 243

6.5 Concluding remarks about quantitative analysis: An exploration from a sensemaking perspective 246

7 REFERENCES 248

APPENDIX A 268

APPENDIX B 269

APPENDIX C 275

LIST OF TABLES

Table 1	A Juxtaposition of new emerging approaches and the functional approach: Research findings	45
Table 2	Conceptual Framework for ISD based on literature review	52
Table 3	Example of the data entry table: Survey instrument	85
Table 4	Variable list and description	103
Table 5	Correlation between aggregate grade 12 results and other variables	112
Table 6	Multiple Regression Estimates	143
Table 7	Simulated scenarios of the interaction of REQV and CED	149
Table 8	Top over-achieving schools	151
Table 9	Distribution of top 50 schools according to poverty quintiles	152
Table 10	Comparison of schools within the same geographical area	154
Table 11	Attainment profile for the ages 15 to 19 per grade, by economic group in the Western Cape	179
Table 12	Attainment profile for the ages 20 to 30 per grade by gender: Western Cape	183
Table 13	Attainment profile for the ages 20 to 30 per grade by Province	185
Table 14	Attainment profile for different cohorts in the Western Cape	187
Table 15	Enrolment by year and gender in the Western Cape (1995 to 2002)	193
Table 16	Repeaters by year and gender in the Western Cape (1995 to 2002)	201
Table 17	Transfers of learners from other systems into the education system of the Western Cape in the year 2002	202
Table 18	Promotees by year and gender (1995 to 2001): Western Cape	204
Table 19	Dropouts by year and gender (1995 to 2001): Western Cape	206
Table 20	Transition rates for male and female learners in the Western Cape	211
Table 21	Transition rates for female learners in the Western Cape	212
Table 22	Transition rates for male learners in the Western Cape	213

Table 23	Summary of all the indicators of internal efficiency according to reconstructed cohort method for model 1	220
Table 24	Survival rates with and without repetition by grade	222
Table 25	Summary of all the indicators of internal efficiency according to reconstructed cohort method for model 2	226
Table 26	Summary of all the indicators of internal efficiency according to reconstructed cohort method for model 3	229
Table 27	Summary of all the indicators of internal efficiency according to reconstructed cohort method for model 4	232
Table 28	Requirements in respect of REQVs	268

LIST OF FIGURES

Figure 1	Structure of thesis	15
Figure 2	Systems theory	48
Figure 3	Entity-Relationship diagram	86
Figure 4	Aggregate vs. number of learners who repeated grade 11 (Repeaters)	106
Figure 5	Aggregate vs. poverty index of the community	107
Figure 6	Aggregate vs. number of teachers per 1000 learners paid by the school (SGB teachers)	108
Figure 7	Aggregate vs. average qualifications of teachers (REQV)	109
Figure 8	Aggregate vs. number of teachers paid by the state (State teachers)	110
Figure 9	Aggregate vs. average experience of teachers (Experience)	111
Figure 10	Over-achiever and under-achiever schools	153
Figure 11	Holistic operation model of efficiency	170
Figure 12	Cumulative percentage of population for ages 15 to 19 per suburb in the Western Cape according to poverty index	178
Figure 13	Attainment profile for ages 15 to 19, by economic group	181
Figure 14	Attainment profile for ages 20 to 30 by gender: Western Cape	183
Figure 15	Attainment profile for the ages 20 to 30 per Province	186
Figure 16	Attainment profile in Western Cape for different age cohorts	188
Figure 17	Enrolment by year and gender per grade in the Western Cape (1995 to 2002)	194
Figure 18	Continuous flow of learners from grade 5 (1995) to grade 12 (2002): Western Cape	196
Figure 19	Age specific ratio (Comparison of Grade 1 enrolment with corresponding age of grade 1 learners in the in the Population): Grade 1 by year (1996 to 2002): Western Cape	197

Figure 20	Age specific ratio (The percentage of learners in the Population with the corresponding age of those in grade 1): Grade 1 by year (1996 to 2002): Western Cape	198
Figure 21	Comparison of Enrolment per grade and population in grade-appropriate age for 2002: Western Cape	199
Figure 22	Age specific ratio 2002 (The comparison of appropriate age in Population with the corresponding age in the respective grades): Western Cape	199
Figure 23	Average repetition rate (1995 to 2001) by gender and grade	208
Figure 24	Average promotion rate in the Western Cape by gender and grade (1995 to 2001)	209
Figure 25	Average dropout rate in the Western Cape by grade and gender (1995 to 2001)	210
Figure 26	The reconstructed cohort method according to method 1	216
Figure 27	Survival rate per grade, based on the first row of Table 24	223
Figure 28	Survival rates and dropouts based on data in Table 24	223
Figure 29	Coefficient of efficiency	224
Figure 30	Input-Output Ratio	224
Figure 31	The reconstructed cohort method according to method 2	227
Figure 32	The reconstructed cohort method according to method 3	230
Figure 33	The reconstructed cohort method according to method 4	233
Figure 34	EMIS Process Model as a Research Outcomes	244

LIST OF ABBREVIATIONS AND ACRONYMS

ABET	Adult Basic Education and Training
ASS	The Annual School Survey
CED	Cape Education Department
CFA	Confirmatory Factor Analysis
DET	Department of Education and Training
EMIS	Education Management Information Systems
GTM	Grade Transition Model
HDI	Human Development Index
HOR	House of Representatives
IO	Information Orientation
IS	Information Systems
ISD	Information Systems Development
ISDLC	Information Systems Development Life Cycle
IT	Information Technology
LSEN	Learners with special educational needs
MIS	Management Information Systems
OCR	Optical Character Recognition
PWP	Professional Work Practice
REQV	Relative Education Qualification Value
SDLC	Systems Development Life Cycle
SEM	Structural Equation Modeling
SES	Socio-Economic Status
SGB	School Governing Body
SRN	School Register of Needs
TAAS	Texas Assessment of Academic Skills
UNESCO	The United Nations Educational, Scientific and Cultural Organization
USR	User Requirement Specifications
WCED	Western Cape Education Department
WLS	Weighted Least Squares

CHAPTER ONE

THESIS OVERVIEW AND RESEARCH METHODOLOGY

1.1 Thesis objectives

The aim of this thesis is to establish a process model that could serve as a roadmap to facilitate Information Systems Development (ISD) for the generation of large-scale datasets. In the development of this process model, the study aims to integrate innovative emerging trends, such as improvisation, bricolage and sensemaking in designing and implementing information systems. The process model will then be applied (tested) in a real-life context to elucidate the establishment of Information Systems Development in everyday work practice. Further, quantitative methods, such as Education Production Function, Reconstructed Cohort Method and Learner Attainment Profiles will be employed to illustrate how the process of knowledge discovery in such large datasets in Education Management Information Systems (EMIS) could be facilitated.

The first objective of this thesis was to develop an information systems model based on a body of information systems development literature that could provide a framework for ISD.

Through an information system, datasets become available for analysis. This framework could underscore that the process of ISD is not always a rational and planned, controlled and repeatable activity. It aims to illustrate that emerging approaches such as bricolage design, improvisational action and the sensemaking process could be key activities to speed up development. These emerging approaches could indicate that ISD is more than just a process-

based exercise, as there is also a human element inherent in the process. The active role of the key actors in the development process will be highlighted.

The second major objective of this thesis was to apply the ISD model empirically and show the flexibility of information systems and the practical role of improvisation, bricolage and sensemaking when implemented in a particular environment, such as an Education Department.

The third objective was to elicit knowledge embedded in datasets through the application of quantitative techniques. To make sense of the data sets, the research intends to show that applying quantitative methods on datasets (data analysis) is not only a numbers-driven exercise. Further, this study aims to identify those variables that could have a significant influence in improving the achievement of students in the matriculation examination.

1.2 Rationale for the study

The South African Government (1995, 1998) is committed to eliminating imbalances of the past through redress and the equitable distribution of resources. The education system spearheaded transformation from an apartheid state to a democratic society. Van der Berg (2005b:1) states that “the [South African] school system is perceived to be the vehicle for transforming a greatly unequal society into a more egalitarian one”. The reform of the South African education system is an example of how policies were implemented to effect transformation. Policies such as the national policy on National Norms and Standards for School Funding (Department of Education:1998) and the Post Provisioning Norms for

Educators (Department of Education: 2002) reflect this targeting of the government on the redistribution of resources for redress, equity and quality.

However, educational reform requires more than just the redistribution of physical resources. A major concern, given the budget constraints and limited resources of the country, is how to invest in education ensuring the optimal application of resources. Data management and data analysis could have a significant influence on the answers to all these concerns. Information systems could increasingly be instrumental for the national and provincial governments to investigate the challenges of poverty, unemployment, and social and educational inequality.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) posited that information systems are integral to the management, planning and evaluation of an education system (Carrizo, Sauvageot & Bella, 2003). However, the effective use of information systems is contingent on quality data¹ that is complete, relevant, accurate, timely, and accessible. This awareness has led to the genesis of Education Management Information Systems (EMIS) around the world (Carrizo, et al. 2003; Wako, 2003). Carrizo et al. (2003:12) suggest that “in some countries where data are available...[in] education management systems, policy-makers hardly use them to guide education policies”. UNESCO (2006) concurs that an Education Management Information System is more than just mere data collection, but should facilitate strategic decision-making, formulation of policies (their management and evaluation) and budgeting in education.

¹ Data in the context of this study will be used in a singular form; refer to par 3.6.4.1 (Chapter 3) for a comprehensive discussion on the characteristics of quality data.

During the period 1996 to 2005 the author increasingly became aware of the lack of quality data for decision-making. Furthermore, during these years it was apparent that education departments generated huge amounts of data on a regular basis. The author's experience during these years was that management in general tended to emphasise routine administration, for which raw data was often sufficient, and that the statistical analyses of the data did not often occur. According to a UNESCO report

“policy-makers and other actors in management and planning need easily understandable and interpretable data. These should be supported by in-depth analysis on the functioning of the system that helps in policy formulation, planning of relevant actions, and in monitoring and evaluation of the latter” (Carrizo et al., 2003:5).

A variety of shortcomings in the education management information system of South Africa could be ascribed to the lack of quality data. The ensuing paragraphs attempt to describe these limitations in the education information systems:

The lack of unique identifiers: One principle that helps to integrate information systems is a commonly used unique identifier. A unique identifier is a common field in database terms that links systems together. The lack of commonly used unique identifiers that allow linkage across data systems could contribute to the unavailability of integrated information systems. It is important to have such a linkage, especially in a provincial government such as the Western Cape, where there are 12 Departments developing systems that could remain isolated and exist as islands of data.

While a great deal of data was collected by different units in the education department in the Western Cape, finance, personnel, examinations, curriculum, and more, it seemed that these

data sets could not always be used to answer many of the questions of managers, politicians and policymakers. Through managing, the data for the Western Cape Education Department (WCED) the author interacted with various role players at all employee levels.

Through this involvement, it became apparent that there was a need for data integration across units. Such an integration exercise can enhance information systems in the department and simultaneously contribute to the establishment of an enterprise-wide information system. An enterprise-wide approach links data from all units into one information system. The lack of a department-wide information system resulted in units developing their own information systems. This fragmented approach and silo frame of mind led to the maintenance of multiple systems and inconsistent data in the department. The result was that most units obtained their data from disparate operational systems and the lack of such data integration in the department resulted in fragmentation of systems with concomitant questions about comparability, validity and reliability.

Data levels: Data that was collected within the department was not always at the same level and the lack of such datasets that were at the same data-level within the department limited data analysis practices. Datasets were compiled at the level of the individual, institution, suburb or municipality that made subsequent integration a complex process.

Lack of skills: The lack of capacity and sound technological skills could act as barriers for the utilisation of data and the resultant lack of data analysis practices. Furthermore, administrators with limited skills were expected to manage activities that required technical knowledge. Management that is uncertain or not skilled enough to utilise available data could have a

negative influence on decision-making, while data gatherers may not know how to analyse, interpret and present data.

Windham (1993:25) expands on the notion of using data for management and states that the “bureaucratic need for greater information about increasingly diverse systems of education and the political advantage of data-based argument have heightened the attention given to both educational data collection and to its uses”.

In a complex post-apartheid South Africa where there is a great demand to do well within fiscal constraints, the utilisation of such information systems is fundamental in order to manage strategically and to deploy resources optimally. The National Department of Education (2004) also recognised this need and has implemented the Education Information Policy to enhance and encourage information use in the education system. This policy provides guidelines to establish EMIS in all provincial governments of South Africa. In 1995 the Western Cape Education Department (WCED) established an Education Management Information System (EMIS). EMIS is the section responsible for maintaining the primary database on public ordinary and independent pre-primary, primary and secondary schools, as well as schools for learners with special educational needs (LSEN). This study attempts to describe the processes, procedures and challenges during the development of EMIS. The research draws extensively from the author’s experience in managing the data for the Western Cape Education Department (WCED), and being an active participant in the development of the system. Thus the present study captures the attempts to develop an integrative Information Systems Development (ISD) model that could address the shortcomings in the current construction and use of EMIS. To this end, the study used the WCED as a case study to

illustrate how islands of data can be transformed and integrated to provide a more comprehensive picture that could inform decision-making.

The aim of this research was to reflect on two key aspects of ISD, namely the management and the analysis of large data sets. The study was conceptualized in two parts:

The first part (chapters two and three) provides an in-depth description of information systems development, using EMIS during the ten years 1996 to 2005 in the Western Cape as a case study. The focus is on the procedures and processes in a real-life (particular) environment and in their natural settings. This part aims to capture the activities and experiences of major role players during the information systems development process.

The rationale for describing these processes and procedures is that beyond the numbers and quantifiable world that can be captured in computer information system there exist a complex reality that, amongst other things, includes the atmosphere, culture and structure of the organization. In attempting to capture these complexities within the EMIS, emerging trends such as improvisation, sensemaking and bricolage² were utilised in the information systems development process.

² Improvisation, sensemaking and bricolage are emergent trends in ISD, see Chapter 2.3.3 for an in-depth discussion of these processes.

“Sensemaking is defined as the process of creating situation awareness in situations of uncertainty”(Leedom 2001:8). The sensemaker for Weick (2001:9) is to “convert a world of experience into an intelligible world”. Sensemaking for Weick (1995:17) is based on seven properties A few of these properties, namely “retrospectiveness” (meaningful lived experience), “ongoing” and “enactiveness” (explain how entities get there in the first place and that people often produce part of the environment they face) are the relevant sensemaking dimensions included in this study.

Bricolage, according to Weick (2001:62) “means to use whatever resources and repertoire one has to perform whatever task one faces.” A Bricoleur therefore “is a person that makes do with whatever tools and materials are at hand”. Improvisation according Ciborra (1999a:78) is “extemporaneous”. “It is situated performance where thinking and action emerge simultaneously and on the spur of the moment (Ciborra, 1999a:78)”.

The second part (chapters four and five) of the study uses quantitative methods, namely education production function and learner flow-through models, to illustrate how the process of knowledge discovery in large data sets in the education management information system (EMIS) could be facilitated. These techniques show how knowledge in an existing education management information system could be used as a decision-support mechanism. Below follows a brief exposition of the methodology used to achieve the objectives of this study.

1.3 Research methodology

The methodology employed in the present study is described below in three parts namely, model development, model application and analysis.

1.3.1 Model development

The approach in this section was to review the subject literature on the different paradigms of ISD in order to facilitate the identification and description of the essential components and activities of a conceptual framework or model for ISD in this thesis. The establishing of an ISD model was based on three paradigms as reviewed in the subject literature, namely the traditional approach, systems theory and emerging approaches in ISD, such as bricolage, improvisation and sensemaking. The traditional paradigm usually describes ISD in terms of phases and as a process with a linear sequence and according to a systems development life cycle (SDLC). The general systems theory is about understanding the inputs, transformation processes and outputs in the context of a particular situation, such as in educational, political, economical, social and cultural contexts. Emerging approaches such as improvisation, bricolage and sensemaking could play an important role to enhance the information systems

development process and to emphasise alternative approaches. There is a growing body of literature on improvisation, sensemaking and bricolage that increasingly influence ISD (Ciborra, 1998; Ciborra, 1999a). Therefore, this study followed the same strategy as Bansler and Havn (2004: 631) that “adopted a sensemaking perspective to analyse the dynamics of this process and showed that improvisational action and bricolage (making do with the materials at hand) played a vital role in the development of the application”.

In developing a framework or model for ISD the intention is not to adopt or to streamline a particular approach, but rather to combine variants of the information systems approaches and utilise these complementarily in the production of an information systems development model. This study aims to illustrate that in establishing such a model it is necessary to develop an information system as a planned, deliberate and systematic activity. However, alternative approaches such as improvisation, bricolage and sensemaking could make positive contributions to the development process.

The nature of the study is such that there are various sections where empirical applications and analysis are required and form part of the ISD process model. Therefore, the literature review in this study will be discussed in the appropriate sections as a precursor for the empirical application.

1.3.2 Model application

The ISD model will subsequently be applied to the WCED as a real-life situation to describe and report on EMIS when using the survey method³. It should be noted that the information systems development in the WCED was not always a rational and normative process. It does not mean that there was a general failure to plan or that a chaotic situation prevailed. Rather, most of the development was based on the business needs and requirements of the department and then implemented without following the conventional controlled and planned method.

Because we live in a stochastic world we can say that the information systems process (education reality) is not so predictable and stable that everything can be planned before hand and nothing would ever change. Policy-making is never perfect (because the future can never be known), and because South Africa was in the middle of an unprecedented transition, the reaction of the EMIS team was optimal, because it tried to optimize at every margin. The information systems development process did not necessarily proceed from a conscious, grand, theoretical perspective based on bricolage and sensemaking. These theoretical frameworks were applied and interpreted on the ISD process afterwards. In hindsight however, most of the ISD activities that were intuitively implemented resonated with the emerging constructs as described in the subject literature.

³ The survey method is used when data gathering takes place through the dissemination and completion of questionnaires. The author was not responsible for any content development, but only used the questionnaire to develop the information system for data collection, data cleaning data storing and data retrieval purposes. However, since 2006 the WCED has implemented an electronic record system to keep track of individual learners through the education system.

Therefore, the Practice-in-Action approach in this study is used to describe how the day-to-day actions and practical experiences of role players contributed to the design, development, implementation, testing, maintenance and improvement of the education information system. Furthermore, this approach emphasises the importance of ISD in a “particular” situation such as education management information system (EMIS).

The focus on a “particular” or “local” ISD was based on what is referred to as a Practice-in-Action approach. Practice-in-Action in the context of this study refers to the process where key role players (actors) are active participants in the processes and directly involved in the development of an EMIS as part of their daily work (practice) in a particular or local context. Using the term “practice” denotes the work activities of the key role players and connotes the notion of doing (action). Goldkuhl (2005) concurs that information systems are contextual phenomena that are used by people in their work contexts. Therefore, according to Goldkuhl (2005: 236) the constituents of ‘workpractice’ refer to “constellations of actors, actions and objects. Such a constellation means that someone (actor) does something (action) and this means dealing with one or more objects”. This approach is deemed appropriate, as the author was a key participant and directly involved in the design, development and implementation of EMIS in the WCED.

The Practice-in-Action approach informed by the conceptual framework generated in the literature review was used as a lens for understanding ISD. In doing so, the author drew extensively from his extended professional association with the WCED and the experiences gleaned while developing and implementing a similar initiative. The study therefore posits that the experiences and the activities (Practice-in-Action) outlined in this study provide valuable

lessons, concurring with Heiskanen and Newman's (1997: 121) view that such practice-generated theories are acceptable for the development of information systems.

In the context of the Practice-in Action approach the author attempted to report on and explored the possibilities of how participants (users and developers) in information systems development are able to describe and analyse their experiences. This was attempted through the lens of sensemaking, improvisation and bricolage, that according to Weick (2001:9), is a process of converting the experiences of such role players into "an intelligible world". The Practice-in-Action approach relied on the kind of sensemaking action which Weick (1995) labels "retrospectiveness" (meaningful lived experience). In similar vein, Choo (1996:333) states that:

"sensemaking is done retrospectively since we cannot make sense of events and actions until they have occurred and we can then glance backward in time to construct their meaning. Current events are compared with past experience in order to construct meaning".

Such experiences were obtained when, as a manager of a team of developers and as practitioner, the author was directly involved from 1996 to 2005 in EMIS. During these years, the author was a key actor in the establishment, expansion and improvement of the system. It is from the perspective as a manager, developer and user of data that this thesis attempts to describe EMIS. The role and collaboration of the author with other role players (users, managers) was particularly significant in the development and improvement of the information system.

The Practice-in-Action came about through the reflection of the experiences gained and could inform other practitioners about EMIS and data utilization. These experiences were gleaned over the course of 9 years in a large government education department in varying roles such as:

- as an education practitioner
- as an information systems developer
- as a user of data
- in interactions with users of EMIS
- in interactions with education management.

The Practice-in-Action approach attempted to move away from the usually traditional and prescriptive approach of ISD to alternative and more flexible design approaches (bricolage, sensemaking and improvisation). Many researchers, when describing ISD use improvisational action, bricolage design and a sensemaking perspective, as a similar approach when referring to the notion of “practice in action”. The concept of action or involvement in practice and the effect on ISD is referred to by Ciborra (1996:23) as “learning by doing”, while Lanzara (1999) regarded it as “designing in action” and Moorman and Miner (1998a: 4) suggested the term “thinking in the midst of action”. Ciborra (1999a:78) further described it as “thinking and action [that] emerge simultaneously and on the spur of the moment”. Bansler and Havn (2004: 633) state that “within organisations, it could be described as the conception of action as it unfolds, drawing on available material, cognitive, affective and social resources”.

1.3.3 Analysis

The second part (chapters four and five) of this study applied quantitative methods of data analysis to illustrate how value can be added to datasets yielded by ISD. Two key statistical methods were used to report on the quality and efficiency of the education system of the WCED:

The first method utilised was the education production function approach. The education production function approach, a mathematical expression of the relationship between inputs and outputs, has become an important methodology for determining the effect of various resources (inputs) on education performance (outcomes). The growing literature in South Africa where education production functions are used in empirical studies is of particular importance for this study (Crouch & Mabogoane, 1998a; Crouch & Mabogoane, 1998b; Case & Deaton, 1996; Fiske & Ladd, 2002, 2003; Gustafsson, 2005; Hosking, 1994; Van der Berg, 2005a; Van der Berg & Burger, 2003; Van der Berg & Louw, 2006). Most of these studies analysed the contribution of education inputs to effective schooling. This underscored the main argument of the present study, which explored the importance of education production functions for data analysis in order to improve management efficiency with regard to the input of resources.

The second method that was applied to analyse the flow of learners through the education system of the Western Cape, was the Grade Transition Model (GTM) or Reconstructed Cohort Method and the Learner Attainment Model. The basic principle of the Reconstructed Cohort Method is that one can construct (estimate) the flow of an entrance cohort of learners through a particular system and predict the eventual outcomes at the end of the cycle. The learner

attainment method was used to simulate dropouts in the Western Cape using the 2001 census data (Statistics South Africa, 2001).

1.4 Overview and structure of the thesis

Figure 1 graphically represents the outline of the thesis.

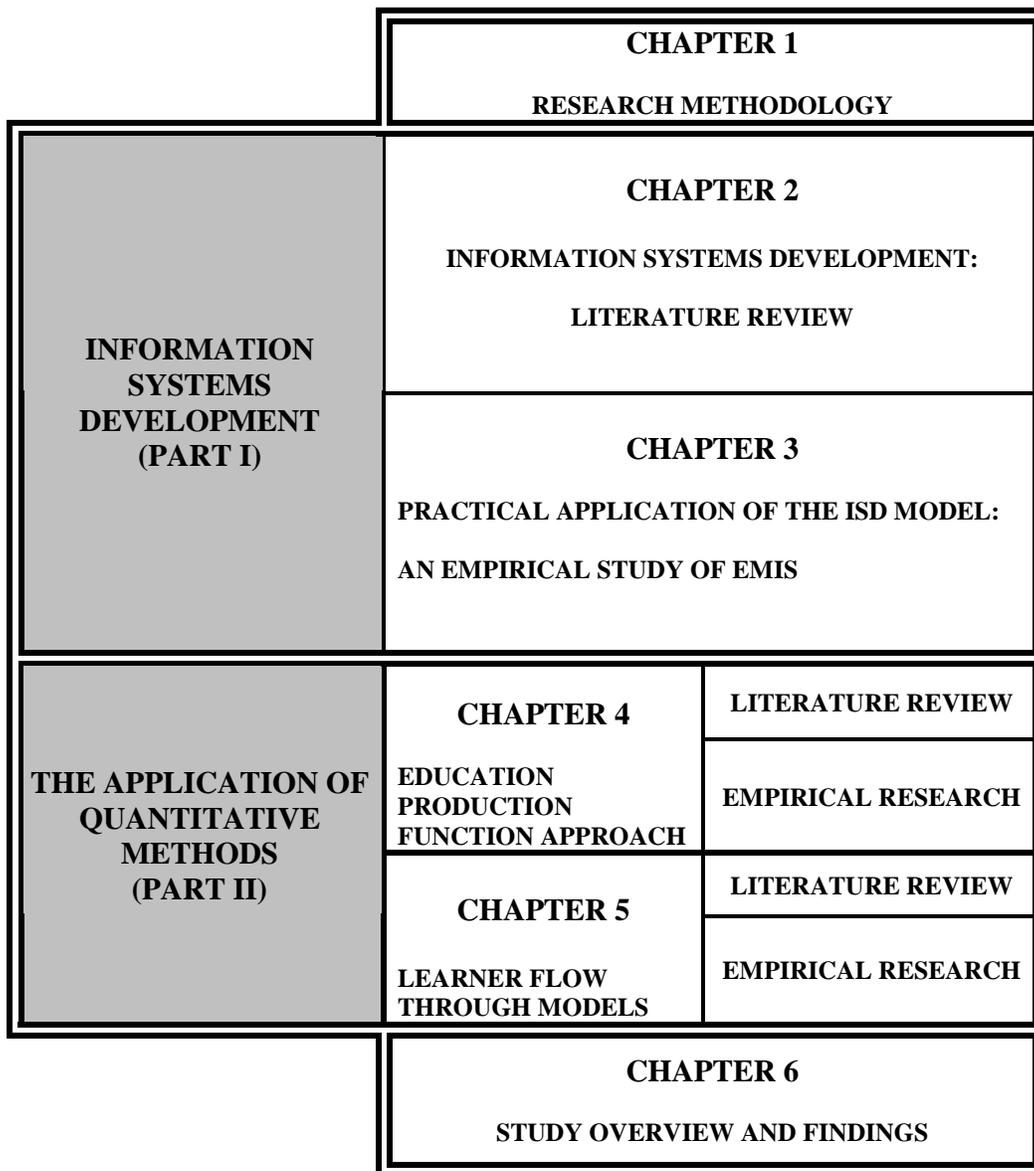


Figure 1: Structure of the thesis

From Figure 1 it becomes evident that there are two major thematic components to the thesis: The first theme (part 1) addresses information systems development and data management. The second theme addresses sensemaking and the role of quantitative techniques in extracting knowledge that would otherwise remain hidden in data sets.

The thesis is organized into six chapters and is structured as follows: Chapter one locates the study by presenting the rationale for the study. It also summarizes the research methodology employed to answer the aims and objectives of the study.

In terms of the first part of the study (chapters two and three), information systems development is outlined in two steps, namely the literature review and an empirical description of the education management information system. Each step is described in separate chapters. These next two chapters address Information Systems Development in order to make quality data available for statistical analysis purposes.

Chapter two entails a review of the literature on the establishment of an Information Systems Development (ISD) model. From this review an ISD model will be developed that provides the conceptual framework including the major components and activities for an ISD model, for the present study.

Chapter three presents an empirical application of this developed ISD model on EMIS and underscore the role of information systems in everyday practice. It describes the working of EMIS when using the survey method and explains the development of a computer application

and how data ought to be collected, stored, retrieved, verified and cleaned for management information.

The next two chapters describe the data analysis and results of the data sets emerging from the EMIS processes resultant from the ISD model, and used quantitative methods to discover knowledge embedded in the information systems.

Chapter four describes the statistical techniques used in, among other things, the education production function. The chapter also includes the findings of these techniques, as well as a discussion thereof.

Chapter five describes quantitative techniques to measure the effect of learner dropout and repetition on internal efficiency of the WCED by means of learner flow through models.

Chapter six summarises the thesis with a diagram and illustrates the EMIS procedures and activities through a process model that encapsulates the systems theory principles of input, transformation processing, output, feedback and control. It further indicates that quantitative analysis is more than just a numbers exercise. In this chapter the study offers a critical perspective on the quantitative analysis and shows that through sensemaking further research and investigation could become essential.

CHAPTER TWO

INFORMATION SYSTEMS DEVELOPMENT: LITERATURE REVIEW

As mentioned in Chapter 1, the literature review in this study will not follow a conventional pattern by grouping it all together. Literature pertaining to particular foci will be discussed in the appropriate chapters as a precursor for the empirical application and analysis. This chapter reviews three approaches to information systems development (ISD) in order to inform a conceptual model of ISD for this thesis. This will be used to develop an education management information system (EMIS) that will be tested empirically in the next chapter.

Information is a necessary resource, produced by information systems and is a key building block to the management and decision-making in any organisation. According to Ahituv and Neumann (1986: 2) “if [it is] properly developed, managed, and used, information systems can provide the most cost-effective resource to the organisation”. O’Brien (1991:16) defines an information system “as a set of people, procedures, and resources that collects, transforms, and disseminates information in an organisation”. Henriksen (2003: iii) further refers to an “information system” as a system that “can include several integrated information technologies as well as organisational use and maintenance practices that collectively comprise a socio-technical phenomenon”.

Therefore, an information system in the context of this study includes the technology, the people, processes and information. Clearly, information systems development (ISD) is more than just a rational representation of reality. It makes provision for representing reality

beyond the measurable, the visible and the world of reason. This chapter attempts to establish an information systems development model that could provide a fundamental conceptual framework of the major components and activities. The subject literature review on the approaches of information systems development is a key step towards the identification and description of the essential components and activities of the conceptual framework.

The ensuing literature review on information systems development explicates how the emerging technology changes over time have influenced the design thereof. The literature review culminates in devising a conceptual framework that focused on the impact of technology changes on information systems development, using improvisation, bricolage and sensemaking.

2.1 Traditional view of information systems development

According to the functional approach (also called the traditional approach) ISD is a planned and rational activity, carried out in a systematic, organized and methodical manner (Ahituv & Neumann, 1986; Bell & Wood-Harper, 2003; Benyon, 1990; Fitzgerald, 1998). The development is usually described according to phases in the development life cycle and has a strong technology focus. Serafeimidis and Smithson (2003: 252) concur with such an approach and assert that the traditional conception deals with an information system (IS) “as if it existed in isolation from its human and organisational components and effects. It also placed excessive emphasis on the technological and accounting/financial aspects at the expense of the organisational and social aspects”.

Authors such as Hirschheim and Klein (1989), Iivari and Hirschheim (1996), Avgerou (2001) and Checkland (1999) have all explored alternative approaches in an attempt to accommodate social and cultural factors besides the technological in information systems development. Below follows a brief outline of the approaches to ISD in the traditional paradigm.

2.1.1 Linear design approach: Information systems development life cycle (ISDLC)

Beynon-Davies and Williams (2003: 31) refer to ISD as the “structured methods that emerged during the 1980’s and initially used a linear model of the development process. Clear phases are identified with clear inputs and outputs from each phase”.

Ahituv and Neuman (1984) were strong proponents of the information systems life cycle (ISDLC) approach in information systems development. Although they agree that ISDLC is “usually treated as a rigid sequence of activities” they propose a flexible approach that

“is now the cornerstone in the management information systems (MIS) literature and a hallmark of every development effort, implying that no MIS activity should be carried on without imposing strict ISDLC procedures, practices and methods on system developers” (Ahituv & Neuman, 1984:69).

They acknowledge that “in practice, however, development processes are not that rigid” (Ahituv & Neuman, 1984: 69). Ahituv and Neuman (1986: 221) pointed out that one must “guard against thinking of the information development process as a linear process, but [that] in practice it is rather an iterative process”. The general scheme for the ISDLC is virtually the same in all research literature (Ahituv & Neuman, 1984, 1986; Kaleremo & Rissanen, 2002; Laudon & Laudon, 1998; Necco, Gordon & Tsai, 1987). According to Ahituv and Neuman

(1984: 69) “it typically contains three major phases which consist of several steps each, namely,

- 1) Definition Phase: a) Preliminary analysis, b) Feasibility study, c) Information analysis, d) System design;
- 2) Construction Phase: a) Programming, b) Development of procedures;
- 3) Implementation Phase: Training, conversion and testing”

(cf. Ahituv & Neuman, 1986: chapter 8):

Ahituv and Neuman (1984:70) describe the factors that affect the ISDLC so that it is not “viewed as a rigid sequence of activities but rather as a more general framework from which the most adequate development plan can be derived for a particular project.” These factors include organisational scope (the number of organisational units connecting to the project), organisation maturity (experience of users in developing information systems), information system policy, structuredness level and technological environment. These factors are “derived from the user and the environment, and also from the nature of the development process” (Ahituv & Neuman, 1984:77). They are clear that the ISDLC should not “be seen as a sequence of steps (though some of them may be iterated), but as a breakdown of each step into various dimensions affected by the factors” (Ahituv & Neuman, 1984: 77).

Laudon and Laudon (1998: 425) maintain that building information systems according to the traditional systems life cycle is “still the predominant method for building large and medium mainframe-based systems today” and based on the author’s experiences it appears as if it is still prevailing practice in 2006. According to these authors

“traditional systems life cycle is the oldest methodology for building an information system. It consists of six stages (project definition, systems study, design,

programming, installation, and post-implementation) that must be completed sequentially” (Laudon & Laudon, 1998: 425).

This linear model approach (sometimes called the “waterfall approach” see Kalermo & Rissanen, 2002) is evident in the work of various researchers (Ahituv & Neuman, 1984; Mahmood, 1987; Necco, Gordon & Tsai, 1987). Necco et al. (1987: 472) agree with Laudon and Laudon (1998) that “organisations will continue to develop their computer based information systems within the framework of a systems development life cycle”. The basic conclusion of such an approach is that distinct phases are identified with specific and clear inputs and outputs for each phase. Numerous researchers in the field of information systems documented limitations of the traditional methods. For example, Tolvaven (cited by Kalermo & Rissanen, 2002:38) asserts that “the improvements caused by methods are modest, methods are considered labour-intensive, methods are difficult to use and learn, and methods have poorly defined and ambiguous concepts”. Truex, Baskerville and Travis (2000) suggest the amethodical systems development that could better fulfil the demands of information systems development set by the emergent organisations and by the changing technological environments. These emergent approaches will be dealt with in the latter part of this chapter.

2.1.2 Management framework : Information Orientation (IO)

It is important to note that Marchand, Kettinger and Rollins (2002:76) use an improved information system life cycle model. To counter the linear and one dimensional approach in ISD, Marchand (2001:1) takes this argument a step further when he refers to a “technology-centric mindset” in organisations where managers “have seen the management of information only within the context of their company’s information technology (IT)”. Marchand (2001:2)

states that there is a well-known saying in management “you can’t manage what you can’t measure” and then comes up with an information approach what he calls “making the invisible visible: seeing information capabilities”. This management framework is called “Information Orientation” (Marchand, 2001; Marchand et al., 2002) to help managers see and actively manage effective information use in organisations. This information orientation is based on the synergy across three unique information capabilities, namely information technology practices, information management practices and information behaviour practices.

Marchand et al. (2002: 1) describe information technology practices as the “capability of a company to effectively manage information technology applications and infrastructure to support operations, business processes, managerial decision making, and innovation”. Information management practices for them are the “capabilities of a company to manage information effectively over the life cycle of information use including sensing, collecting, organising, processing, and maintaining information” (Marchand et al, 2002:1). They refer to it as the “circuitous set of phases” (Marchand et al., 2002:75). They however introduced information practices as a life cycle with interactive phases with sensing as a fifth new phase.

They further explain information behaviours and values as the “capability of a company to instil and promote behaviours and values in its people for effective use of information” (Marchand et al., 2002:1). They emphasise that the key behaviours and values that lead to effective information use by people are integrity, formality, control, transparency, sharing and proactiveness and should be followed, encouraged and supported by managers (Marchand et al., 2002:102).

Information orientation is about the effective use of information in organisations. In their review of the history of these three schools of thought, namely the history and evolution of IT, Information Management, Behaviour and Control, they realised that there “have been so few real interactions or connections between the three” (Marchand et al., 2002:46). In his discussion of “information orientation: the link to business performance”, Marchand (2001:2) makes it clear that “IT practices alone do not result in superior business performance. But our research does suggest that all three Information Capabilities - IT practices, management of information and information behaviours - must be strong and working together if superior business performance is to be achieved”. Marchand (2001:2) states that “these three information capabilities together – which we call Information Orientation or IO - thus provide a critical link to business performance and can be used as a new management measure”.

Marchand et al. (2002:47) concur that “if we manage what we measure, then it is not surprising that the management of the interactions between people, information, and IT to improve business performance is not well understood by executives, since effective information use in companies is seldom measured”.

Marchand et al. (2002) in their information orientation model, in determining if effective use of information leads to better business performance, adopts a statistical analysis approach. Their study uses confirmatory factor analysis (CFA) and structural equation modeling (SEM) to analyse the data. They successfully “determine whether the presence of first order factors (IT practices, information management practices, information behaviours and values) or a second order factor (IO) best predicts an increase in a business performance criterion”

Marchand (2001:2) (cf. Confirmatory Factor Model of Information Orientation in (Marchand et al., 2002: 143).

In a review of the Information Orientation Model, Marchand (2001:3) summarises it as follows: “From these results we confirmed the existence of three ‘information capabilities’ consisting of fifteen specific competencies associated with effective information use”.

Davenport and Prusak (1997:4) have a similar approach as the Information Orientation model; however, they regard information management more than just a rational and quantifiable process. They refer to it “as information ecology” that emphasises an organisation’s entire information environment that includes

“the organisations beliefs and values about information (culture); how people actually use information and what they do with it (behaviour and work processes); the pitfalls that can interfere with information sharing (politics); and what information systems are already in place (technology)” (Davenport & Prusak, 1997: 4).

Davenport and Prusak (1997:5) are also opposed to the linear approach of seeing information only as the quantifiable positivistic approach. They suggest that an ecological approach can be more comprehensive and includes information as something that is not easily stored on computers and which is different from “data”. The ecological approach for them indicates that the more complex an information model becomes, the less useful it will be. Information, according to this approach can take on many meanings in an organisation and that technology is only one component of the information environment and often not the right way to create change.

The information life cycle is not a sequential process where pre-defined phases including certain activities follow one after the other. If it was so then it would have been a very narrow approach that would limit the ISD process.

2.1.3 Functionalism

Hirschheim and Klein (1989) identify four major kinds of systems development approaches (functionalism, social relativism, radical structuralism and neohumanism) and discuss how they lead to different outcomes. Hirschheim and Klein (1989:1201) discuss these approaches as “paradigms” which they define as “assumptions about knowledge and how to acquire it, and about the physical and social world”. According to these authors the functionalist paradigm is in line with the traditional and rational approach of information system development as discussed in paragraph 2.1. They delineate and map it out as a paradigm “concerned with providing explanations of the status quo, social order, social integration, consensus, need satisfaction, and rational choice” (Hirschheim & Klein, 1989:1201). They convey functionalism as a story with management, the system developer and users as the key actors. The story-line (narrative) supports the notion of “rational organisational operation” (Hirschheim & Klein, 1989: 1203). The idea of observable reality in information system development “proceeds through the application of ‘naive realism’- the notion that the validity of system specifications, data models, decision models, and system output can be established by checking if they correspond to reality. Reality consists of objects, properties, processes that are directly observable” (Hirschheim and Klein, 1989:1203). Henriksen (2003:21) underscores the positivistic notion of the functionalism paradigm of Hirschheim and Klein, when he refers to it as “organisational reality made up of objects, properties, and orderly goal-

oriented patterns that are directly observable and predictable. The world can (and should) be engineered and mathematically modelled in a rational and scientific manner.” The underlying assumptions of functionalism are that epistemology is based on positivism in that “the developer gains knowledge about the organisation by searching for measurable cause-effect relationships” (Hirschheim & Klein, 1989: 1203). The entire notion of functionalism presented by them stipulates that it is a technical process that makes system development more formal and rational, “placing less reliance on human intuition, judgement, and politics” (Hirschheim & Klein, 1989: 1203). The entire systems design is primarily a technical process. The applicability to education policy and education reality (translated into quantitative, financial goals and systems performance characteristics) “allow system objectives to be derived in an objective, verifiable, and rational way” (Hirschheim & Klein, 1989: 1203). From a social and economic policy perspective, according to Hirschheim and Klein (1989: 1203), “it is therefore unwise to question the legitimacy of management in deciding system objectives”. The epistemological assumptions of the functional approach is predictable and quantifiable and based on positivism – observation, experiment is accepted as the only way of acquiring knowledge, all other knowledge is subjective. The sensemaking process is offered as an alternative with the focus on human actors and the social side of development.

2.2 The application of the systems theory approach in information systems development

Systems thinking posits that an information system be viewed as a system, with data as the input and data analysis processes that transform them into information and knowledge as the output. The basic premise is that the general systems theory concepts underlie the information systems development (Ahituv & Neumann, 1986; Bell & Wood-Harper, 2003; Benyon, 1990;

Checkland, 1999; Laudon & Laudon, 1995, 1998; O'Brien, 1991). Truex, Baskerville and Travis (2000: 56) confirm this systems thinking paradigm in ISD when they maintain, "information systems development and information systems development method are completely merged in systems development literature". Khazanchi and Munkvold (2000: 31) concur when they define a system as "a collection of interrelated components that work together for a common purpose" (cf. Benyon, 1990: 2; O'Brien, 1991:11) and refer to an "information system as a collection of interrelated components (hardware, software, procedures, people and databases) that work together". Laudon and Laudon (1998) take it a step further when they emphasize that an information system "consists of three basic activities – input, processing, and output – that transform raw data into useful information. Feedback is output that is fed back to appropriate people or activities" (Laudon & Laudon, 1998: 6). Although systems theory approach is highly technology focused, it is included in this study to give structure to the information systems development model. It was in this regard that O'Brien (1991: 11) declares that "the knowledge of systems concepts is vital to a proper understanding of the technology, applications, development, and management of information systems".

Numerous authors have examined how the general systems theory approach has been applied to ISD (Ahituv & Neuman, 1986; Bell & Wood-Harper, 2003; Benyon, 1990; Checkland, 1999; Lundeberg, GoldKuhl & Nilsson, 1981; O'Brien, 1991; Oz, 2002; Skidmore & Eva, 2004).

ISD is not a one dimensional and linear action where only technical considerations are taken into account. Human perspectives (culture, behaviour, values, experience and emotions) have

a major impact on the information systems analysis, design, development, implementation and maintenance. Warren and Adman (1999: 224) regard the systems-based approaches as an alternative to information systems development, because of the inability of the “traditional, ‘rational’, ‘hard models’ to cope with the increasingly complex and turbulent organisational world, with its conflicting human perspectives”.

There is a vast body of literature that outlines the general systems theory (Ahituv & Neuman, 1986; Bell & Wood-Harper, 2003; Benyon, 1990; Checkland, 1999; Jackson, 2004; von Bertalanffy, 1968). In this approach the organisation is conceptualised as a total system, therefore understanding the systemic inputs, transformation processes, the outputs, the numerous linkages between system objects and interrelated and interacting subsystems in context of the organisational operations are essential to comprehend the issues relating to information systems development.

Ahituv and Neumann (1986:76) apply the systems approach in ISD and describe the basic components of the general systems theory as was first formulated by von Bertalanffy (1968). They suggest that systems approach is also the cornerstone of the life cycle phases in the process of developing an information system. They use the systems approach to compare and analyse the three phases of information systems development. According to them information system analysis is the first phase and includes 1) the defining of the problem and 2) the gathering of relevant data to solve the problem development (Ahituv & Neuman, 1986:78). In typical systems approach manner, when defining the problem it is important to identify the environment of the system and determining the inputs to the system, and the processes that transform these inputs into outputs, and the attributes of the outputs. In the data gathering step,

Ahituv and Neuman (1986:78) again confirm the systems approach in the development life cycle, when they state “that a problem is understood and studied by collecting data about its goals, inputs, outputs, environment, transformation processes, and constraints.

The second phase is the design phase that includes 1) identifying alternative solutions, 2) evaluating the cost effectiveness of alternatives, and 3) selecting the best solution. The third phase is the system implementation that includes the implementing and monitoring of the selected solution (Ahituv and Neuman, 1986:78).

Bell and Wood-Harper (2003: 1) emphasise the systems theory principles when they described the information system in terms of the transformation of inputs (data) into information and eventually into knowledge. They refer to the stereotypical definition of positivism/reductionism, which they regard as “fixed, knowable, measurable and, therefore predictable. If it is not these things, then it is not worth knowing” (Bell and Wood-Harper, 2003:21).

Similarly they applied the systems approach on the phases of ISD. They proposed the following steps in designing the information system which are closely aligned with the phases identified by Ahituv and Neumann (1986), namely 1) discover what the information problem is, 2) discover what is the setting for the problem, 3) what resources and constraints are evident, 4) what are the major information components of the problem, 5) structure the problem into a model, 6) design model solutions for the problem, 7) test and cost the model, 8) implement the model as appropriate, and 9) monitor and evaluate the result (Bell & Wood-Harper, 2003 : 15).

O’Brien (1991:11-12) also applies the systems concept when he views the information system “as a system that accepts *data resources* as *input* and *processes* them into *information*

products as output". He further emphasises the systems approach by including two more systems characteristics, feedback and control. For O'Brien (1991:11) "feedback is data or information concerning the performance of a system" and "control is a major system function that monitors and evaluates feedback to determine whether the system is moving toward the achievement of its goal". The following sections explore the shortcomings of the previous approaches.

2.3 Information systems development in the knowledge economy

An alternative epistemological approach to the traditional and functional approach to ISD, according to Hirschheim and Klein (1989:1205), "is that of anti-positivism reflecting the belief that the search for causal, empirical explanations for social phenomena is misguided and should be replaced by sense-making". Henriksen (2003:23) agrees with this epistemological approach when he states

"reality is always viewed as socially constructed through knowledgeable human interaction and creative sense-making processes. Ideally the systems developer engages in and facilitates sense-making processes to produce system objectives and organisational changes that are desirable and feasible for both management and users".

Laudon and Laudon (1998: 8) concur when they state that "an information system is an integral part of an organisation and is a product of three components: technology, organisations, and people". The following discussion outlines the emergent and transient approaches as alternatives to the traditional development approaches.

2.3.1 Information systems development is more than a rational activity

Ciborra (1996) has the understanding that beyond the technical world is another reality that is difficult to quantify and to observe. He therefore started his argument distinguishing between data, information and knowledge. **Data** for Ciborra (1996:30) “record signs and observations in or on a medium” and “**information** shapes the meaning of data in relation to a specific context of action or speech”. **Knowledge** on the other hand is the “capability of an individual or an organisation to relate complex structures of information to new contexts of action” (Ciborra, 1996: 30). “In order to turn knowledge into information and information into data, one has to make explicit what has hitherto been tacit in competence of those having the knowledge and in the situation where such knowledge was put to work” (Dahlblom & Mathiassen, 1991, cited in Ciborra, 1996:30). This remains the challenge in all information systems.

Ciborra (1996:21) outlines “four different ways of looking at rational processes of production, communication and accumulation of knowledge and information in organisations”.

Unlimited rationality and fully accessible knowledge is “according to this view, individual and organisational knowledge is fully accessible to an expert analyst, and structured methods have been put forward to capture and include it in formalised (computerised) routines” (Ciborra, 1996: 21). Ciborra (1996: 22) indicates that according to this view “the management of an organisation focuses on the concepts of planning and control, and information systems support such a central concern”.

Limited rationality and tacit knowledge, according to Ciborra (1996: 23) “is to recall the tacit nature of individual and organisational knowledge”. He stresses that “information technology needs to come to terms with the tacit nature of knowledge in order to find a realistic role in supporting human work” (Ciborra, 1996: 23). Otherwise, information systems will remain limited and incomplete for decision-maker. The objective is to make the invisible visible (Marchand, 2001). Ciborra (1996: 23) emphasises that “adaptation of individual and organisation behaviour occurs through searching and by learning and doing” and in the same vein mentions that “of a similar kind is learning by using”. The people (designers and users) are key in system development, because

“user-participation strategies and the socio-technical approach are based on the notion of limited rationality of the analyst and the recognition of the superior knowledge of the users have regarding the organisational routines as they are carried out in practice” (Ciborra, 1996: 24).

We must remember that in the development life cycle of a system different people perform different functions (actions). Everyone adds a different dimension to the information system, based on their involvement, according to their own interpretation, behaviour and knowledge.

Strategic rationality and opportunistic information happen when there is no full collaboration and teamwork among members in an organisation and someone strategically uses tacit knowledge. Here the old notion of knowledge is power applies instead of the alternative approach in the information society, namely that knowledge sharing is power. Opportunistic information processing then also deprives the information system from knowledge, because it

“consists of a selective disclosure or deliberate distortion of the factual data to which each party has unique access during transacting”. Understandably such behaviour and

the “way knowledge, information and data are processed by their members has far-reaching implications for the role of systems and their design” (Ciborra, 1996: 26).

Adaptive and limited learning is an important aspect to consider in information systems development. According to Ciborra (1996: 27) “*organisational change*” refers to “the organisational process of interest [and] is not the execution of routines but their transformation, that may be required by the turbulent environment” (Ciborra, 1996: 27). Here Ciborra (1996:29) employs the notion of formative context which he defines as “the set of institutional arrangements, cognitive frames and imageries that actors bring and routinely enact in a situation”. According to him “a formative context has far-reaching influences, for it constitutes the background condition for action, enforcing cognitive and practical constraints, giving direction and meaning, and setting the range of opportunities for undertaking action” (Ciborra, 1996:29). Based on the aforementioned views, Ciborra (1996) suggests that information systems should be developed at two distinct levels, namely one of formed routines and one of the formative context. “The first is concerned with restructuring of routines, transactions and the boundaries that define the modes of production, co-ordination and division of labour between the members of the organisation. The second one deals with the frames, beliefs and consolidated ways of thinking of those who work in the old, or the new, computer-supported organisation” (Ciborra, 1996:31).

Davenport and Prusak (1997:16) argue along similar lines when they indicate that information management includes four different approaches, namely “unstructured information, intellectual capital or knowledge, structured information on paper and structured information on computers”. They focus on the last of these approaches, namely management of structured information on computers. They claim that “IT professionals have mainly concerned

themselves with managing computer-based data, rather than information more broadly defined. This has its uses; yet the heavy emphasis on what can be represented on a computer has often ended up denying that unstructured information or knowledge has any value” (Davenport & Prusak, 1997:21).

Ciborra (1998:6) affirms the notion of emerging technologies in ISD when he asserts that the “internet has emerged as flexible infrastructure outside any strategic master plan: it allows us to share knowledge in ways our textbooks had not even imagined”. Meaning we cannot continue to do ISD in a predefined, structured and traditional way, according to a specific method. What Ciborra (1998: 6) makes clear is that information systems design methods “may be the most diffused methodology on earth accompanying the introduction of a new technology, but they work only in part”. He emphasises tacit knowledge as part of the information systems development and points out that we cannot use the same representations as the physical systems (the computer) in information system. Ciborra (1998:8) highlights the importance of this when he states that “we tend to forget the role of human choice behind the technical artefacts, and study the user side of information systems by adopting the methods of the natural sciences”. He suggests that the origin in the crisis in information systems development is due to the “dominating style of research based on the scientific paradigm” (Ciborra, 1998:15). The current technology “the internet, the emergence of global IT infrastructure, all seem to suggest that technology may require us speak another language, less formal and structured” (Ciborra, 1998: 16).

2.3.2 Revisiting methods and methodology in growing and emerging organisations

Truex et al. (2000: 53) questions whether “information systems development methods really describe ongoing systems development practice”. They are further concerned about the relevance of information systems development methods to function as “frameworks, formulas or templates for building successful information systems” (Truex et al., 2000: 53). Their clear objective is to revisit the concept of method without causing chaos in the building of ISD. That brings them to the term *amethodical* which means “management and orchestration of systems development without a predefined sequence, control, rationality, or claims to universality” (Truex et al., 2000: 54). They compare the assumptions and ideals of methodical and amethodical ISD and describe methodical ISD as a “managed, controlled process”, while assumption of the amethodical development is a “random, opportunistic process”. For them ISD according to a method is “a linear, sequential process” and amethodical is “simultaneous and overlapping and there are no gaps”. A method assumes that ISD is “rational, determined, and goal-driven”, while amethodical processes are “negotiated, compromised and capricious” (Truex et al., 2000:54).

Fitzgerald (1998: 318) concurs with this approach when he clearly states that “system development is not actually an orderly, rational process, but most systems development methodologies treat it as such, with major emphasis on technical rationality at the expense of social aspects”. Fitzgerald (1998: 319) substantiates his argument by referring to the changing nature of the business environment (see also Ciborra 1998, Lanzara 1999, Necco et al., 1987) with the explicit consideration that the traditional “life-cycle approaches that resulted in eventual delivery of systems after several years are no longer appropriate”. His reasoning

behind this is that many of the current and probably outdated methodologies are derived from “practices and concepts relevant to old organisational environment, and there is a need to reconsider their role in today’s environment” (Fitzgerald, 1998: 326). Today’s environment includes technological advances such as the Internet, World Wide Web, mobile technology, wireless technology, and more.

Necco et al. (1987) emphasised the change in ISD because of the dramatic change in hardware technology in terms of improved capabilities and lower costs. They listed the reasons why information systems have failed over the last few years, namely, “1) system development did not meet the user’s requirements, 2) system development exceeded the estimated time schedule 3) the cost of the system developed was too great compared to the value of the benefits realised and 4) the system developed was excessively costly and difficult to maintain” (Necco et al., 1987:461).

Alatalo, Oinas-Kukkonen, Kurkela and Siponen (2002) refer to this phenomenon as information systems development in emergent organisations. They cited the ideas of Truex et al. (1999) that perceived the process for designing stable information systems (IS) as different from that of emergent IS with goals contradictory to each other. The goals in stable organisations are: “1) economic advantages of lengthy analysis, 2) user satisfaction, 3) abstract requirements, 4) complete and unambiguous specifications and 5) new system projects as achievements” (Alatalo et al., 2002: 2). According to Truex et al. (Alatalo et al., 2002: 2), the goals of ISD in emergent organisations are “1) always analysis, 2) dynamic requirements negotiations, 3) incomplete, usefully ambiguous specifications and 4) continuous development”.

2.3.3 A sensemaking perspective and the important role of improvisation and bricolage as emerging concepts in information systems development

Improvisation, sensemaking and bricolage (“making do with the materials at hand”) are emerging concepts that are included and taken into account in ISD (Bansler & Havn, 2004; Barrett, 1998; Büscher, Gill, Mogensen & Shapiro, 2001; Ciborra, 1998, 1999a; Hirschheim & Klein, 1989; Kalermo & Rissanen, 2002; Lanzara, 1999; Louridas, 1999; Miner, Bassof & Moorman, 2001; Moorman & Miner, 1998a, 1998b; Peplowski, 1998; ; Serafeimidis & Smithson , 2003; Weick, 1995, 1998, 2001).

Bansler and Havn (2004: 632) refer to it as a “perspective that posits improvisation and emergent rather than methodical behaviour and planned change as fundamental aspects of ISD in organisations”. They underline that “information systems are no longer stable, discrete entities, but part of elaborate networks and ‘information infrastructures’ that are subject to constant adjustment and adaptation” (Ciborra et al 2000 as cited in Bansler & Havn, 2004: 631). Iivari and Hirschheim (1996: 569) conclude “that the newer, emerging approaches arose because the established traditions failed to deliver on their stated promises. Indeed, the emerging approaches have adopted very different sets of assumptions in an attempt to overcome the limitations of the ‘orthodox’ approaches”. According to them “the evolution of ISD towards the understanding of the social nature of information systems and more intersubjective view of information requirements is partly explained by technological developments and the evolution of application areas to cover knowledge work” (Iivari & Hirschheim, 1996: 571). They highlight that “most of the emerging approaches emphasise the social nature of information systems” (Iivari & Hirschheim, 1996: 551). According to Iivari

and Hirschheim (1996) the organisational role of information is not a linear and one-dimensional approach. They made provision for all the levels in an organisation when they analysed the organisational role of information and distinguished between three alternatives: a technical, sociotechnical, and social view.

“A technical view regards an information system predominantly as a technical artefact, and assumes that its connections with its organizational environment can be reduced to well-defined inputs and outputs and ergonomic interface questions. The social view considers an information system primarily as an organizational and social system; an information system is seen as an integral, constitutive part of organizational communication, control, coordination, cooperation and work arrangements and not only as a separate support system for these organizational activities. In the more theoretical terms of structuration theory, an information system as a social system can be characterized as an embodiment of interpretive schemas, facilities of coordination and organizational/social norms. The sociotechnical view is based on the assumption of interdependent subsystems, the technical subsystem and the social subsystem which are designed jointly” (Iivari & Hirschheim, 1996: 553).

According to Lyyntin and Robey (1999: 85)

“the practice of information systems development has undergone a radical transformation during the last decade. Advancing technologies have encouraged a migration away from the traditional, life cycle methods of development toward more flexible and dynamic approaches in which reusable components are assembled into working systems in a radically shorter time.”

2.3.3.1 Improvisation

Bansler and Havn (2004) discussed the role of extemporaneous action and bricolage in designing and implementing information systems in organisations. They are clear about their approach when they state that

“when adherence to methods is taken for granted, activities and situations that do not fit within a methodical frame become marginalised and practically invisible, e.g. ISD is subject to fortuity, circumstance, human whims, talents and the personal goals of the managers, designers and users involved” (Bansler & Havn, 2004: 632).

Based on this approach they elucidate the role of improvisation in organisations using jazz music as a metaphor for ISD. “Within organisations, it could be described as the conception of action as it unfolds, drawing on available material, cognitive, affective and social resources” (Bansler & Havn, 2004: 633). These principles form an integral part of the fundamental conceptual framework. They ground their argument on findings from a longitudinal case study of the development of a Web-based groupware system in a large multinational corporation” (Bansler & Havn, 2004: 632). For them improvisation is firstly “deliberate, meaning that it is the result of intentional efforts on the behalf of the organisation and/or any of its members”. Secondly, improvisation “deals with the unforeseen; it works without a prior plan and without blueprints and methods“(Bansler & Havn, 2004:633).

Finally they state that “improvisation implies the pre-existence of a set of resources, e.g. a plan of action, tools and technologies, knowledge or a social structure, upon which variations can be built” (Bansler & Havn, 2004:633). They describe improvisation as part of the development process and emphasise that the development of the system was ‘not guided by a preconceived plan or a systematic method’ (Bansler & Havn, 2004: 642). The people involved

(users and developers) “depended on improvisation and extemporaneous action in order to cope with unexpected problems, unanticipated opportunities, multiple meanings, and transient organisational requirements” (Bansler & Havn, 2004: 642). They were very specific about the role of improvisation in the development process. For them “ISD is more an act of interpretation rather than an act of decision-making”. There are always things happening that require immediate attention and action.

Ciborra (1999a: 78) further elaborates on improvisation as a highly grounded process that happens “intentionally but extemporaneous, that is, happening almost unexpectedly (‘ex tempore – outside the flow of time), and with little cause or relationship” that as he puts it “opens up alternative approaches to cope with time in business”. Moorman and Miner (1998a: 3) in describing the importance of improvisation in new product development cited Irby (1992: 630) who referred to it as “thinking in the midst of action”.

The purpose of the study of Weick (1998: 543) was to improve on the way people talk about organizational improvisation. In doing so he describes jazz improvisation. Here he reviews holistically several definitions to indicate what is happening when people improvise. He continues his argument about improvisation by taking a closer look at degrees of improvisation, forms of improvisation and the cognition of improvisation. He concludes his argument with the implications of the above on theory and practice. Similarly, Frank Barrett (1998) uses jazz music as a metaphor for improvisation. He states that “improvisation involves exploring, continual experimenting, tinkering with possibilities without knowing where one’s queries will lead or how action will unfold” (Barrett, 1998: 606).

2.3.3.2 Bricolage Design

Another emerging concept in information systems development is, *Bricolage*. Lanzara (1999) is well aware of the challenges in coping with the ever-changing technology when he describes the relationship between transient constructs and persistent structures. For Lanzara (1999: 332) “designing in action is a practical, situated context-sensitive mode of design that feeds on the dynamic tension between the requirements of change and stability”. Therefore a system cannot be designed as a complete entity. Users, designers and managers need to stay involved. Information systems evolve over time and is not analysed, designed and developed as a complete product. The business world is too dynamic and changing to develop complete information systems. It is for this reason that Lanzara (1999) has an extensive outline on the logic of bricolage. According to him a large part of designing “consists in transforming and reshaping what is already in use” (Lanzara, 1999:346). All the role players in action should know the business.

Bansler and Havn (2004: 633) further explain organizational improvisation in information systems development by explaining the concept of bricolage, “i.e. the ability to use whatever resources and repertoire one has to perform whatever task one faces” (see also Lanzara, 1999; Louridas, 1999; Weick cited in Bansler & Havn,2004: 633).

2.3.3.3 Sensemaking

Sensemaking as part of the development process is, according to Bansler and Havn (2004: 641), “a process where people strive to convert a world of experience into an intelligible and

meaningful world”. The essence of sensemaking according to them “is the idea that understanding lies in the path of action. Action precedes understanding and focuses interpretation. Sensemaking by means of manipulation is for Bansler and Havn (2004) the “acting in ways that create something (e.g. a new technology) that people can then comprehend and manage” (Bansler & Havn, 2004: 641). The key point for Bansler and Havn (2004: 642) is “that sensemaking is an active process and that action is a precondition for sensemaking”.

The social relativist paradigm of Hirschheim and Klein (1989) in system development is described as a sensemaking process. Sensemaking emerges through a social interaction of the system developer with management and users. Central to this paradigm, systems development is an evolutionary process, because the social environment is under continuous evolution. According to Hirschheim and Klein (1989) sensemaking should be the focus in ISD rather than the preoccupation with the search for causal and objective explanations for social phenomena.

In conclusion it is obvious that the development process that Bansler and Havn (2004) describe differs remarkably from the orderly, structured paths that most IS theories and methods tend to assume. They however, have observed a

“more emergent, more spontaneous, more open-ended and more continuous process involving bricolage, unjustified trial and error, small-scale practical experiments, local readjustments, and improvisations. The process has been shaped more by action than by plans, and more by attention than by intention” (Bansler & Havn, 2004: 643).

They highlight the importance of users of the system as significant sources of innovation.

Bansler and Havn (2004: 641) emphasised that

“unlike most managers (and IT specialists), they have detailed, first hand knowledge of how work is actually done and they personally experience the frustrations, troubles, and breakdowns caused by inexpedient work procedures, rigid rules and inadequate or outdated technologies. As a consequence, they are often both motivated and able to come up with creative solutions to recurring problems and discover opportunities to improve their practices by exploiting new technologies.”

We know that a computer system can only process the quantifiable reality which leads to rational management. That is why these behaviours, values, experiences of developers and users of the information systems can never be included or anticipated in a formal information system neither can it be captured in a database. However, it is a crucial part of the information system that contributes to meaningful decision making. According to Bansler and Havn (2004) they adopt a sensemaking perspective to analyse the dynamics of the information systems development process and show that improvisational action and bricolage (making do with what you have and then building on it) play an essential role in the development of the application. They underscore three important aspects in the longitudinal case study of the development of a Web-based groupware system. This is important to this study, because a similar approach will be followed in the empirical model of the WCED information systems development process. Table 1 summarises the major differences between traditional and emergent approaches to ISD.

Table 1: A Juxtaposition of New Emerging Approaches and the Functional Approach:

Research Findings

	Emergent	Traditional
Leadership	Action/involved	Rational/predictable/controllable
Action	Doing and thinking	Thinking and doing
Method	Extemporaneous, improvisation, Bricolage, random and opportunistic	Planned, managed and controlled
Analysis	Sensemaking	Problem solving
Knowledge Paradigm	Doing the right thing Tacit	Doing things right Explicit
Systems	Throwness Holistic	Linear Entities
Process	Flexible/Uncertain	Technology focus
Epistemological Assumptions	Sensemaking Improvisation	Quantifiable Positivism/Predictable

From the above table, it becomes evident how traditional approaches often do not approximate real-life situations. ISD must always guard against mismatching the reality (the life-world) and the information system that is used to represent it. The present study aimed to develop an EMIS that would more accurately reflect the life-world as experienced by the key actors in a particular context during the ISD process.

2.4 Comparative summary of the subject literature

The literature review has revealed some shifts in approaches and steps, components and activities that informed in the development of information system development. Particular activities and components from each information systems development approach will be maintained to build the fundamental conceptual framework or model applicable to this study.

Through the different ISD approaches the research defines “a set of assumptions about knowledge and how to acquire it (epistemology) and about the social and technical world in the world (ontology)” (Henriksen, 2003: 21).

The objective of this study is not to adopt or streamline a particular approach, but to combine variants of the information system approaches and utilise them in a complementary way to produce an information systems development model. The traditional ISD methods require mechanistic, predictable and rationalistic approaches and therefore are too rigid to be used in the flexible environment of emerging technologies as a method on its own. Bansler and Havn (2004: 632) duly recognise that,

“ISD can be and often is performed as a deliberate, purposeful project with formal governance structures, requirement specifications, milestones and substantial technological and organisational resources...underlin[ing] the fact that ISD development happens in a myriad of other ways...These development processes are more emergent, more continuous and more filled with surprise, more difficult to control, more tied to their circumstances and more affected by what people pay attention to than by intentions, plans and methodologies”.

The same strategy will be followed here to firstly establish an information systems model as a prototype based on the information development life cycle and the principles of system theory. Then include concepts such as improvisation, reflective sense-making, bricolage and other emergent changes to further enhance and develop a comprehensive information system for data analysis purposes.

To provide structure to the information systems development model, general systems theory approach supports the background of this research. The establishment of this system shows that it is necessary to develop an information system as a planned, deliberate and systematic activity according to specific method. However, once the prototype is established improvisation, emergent change and unanticipated outcomes may play a pivotal role in ISD. To be able to do this certain criteria and structures should be adhered to. These are the important requirements that will be described in this section in order to be successful in the eventual ISD process.

This research attempted to shed light on ISD and shows that although system development needs specific guidelines and methods, improvisation, sensemaking and bricolage could play an important role to enhance the information systems development process.

The literature review confirms that according to the traditional approach information systems are developed according to a life cycle. Although, according to Laudon and Laudon (1995) this method is still the predominant method for building information systems in most organisations today, this life cycle metaphor will only be used in this study to determine the information system phases. In this way it helps to identify the roles and functions of the

business (education in this study) and technical specialist involved in the information systems development. Although the development process is not regarded as a rigid and linear sequence of steps particular phases have been distinguished. A similar approach will be followed in the empirical model with

- phase one as information systems analysis and design,
- phase two as application development,
- phase three as information systems data processing and
- phase four as data analysis and reporting.

The system theory principles that will be applied in the information system development model are inputs, transformation processes, outputs, linkages and interrelationships between subsystems, feedback and control. The systems theory will help to give structure to the model design. Weick (2001:43) in the language of systems theory represented organisations as illustrated by Figure 2.



Figure 2: Systems Theory

This diagram serves as a guideline for the structure and facilitates the systems thinking approach. Barrett (1998:615) is spot on about the importance of the new conceptual frameworks in the modern, emerging organisations when he states that “organisations tend to

forget how much improvisation, bricolage, and retrospective sense making are required to complete daily tasks”.

Truex et al. (2000: 59) describe methodical information systems development as a “managed, controlled process”, while assumption of the amethodical development is a “random, opportunistic process”. For them information systems development according to a method is a “linear, sequential process” and amethodical is “simultaneous and overlapping”. The following outline is an attempt to capture definitionally what is common among the different researchers in the field of emerging approaches.

Improvisation

Developers and users of an information system, are key role players in this regard. Depending on their level of expertise and experience they can improvise. In this regard Kalermo and Rissanen (2002) distinguished between skilled and novice improvisation.

“Skilled improvisation is performed by experienced people those are very familiar with the domain and used technology, and that have a broad and holistic picture of general development process and different phases and tasks related to it. They also have a good understanding of the business environment” (p. 20).

It must be remembered that in the development life cycle of a system different people perform different functions (actions). Everyone brings his own interpretation, behaviour and knowledge to the table. Many researchers (Bansler & Havn, 2004; Barrett, 1998; Ciborra, 1999a, 1999b; Kalermo & Rissanen, 2002; Miner, Bassoff & Moorman, 2001; Moorman & Miner, 1998a, 1998b; Weick, 1998) have explained the effect of improvisation on information

systems development and regarded it as “extemporaneous” (dealing with the unforeseen, unexpected, no blue print and predefined plans and methods).

Sensemaking

Sensemaking (Bansler & Havn 2004; Hirschheim & Klein, 1989; Weick, 1995, 2001) is about interpreting and sizing up reality and converting a world of experience into an intelligible and meaningful world. The sensemaking process is an alternative where epistemological assumptions are based on a socially constructed reality and human interaction.

In contrast the functional approach rests upon a positivistic notion that is based on a reality that is predictable and quantifiable.

Bricolage

Bricolage uses whatever technologies, resources, and existing information systems are available and build on it. Transforming and reshaping what is already in use. According to this paradigm one cannot improvise on nothing, only that which is existent can be improved. That is why improvisations are closely associated with the concept of bricolage (Bansler & Havn, 2004; Barrett, 1998; Henriksen 2003; Lanzara, 1999; Loudiras, 1999; Moorman & Miner, 1998b).

2.5 Components and activities of the fundamental conceptual framework or model

After examining the numerous ISD methodologies cited in the subject literature the proposed framework suggests the following stages in ISD life cycle.

The model, in Table 2 defines four distinct phases in information systems development when using the survey method based on the systems development life cycle (SDLC) approach. The phases in the model are the System Analysis and Design; Application Development; Information Systems Data Processing; and Data Analysis and Reporting. Furthermore, each phase is described according to the basic principles of the general systems theory approach of input, transformation processing, output and feedback and control. The aim is not to offer a comprehensive treatise on systems theory, but to harness the essential principles that facilitate an information systems model. It provides a framework for a strategic management information system to apply quantitative methods for discovery of knowledge embedded in the data. Each phase further highlights the important role of emerging concepts of improvisation, sensemaking and bricolage in information systems development.

Table 2: Conceptual Framework for ISD based on Literature Review

Design Phases	Design Structure: Systems Thinking	Design Approaches: Emergent Concepts	Actors involved
Systems Analysis and Design (Survey Instrument -Questionnaire)	Input Questionnaire Content Development	Sensemaking: -Inclusion of tables for education indicators - integration of data sets (PERSAL, Census, EMIS) - Decision-making Improvisation: - technical requirements Bricolage:	Education management & Users
	Transformation Processing (Technical Requirements)	Sensemaking: Bricolage pre-printing of questionnaires Improvisation:	Users and Systems developers
	Output - Designed Instruments - Completed survey form	Sensemaking: - cross reference tables for verification Improvisation	Management & school principals
	Feedback and Control	- Quality of the information provided - Outstanding survey forms	Education management

<p>Application Development (The database)</p>	<p>Input :The designed survey form (URS)</p>	<p>Bricolage: Use the existing designed survey form as user requirement specifications</p>	<p>System Developers</p>
	<p>- Transformation Processing: Data Modelling</p>	<p>Sensemaking: Map the data model from the survey form – use tables as entities and labels of table rows and columns as attributes.</p>	
	<p>Output: web-application</p>		
	<p>Feedback and Control</p>		
<p>Information Systems Data Processing</p> <ul style="list-style-type: none"> - Capturing - Verification - Validation 	<p>Input (Survey Form)</p>	<p>Bricolage: Form completed by schools</p>	<p>Data captures and users</p>
	<p>Transformation Processing(data capturing and cleaning)</p>	<p>Sensemaking: Verification and Validation</p>	
	<p>Output: Quality data (complete, relevant, accurate, timely, accessible)</p>		
	<p>Feedback and control</p>		

Data Analysis and Reporting			Policy makers Education management
	Input: data sets from PERSAL, examinations, EMIS and Census 1996 and 2001	Sensemaking as an ongoing process. Bricolage: integration of different data sets	
	Transformation: From data to knowledge through quantitative methods	Sensemaking through the application of quantitative methods	
	Output: Findings of quantitative analysis		
	Feedback		

2.6 Conclusion

The conceptual framework model is proposed as an alternative to the traditional and normative approach to ISD. The conceptual framework put forward here combines the phases of the SDLC, the principles of systems thinking (input, transformation processing, output and feedback), emergent design concepts (improvisational action, bricolage design and sensemaking process), and the key actors (users, systems developers and management) involved in the development process.

The conceptual framework that is developed focuses on the impact of technology changes on information systems development, using improvisation, bricolage and sensemaking as outcomes. With this exploration the research wants to indicate that although information

systems development takes place within a normative framework it allows for systems thinking and shows that beyond the rational and quantifiable world there is a reality that includes bricolage, improvisation and sensemaking. The development of such a model proposes that a new approach emerging in ISD elucidates that there is more to the rational approach in information systems than just the data that is in the database. Beyond the numbers and quantifiable world is a complex reality that amongst other things includes the atmosphere, culture and structure of the organization, together with the behaviour, emotions, knowledge and experience of all the people who in one way or another interact with the information system. A computer system can only process the measurable, tangible and quantifiable reality, which leads to rational management - a very one dimensional form of thinking which needs increasingly more numbers and figures. Benyon (1990:9) explains that “the computer has a tremendous capacity to manipulate data with extreme accuracy and speed...But it still only deals with a limited type of data – principally alphabetic and numeric symbols”. He stresses the fact that “outside the computer system a host of other symbols are used to communicate information”.

Ciborra (1998:16) states that we could not continue with the old methods when developing information systems. Current technologies (the Internet, World Wide Web and emergence of global IT infrastructures) require us to be less formal and structured and to devise our traditional, rational, structured, systematic methodologies. Lanzara (1999: 332), when comparing transient constructs and persistent structures, asserted that “nowadays we live in an increasingly artificial, computer-permeated world which is dynamic, ever changing, and in constant state of flux” that requires designing systems action.

This model will be used to determine the structure and the content of the next chapters. Chapter three, the empirical framework of the activities and processes of EMIS of WCED, will be based on the first three phases of the model. Chapters four and five, the application of quantitative methods, are based on the data analysis and reporting phase of the ISD model.

CHAPTER THREE

PRACTICAL APPLICATION OF THE ISD MODEL: AN EMPIRICAL STUDY OF EMIS

3.1 Introduction

This chapter uses the ISD model developed in Chapter 2 to develop EMIS according to the annual survey method. In this chapter, a practical example of an information system is used to illustrate the major aspects of EMIS and shows its role in everyday practice. The first three phases of the conceptual model for ISD based on the subject literature is used to describe the actual EMIS in the Western Cape Education Department based on the identified essential elements of the conceptual framework or model. Each of these three phases of the information systems development (System Analysis and Design; Application Development; and Information Systems Data Processing) is described according to procedures, activities and the key role players when using the survey method.

The information systems development (ISD) model developed in the previous chapter (Table 2) will be used to organise this chapter. The practical study was an example of a study on how information systems can be realised in a real-life (particular) situation. Other studies of ISD (e.g. Bansler & Bodker, 1993; Bansler & Havn, 2004; Büscher, Gill, Mogensen & Shapiro, 2001; Carstensen & Vogelsang, 2001; Henriksen, 2003; Kalermo & Rissanen, 2002) have also used the constitutive elements of information system structures and implemented it in formal projects of ISD to evaluate how they are realized in particular situations.

The notion of ‘particularity’ is addressed in a few other studies. Henriksen (2003) approaches ISD as an ‘empirical matter’ and refers to it as ‘radical localism’ and Leedom (2001) describes it as ‘situation awareness’. For Henriksen (2003:3) “the word ‘radical’ implies a drastic deviation from the established normative approach of ISD through methods, models, and guidelines. He contrasted ‘localism’ with the universalizing tendencies of ISD research by always attending to the local practices and specificities through which a particular information system comes into being”. Leedom (2001: 8) similarly defines ‘situation awareness’ as “dynamic situated knowledge, or the capacity to act effectively in a given specific situation” (e.g. knowing what’s going on so you can figure out what to do). Likewise Lind (2005) refers to the importance of a “contextual understanding of information systems”.

The focus on ‘particular’ or ‘local’ systems design in this study was to unlock the constitutive elements of the universal structures of the information systems that could be applied in a particular systems development environment in this case an educational context. The objective of this chapter was to show how the constitutive elements of ISD when practiced in the WCED as a particular context, used bricolage, improvisation and sensemaking.

3.2 Practice-in-Action approach method

ISD in a particular setup or real-life situation was in this study referred to as Practice-in-Action. Practice-in-Action in the context of this study includes the key role players as active participants in the development of EMIS. This approach is used to describe how the day-to-day actions and practical experiences of role players contributed to the design, development, implementation, testing, maintenance and improvement of the education information system.

It should be noted here that the information systems development in the WCED was not always a rational and normative process. Most of the development was based on the business needs of the department and then implemented without following a controlled and planned method. However, most of the activities implemented intuitively were in alignment with what the literature reports. The applicability of the approach is not argued for to make an ex-post *a virtue out of necessity*. It was grounded on the assumption that policy and policy making is never a perfectly rational process, particularly in fast-transforming environments, such as South Africa in the late 1990's. Acting as a bricoleur enables developers and users of information systems to build on existing practices and enhance the meaning and usefulness of existing data and systems. It does not thereby lessen the need for quality data, as reliable and valid as possible. Sensemaking enables the use of additional and alternative perspectives of the various roleplayers' theoretical and practical experience to inform the data and process. This concurs with Barry's (2005: 3) view that although the "design activity may appear chaotic and perhaps slightly out of control ... the process is directed by 'hidden rationality' of skilled individuals" (p.3). Anarchy is avoided through experienced decision-making and what Ciborra (1999b) refers to as "smart improvisation", and Kalermo and Rissanen (2002) as "skilled improvisation".

The collaboration of the key actors in the everyday practice of their duties allowed for improvisation and sensemaking. The primary modes of data collection and collaboration were interviews, formal and informal discussions, meetings, workshops, seminars, conferences, e-mails, telephone conversations and surveys.

The Practice-in-Action approach focused on activities in their natural settings and included the rich array of experiences and behaviours of the different organization actors during the development process. The Practice-in-Action method was adopted for management information systems development in WCED. Practice-in-Action is addressed in a few other studies when they refer to the notion of Workpractice (Casey & Brugha, 2005; Clarke, 2005; Goldkuhl, 2005; Jacucci, 2005; Lind 2005; Nurminen, 2005).

The Practice-in-Action approach relates to and has common characteristics with action research although not entirely similar in approach. Action research is a particular qualitative research method that emphasizes collaboration between researchers and practitioners, such as was done in this study and is particularly useful for the information systems field (Avison, Baskerville & Myers, 2001; Avison, Lau, Myers & Nielsen, 1999; Baskerville, 1999; Chiasson & Dexter, 2001; Mckay & Marshall, 2001; Mumford, 2001; Rose, 2000). Brody (2006:1) asserts that action research “is collaborative, critical, and self-critical inquiry by practitioners into a major problem, issue or concern in their own practice” as was the approach in this empirical study. Avison et al. (1999:94) suggest that “action research combines theory and practice (and researchers and practitioners) through change and reflection in an immediate problematic situation within a mutually acceptable ethical framework. Action research is an iterative process involving researchers and practitioners acting together on a particular cycle of activities, including problem diagnosis, action intervention, and reflective learning”.

3.3 Key actors in developing EMIS

In this study, the Practice-in-Action approach referred to a methodology used to design, develop and test information systems through a continued collaboration between key actors, namely the developers (of which the author was the manager) and the users as part of their daily actions (practice). It involved the development of the system, developing a planned change in the system, and conducting a continuous strategy of testing, refining, and improvement of the current system. This method was used to implement the information systems model in a real-life (particular) setting.

The role of these actors (users, developers and managers) was decisive during the development process, because they brought with them culture, emotions, behaviour and beliefs that are embedded in the system structure. Their interpretation and articulation influenced the finished product. Throughout this research, it was assumed that the action of the individual participants in the information system development is crucial for sensemaking, improvisation and bricolage. These assumptions are in line with the principles of sensemaking as described by Weick (1995; 2001).

The nature of this empirical approach required a complete exposition of the functions of each of the actors in the development of EMIS. The actors involved in the information systems development process of the Western Cape Education Department were as follow:

- The **developers** played a major technical role and were responsible for the development of the applications. Their contribution in the content development and technical requirements of which questionnaires were invaluable. The author played a leading role as manager in

the development of the information system. It was from a development point of view (technical) that the information systems process was described. As manager, the author was part of the developers that led the department through a series of upgrades to capitalise on emerging technologies and application enhancements, including the replacement of obsolete systems. The author's knowledge in the design and delivery of cost-effective high-performance information systems and web applications helped to address complex business (education) problems. Through this knowledge various web-based applications, such as the Annual School Survey, Adult Basic Education and Training, Learners with Special Educational Needs, and many more were initiated.

- The **users** were those people for whom the system was built and who would use the education management information system for the day-to-day practice. They performed multiple functions in the system.

The **users** consisted of the following groups:

- The **EMIS staff** whose primary function was to use the applications to provide information to management. They also provided valuable input in the design and development of the questionnaire. They played a leading role in the validation and verification of the data to ensure quality information.
- The **managers** (MEC, HOD, senior managers) in education set the guidelines and determined the content and data requirements for educational decision-making, policymaking, monitoring and evaluation.

Collaboration between the users and the developers enhanced the quality and the acceptance of the system.

Ives and Ohlsen (1984: 587) listed reasons why user participation improves quality of the system, which was also the case in the development of EMIS:

1. It provides a more accurate and complete assessment of user information requirements;
 2. It offers expertise about the organization the system is to support, expertise usually unavailable within the information systems group (the developers in this case);
 3. It avoids development of unacceptable or unimportant features;
 4. It improves user understanding of the system.
- The **individuals (teachers) at school level** who completed the questionnaire that was sent to all schools annually had to be well informed, as this was the point from where the data originated. Such local level influence can determine the quality of the information. These individuals therefore needed to make sense in terms of the reasons why questions are included and what purpose it has in the education decision-making and education development. Training sessions (road shows as it is called in the WCED) on the **how** to complete the questionnaires were conducted and was critical in the process.
 - The **data capturers** were responsible for the entry of the data in the system. Their competency levels determined the quality of the data and expedited the availability of the data. The scanning of the multitude of questionnaires was a time consuming exercise and a highly technical function that required intensive training and technical knowledge. The EMIS staff played a vital role in this process.

At this point, it is necessary to elaborate on the processes and procedures that informed the development of EMIS. The WCED annually sends a survey questionnaire to all schools for completion. The content of the survey questionnaire is compiled in collaboration with the National Department of Education. The content of the questionnaire consists of categories such as general information of the school (name, address, sector, grades), learner information

(enrolment, medium of instruction, home language, age and subjects offered), staff information and physical resources of the school. Schools in different sectors of education, pre-primary, primary, secondary, special education and adult education in public and private institutions must complete the form and send it back before a specific date. These data elements are considered imperative in the provision and management of education in the country.

The Education Department within the Provincial Government of the Western Cape is responsible for over 2000 public and private schools. To manage and provide education effectively and efficiently the WCED needs complete, accurate and timely information that is accessible to the relevant role players to make strategic decisions.

This study used the ISD model (Table 2) as the lens through which to investigate and describe EMIS when using the survey method for data collection, data cleaning and data analysis. The life cycle according to which this study described the education management information system of the Western Cape Education Department consists of four phases, namely, systems analysis and design (survey questionnaire), application development (database), information systems data processing and data analysis and reporting. In the following sections these phases are discussed. In each phase, basic systems theory principles of input, transformation processing, output, feedback and control were applied, together with the key activities of sensemaking, improvisation and bricolage. The actions of the people who were key role players in the entire process in the development process were continuously emphasised.

3.4 PHASE ONE: Information systems analysis and design (survey questionnaire)

3.4.1 Introduction

This section will examine the processes and procedures in the design and development of the survey questionnaire. It is during this phase that essential data elements that need to be included in the information system and the data gathering methods are determined.

There are various ways to gather and capture data in education information systems in the world. Emerging technologies such as the internet has brought new and faster approaches of data gathering. The internet (web) technology is more flexible and allows for on-line⁴ data entry that makes quality data available at a faster rate. Web technology however, brings along other challenges such as system security and confidentiality issues. Optical character recognition is another emerging technology that could speedup the capturing of data. Data collection through surveys is an old and well tested approach. The survey method is also the approach used in this study to develop information systems and for the collection of data from schools in the WCED.

This section provides an overview of the processes that were used in designing and developing of the annual survey instrument which each school has to complete. The design and development of the survey instrument consisted of the following steps:

- the content development process
- technical requirements

⁴“On-line” means the system is web-based and data is entered directly over the internet. Web technology makes it possible that data is entered directly into the system at the point of origin without intervention of third parties.

- printing and the distribution of the survey forms to the schools

This phase will be described according to the characteristics of the general systems theory, such as input, transformation processing, output, and feedback and control.

3.4.2 Input as an information systems constituent

The design of this survey instrument is important because it is used as the User Requirement Specifications (URS)⁵ to develop the database (input of next phase). In this section a detailed description of the construction of the questionnaire from a developer's perspective follows. At the heart of the education management information system lies the survey questionnaire. Through the questionnaire the input data from the schools is obtained.

In most countries the data required for education management information systems is obtained through surveys. The data collected from schools, the primary data source, via a survey is usually done with a questionnaire that serves as an instrument to develop the information system. The collection of data through survey questionnaires was the primary method in the WCED. The survey method however has certain disadvantages. It takes too long to distribute the survey form. Capturing of data and the cleaning of the data are time-consuming processes. The result is that quality and reliable data take too long to be available for education planning and decision making purposes. However, since 2006 the WCED has implemented an electronic record keeping system to keep track of individual learners through the education system.

⁵ A URS is a document that the developers of the database use as a guide to determine which functions need to be included in the system and to create the data model.

The design and development of the survey instrument demonstrated what an important source of innovation the users are. They use the system on a regular basis in performing their day-to-day functions and have detailed first hand knowledge of how the work is actually done. They are often the first to realise the limitations in the system and as a result they often generate original ideas to solve frequent problems and discover opportunities to enhance their work practice.

In typical systems theory fashion the content development process is analogous to the input of this phase. Three key aspects of the education management information system, namely sensemaking, improvisation and bricolage in the Western Cape are noticeable when we compare it to the way ISD is usually described.

3.4.2.1 A sensemaking perspective of the survey design

Weick (1995) describes enactment as one of the attributes of sensemaking. The content development process of the survey instrument includes the characteristic of enactment because according to Weick (1995:30) and also in the development process of the WCED, “sensemaking better explains how entities get there in the first place”. He continues to underscore this notion when he states that “people often produce part of the environment they face” (Weick, 1995:30). The **content development** of the survey instrument of the WCED was fundamentally based on this sensemaking principle. The elements included in the survey form were not random. The sense for including the data elements included in the survey instrument was based on the following considerations:

- 1) To make integration with other data sources possible specific data elements were included as unique identifiers (data elements with the same meaning and that are common with other data sets). These identifiers and common fields that link with other data sets, such as Census, PERSAL (a national database used as a personnel and salary system for all teachers in South Africa), Examinations and other provincial data sets. Common information standards across all information systems, such as a data dictionary and metadata, would make data integration possible. These data standards (code lists and data elements) are considered as probably the most effective way to ensure information sharing amongst provinces, departments, sectors and organisations.
- 2) Certain data elements were included in the questionnaire that could be utilised for education planning purposes. The data collected on the survey form facilitates the equitable distribution of resources. The school funding norms and standards (South African Schools Act of 1996), where schools are weighted according to certain criteria, require that resources and budget allocations are distributed in such a manner that institutions with the highest needs benefit the most. Information systems that can provide quality data help in this allocation. The provision of staff, building of schools, allocation of textbooks and resource allocation are management functions that require planning and monitoring activities in order to improve education development. Quality (reliable, accurate and complete) information systems are necessary for these decision making processes in education. The individuals at school level should therefore have an understanding of the purpose of the data elements that are included in the questionnaire in order to make sense of them.
- 3) Cleaning of the data contributes to the quality of the data. Certain data elements have been included in the survey instrument to cross reference totals in order to make sure that

information completed on the form is correct. Without this prior knowledge these questions on the questionnaire make no sense at all.

3.4.2.2 Survey design as improvisation

The development and the design of the survey questionnaire were not only based on conventional information system analysis principles (control and repeatability), but sometimes required improvisational action. When the need arose to add something in terms of content development or technical requirements, it was not necessary to go through all the steps of systems analysis, but rather to improvise. The development and design of the survey instrument took place according to the “Practice-in-Action” approach. The direct involvement and constant availability of the users and the developers provided opportunities to build functions and requirements into the system rapidly (unanticipated). We live in a stochastic world, and so is the education environment, that sometimes requires improvisational and extemporaneous action. According to the Concise Oxford Dictionary the word “stochastic” means “determined by a random distribution of probabilities” or “governed by the laws of probabilities”. In essence it means that the information systems process (education reality) is not so predictable and stable that everything can be planned before hand and nothing would ever change. Weick (2001: 152) pertinently backs up this notion when he quotes Davis and Taylor who suggest

“that previous industrial-era technologies were deterministic, with clear cause-effect relationships among what was to be done, how it was to be done, and when it was to be done. Newer technologies no longer are dominated by determinism. Instead, people operate in an environment whose ‘important events’ are randomly occurring and unpredictable”.

Weick (2001:153) is even more specific when he stated that the

“skill requirement of a stochastic environment are unique. A large repertoire of skills must be maintained, even though they are used infrequently; people are usually on standby, giving special attention to startup and to anticipating faults that may lead to downtime; the distinction between operations and maintenance is blurred; skills in monitoring and diagnostics are crucial; people must be committed to do what is necessary on their own initiative and have the autonomy to do so; and people have now assumed the role of variance absorber, dealing with and counteracting the unexpected”.

This is probably why the information systems development process in the WCED is reasonably successful (Practice-in-Action approach). The development of the information system was based on day-to-day activities of the key role players. It means that in the ISD process there needs to be individuals with the necessary technical and business skills on-site that are available to intervene when the unexpected occurs. Through the collaboration between users and developers the inclusion of relevant data elements were made possible in the survey questionnaire that improved and refined the process. Technical considerations, education content requirements, verification of the data, questionnaire printing considerations and management decisions sometimes necessitate improvisational action – “a doing and thinking” instead of “thinking and doing” approach – method in action.

3.4.2.3 Survey construction as bricolage design

Bricolage, as was explained already in the earlier sections, is the process that builds on existing material. The content development of the questionnaire in the WCED is also grounded on this principle. Constant collaboration between the key role players, such as users

and developers led to the realisation of the data needs and technical limitations within the system. The result was the continuous improvement and changing of the system.

3.4.2.4 Key actors in the questionnaire content development process

The education management and users of the system played a decisive role in the content development of the survey instrument. Although there was a close collaboration between users, developers and management throughout the development process, the users and management were those with the first hand knowledge of the data requirements and information needed in an education management system. However, there was a constant interaction between the developers and the users to determine the questionnaire layout and the technical feasibility when specific questions were included. The survey instrument would eventually become the user requirement specification (URS). The URS is a document that provides the specifications for the development of the database. The management and users therefore fulfilled the role of the systems analysts as they were responsible to make sense of the data.

3.4.3 Transformation processing as a necessary characteristic of ISD

The basic requirements of this phase (survey questionnaire analysis and design) are inputs, outputs and processes. Transformation processes take the data elements and produce a survey questionnaire that could be used as specifications that are suitable for database design and as a data gathering instrument. Since the output (survey questionnaire) of one phase may be the input to another phase (URS for database design).

The transformation processes included in the questionnaire development were the design of the survey form, printing and distribution of the survey forms. These transformation processes were largely based on technical requirements that were further based on sensemaking and improvisational actions.

3.4.3.1 The transformation as a sensemaking process

Systems transform inputs into outputs (survey questionnaire) through technical processes. The capturing method used to a large extent determined the design basics and the presentation and appearance of the survey instrument in order to adhere to the technical requirements. These technical requirements included considerations for database development and optical character recognition as a capturing method. The users (EMIS staff) who designed the questionnaire needed to understand these technical requirements and adapted the layout accordingly. The following technical requirements in designing of the survey questionnaire were sensemaking processes:

The presentation of the questions on the survey form is in the format of a table (rows and columns). The entire data model (from which the database is designed), the backbone of the web application, is derived from this structure. The mapping of the data model from the survey layout is primarily a sensemaking process. The attributes of the data entities for the data model were derived from the labels of the rows and columns of the tables on the questionnaire.

3.4.3.2 Bricolage design

The computer applications in the information system improved over time from conventional desk-top applications to interactive web-based applications. We could learn from experience and based on existing technology determined what would be the most suitable technology to be used.

Pre-printing was the process where information that was already available in the information system and was consistent over a long period of time was printed on the survey questionnaire. The individuals that completed the questionnaire could then make the required changes where applicable and add the additional information. This was a bricolage design approach that contributed to the integrity of the information system.

3.4.3.3 Improvisational action in the transformation processing

The design, development, printing and the distribution of the survey instrument were not always based on planned, normative and rational activities. The education environment is unpredictable and complex with lots of problems that arise during the course of the development process. Extemporaneous and spontaneous action is sometimes required.

To manage the large number of forms manually and the amount of information displayed on the form were time-consuming processes. The development of computer programmes was a result of an urgent need and required action without extensive prior planning (method in action approach).

To further underscore the improvisational approach (Practice-in-Action) it could be mentioned that certain data elements on the survey were included with data verification and validation in mind. Actions which happened based on a need as the information system was developed. The Education Department put a lot of effort into producing quality information. The size of the information system made it virtually impossible to manually verify every bit of information captured and to validate the accuracy of the information received from the schools. These data elements were included in the survey form to cross reference and to verify the data entered. Cross referenced symbols were included on the survey to give the schools indications which data tables needed to balance. (This was an unplanned exercise based on previous experience and immediate needs in order to speedup up the data cleaning process).

3.4.3.4 Key Actors in the transformation process

The presentation of the content on the survey instrument (layout) was largely based on technical requirements. The web applications and the OCR system were developed according to the questionnaire layout. The system developers with the technical knowledge, in collaboration with the users with business knowledge, were key in the successful development of the information system. A crucial factor was the on-site presence of the developers who could implement any technical requirements as the need arose. The users who designed the questionnaire layout needed to take these technical requirements into consideration, especially with regard to the data capturing method and the system development platform (in this case an interactive web-based application). Based on these foundations the technical considerations of the survey instrument were based on an improvisational action (method in action), a bricolage approach (tried out and improved on existing technologies) and a sense making process (understand what is required, interpreting the questionnaire layout, experience and knowledge

of the education environment). As Weick (2001:9) avers “it is the job of the sensemaker (in this case the systems developer) to convert a world of experience into an intelligible world”. He delineates sensemaking as a metaphor that “focuses attention upon the idea that reality of everyday life must be seen as an ongoing ‘accomplishment’, which takes particular shape as individuals attempt to create order and make retrospective sense of the situations in which they find themselves” (p.11).

This kind of development uses what was earlier described in this study as agile methodologies. Nerur, Mahapatra and Mangalara (1998) compare agile methodologies with the traditional way of developing information systems. They emphasise the important role of the developer as the sensemaker by stating that “agile methodologies require a shift from command-and-control management to leadership-and-collaboration” (p. 76). They further state that “much of the knowledge in agile development is tacit and resides in the heads of the development team members, an assumption which makes the organisation heavily dependent on the development team members” (Nerur et al., 1998:76).

Hirschheim and Klein (1989) describe systems development as sensemaking and maintain that it emerges through “social interaction”. Henriksen (2003) refers to this continuous interaction between key actors as “articulation work”. According to him “Articulation work is a term for talking about some of the ongoing accommodation, adjustments, moment-to-moment activities people engage in to make things work and move on with the task at hand” (Henriksen, 2003: 43). Barry (2005: 3) asserts that “software development is complex, unpredictable and there is no clear procedure for resolving all the problems that arise during systems development”. He

further states that although the “design activity may appear chaotic and perhaps slightly out of control ... the process is directed by ‘hidden rationality’ of skilled individuals” (p.3).

3.4.4 Output of the information systems design phase

Weick (1995:30) uses the concept of enactment to indicate how people “often produce part of the environment they face” a notion that is probably more appropriate in the output of the phase than at in any other stage. The key role players created an environment (content development) with which they interacted throughout the information systems process. The output of the design phase was the manifestation of this enacted world. According Weick (1995:43) this “ongoing” process is one of the characteristics of sensemaking. This section will demonstrate that the information systems development in WCED was an ongoing process. This ongoing notion is supported by the fact that the output of this phase (the designed questionnaire) would be used as input to two other phases. Although the information systems development process is not linear and sequential, certain tasks should be completed before others can start. The start of the one activity is dependent on the finishing of others. For example, in this study the designed questionnaire provided the technical requirements for the development of the data model and the completed survey questionnaire became the input for the data processing (capturing, verification and validation) phase.

3.4.4.1 Output from a sensemaking perspective

The sensemaking aspects of the survey questionnaire included the following:

- 1) The schools which have to complete the survey form needed to make sense of the data elements included on the form. They needed to understand that some of these data

elements were used to cross reference totals for accuracy purposes. Enrolment information was used to verify the entries of other information such as learners according to age groups, learners according to home language, learners according to medium of instruction, learner totals according to subjects offered, and more.

- 2) The labels of the rows and columns of the table grids on the survey instrument were used to create the data model for database development. The users (EMIS staff) of the information system in the WCED, who are responsible for designing the survey questionnaire needed to be aware of this process. There should be close collaboration between the designers of the questionnaire and the developers of the database. Weick (1995: 30) refers to this interaction between different role players as the “social” dimension of sensemaking and describes it as “a network of intersubjectively shared meanings that are sustained through the development and use of a common language and everyday social interaction”. Hirschheim and Klein (1989:1205) elaborate on this social interaction when they state that “users are the organisational agents who interpret and make sense of their surroundings. The system developer is the change agent who helps users make sense of the new system and its environment.”
- 3) The grid layout of the form (boxes on the form) was included on the form in which the data was entered, but was also a technical requirement for the OCR process. All the participants (users, system developers and individuals who completed the survey form) had to be knowledgeable about these information systems development requirements.

3.4.4.2 Key Actors

The successful completion of the survey instrument and the quality of the information depended on how the schools understood the questions and the reason why they had to

complete the survey forms. Decision-making in the Education Department is largely dependent on the information received via this survey census each year. The provision of teachers, the school funding norms and standards, building of schools, provision of text books are all enrolment driven - information that is obtained through the survey census. Managers within the Education Department therefore play a pivotal role to enlighten school principals about the purpose of the annual school survey. This was done through road shows (gatherings with principals at different places), circulars, e-mails, news bulletins and more.

3.4.5 Questionnaire feedback and control

It was essential to give feedback to schools by informing them about the school survey instrument in terms of:

- the quality of the information provided
- still outstanding survey forms

Control and feedback helped to improve the general data quality and decision-making process. The education managers (school principals, circuit⁶ and district managers, and education management at the head office of the department) were key role players in this regard.

3.4.6 Concluding remarks on survey design phase

In summary, this first phase of information systems development provided key foundations for other phases that followed. The systems thinking approach followed in this phase focused on

⁶ A circuit comprises a pre-defined number of schools within a specific area. A school district is a number of circuits grouped together within a demarcated geographical area.

the relevant activities. The constructs of sensemaking, bricolage and improvisation provided a meaningful conceptualisation of the information system development process in WCED. Throughout this phase it was apparent that apart from the normative, planned and controlled activities alternative approaches, such as “method in action” and extemporaneous actions could help to improve and enhance the system. The output of this phase (the survey questionnaire) was used as input to the next phase. The next paragraphs will outline the work practices in developing the computer application.

3.5 PHASE TWO: Application development

3.5.1 Introduction: Interactive web-based information system

The second phase (the practical application development) of the conceptual model developed in Chapter 2 was used to organize and structure this section. This study considered the problem of ISD and followed a practical approach (Practice-in-Action). The experiences and actions of the systems developers and the users in their engagement with the information system during their day-to-day work practice generated much knowledge. This knowledge was then utilized to improve and enhance the existing information systems in an attempt to adopt a more agile approach.

This section outlines the steps that were used to develop a computer application as an integral part of the management information system. According to Benyon (1990:49) “a database [and the computer application] is central to the information system and its design crucial to the efficacy of that system” and supports the actions of the participants in performing their

functions. Goldkuhl and Röstlinger cited by Lind (2005:4) regards “a computerized system as an action system”. For Lind (2005:4) an information system “is both an instrument for performance of action and a support tool for humans to perform their actions”. This notion of action corresponds with the Practice-in-Action approach followed in this study. Lind (2005:4) concurs that “information systems should be actable”. He agrees with Goldkuhl and Röstlinger that defined the information systems actability “as an information system’s ability to perform actions, and to permit, promote and facilitate the performance of actions by users both through the system and based on information from the system, in some business context” (p. 5). This applied in the education environment of the Western Cape. Andersen et al. (cited Casey and Brugha (2005)) described a similar approach, namely Professional Work Practice (PWP). The PWP according to Casey and Brugha (2005:58) “emerged from empirical analysis of what systems practitioners actually do in practice”.

Providing information for the management of the education system for planning, monitoring and decision-making purposes was a central idea in this study. To provide complete, relevant, accurate and timely information as far as possible that is accessible to the entire organization has become a tremendous challenge. Emergent organizations, such as WCED currently content with increasing change as a result of changing environments and technology, and this contributes to the complexity of the education environment. Consequently, information systems have become complex and much more demand is put on the developers to shorten the development cycle and make quality information available in a shorter time. The research objective was to demonstrate that by using local design improvisations⁷ and sensemaking

⁷ Actions by key actors (users and developers) who are not so far removed from the information system could implement changes and extra requirements almost instantaneously as the information system developed.

processes instead of the formal, traditional ISD methods, the application development time and the timeframes of the development cycle could be fast tracked considerably.

The development of a computer application is a highly technical exercise and most of the detail falls outside the scope of this study. One of the great benefits of web technology is its ability to present information to the user in a simple way through a web browser. Furthermore, a web-based application makes information accessible from anywhere within the organization and is accessible to larger audiences. Web-based applications in comparison with desktop or mainframe applications require new approaches in design and development. Application development is an important phase in the overall information system development process. Applications in the EMIS of WCED, such as school surveys (Public schools, LSEN and ABET) integrated with other applications (CENSUS, PERSAL, examination, etc.) make more comprehensive datasets available for knowledge discovery through analysis.

3.5.2 Input of the application development phase

The input of this phase was the survey questionnaire as discussed in 3.4. Therefore, the development of the computer application relies mainly on the sensemaking dimension which Weick (1995) labels as “ongoing”, because the output of the previous phase (survey questionnaire) becomes the input to this phase.

A specific layout structure of the survey questionnaire was an essential requirement for the application development. The survey questionnaire was used as the user specifications to create the data model and to develop the database (output of the phase). The questionnaire was

used as the input to this phase. The survey questionnaire as described in phase one serves as the user requirements specifications (URS). The URS is used to create the data model from which the database is developed.

3.5.3 Transformation processing: Steps in building a web-application

The database design process has three specific steps, namely the conceptual design, the logical design and the physical design. During the conceptual design phase, a model of the empirical life-world is created or as Benyon (1990:55) stated “it deals with the inherent structure of the enterprise” - in this case, the education environment within the Western Cape. The logical design phase is the mapping of the conceptual phase into logical components. In the education environment, it would be entities such as school, learner, staff, infrastructure and subjects. The entities then become tables in the database and the attributes of the entities the table columns. The physical design phase is the actual development of the database using a specific database environment. The WCED used a database technology that made web-based applications possible.

The first step in the database design and development phase is the data model. The **data** model⁸ is part of the conceptual design phase which will be reviewed since all the other steps (logical and physical) are reliant on the effectiveness of the data model. The technical detail of the data model is beyond the scope of this study.

⁸ In the overview of the data model the [Information Technology Services](http://www.utexas.edu/its/windows/database/datamodeling/dm/overview.html) at [The University of Texas at Austin](http://www.utexas.edu) describes a data model as a “conceptual representation of the data structures that are required by a database. The data structures include the data objects, the associations between data objects, and the rules which govern operations on the objects. As the name implies, the data model focuses on what data is required and how it should be organized rather than what operations will be performed on the data”. Available on line website: <http://www.utexas.edu/its/windows/database/datamodeling/dm/overview.html>

This section outlines the effective use of data models in building web applications and describes sensemaking in information system development. A data model outlines the information requirements (the survey questionnaire) needed to develop the database. The data model represents the structure of the education department by modeling it in terms of entities and the relationships between those entities (Benyon, 1990: 54). The data model included the entities of the Western Cape Education Department and attributes (the characteristics of the entities) and from this data model, the database was developed.

The aim of this section is to show how data models were developed in practice and presented a method to create data models from a survey questionnaire. The aim of data models was to produce a specification (Entity-Relation diagram) in enough detail to develop the database application and it was attempted to achieve this without too lengthy technical detail. This is a very important part of the application development and captures the sensemaking perspective. The Entity-Relation-model (E-R model) is probably the most used method to build data models for database development. The Entity-Relationship Model (Benyon, 1990) is a graphical representation of the real world (education in this study) as depicted by different entities and relationships between them. The *attributes* (Benyon, 1990) are characteristics of the entity. Each of these entities of significance has attributes and behaviours, for example a person has a name, surname, date of birth and address. An attribute is a data element, and not an abstract concept, that can be managed. Data therefore can be defined as the attributes of the entities (real-world phenomena) of interest to an organization. For the purpose of this study these attributes will be used as the data elements of an organization. A *relationship* (Benyon, 1990) in database terms is an association and dependency between two or more entities. For example, there is a relationship between the entity learner and the entity school.

The cardinality of such a relationship could be one-to-one, one-to-many and many-to-many. *Normalisation* is basically the process of efficiently and effectively organizing the data in a database by simplifying the complicated relationships so that no redundancies occur. The secret is to split large tables into smaller and more efficient ones. The following section describes the application of these data modeling concepts.

The data model from a Sensemaking perspective

The central idea of this section is to describe the development of a data model from a sensemaking perspective. The technical aspects of the data model and how it relates to database development is beyond the scope of this study. However, it is important to illustrate how a data model was mapped from the data elements of the questionnaire to emphasize the importance of the sensemaking process. The collaboration between the users (who designed the questionnaire) and the developers (who technically designed the database) had to make sense of the data elements included in the questionnaire. They needed to understand that the layout (grid format) on the survey questionnaire contained all the entities and attributes needed to create the data model. The users therefore had to interact with the developers to produce a form that was in line with the development requirements. Bansler and Havn (2004: 641) affirm this notion when they state that “at the heart of sensemaking is the idea that understanding lies in the path of action. Action precedes understanding and focuses interpretation”. In all probability, the most valuable contribution of such a sensemaking approach is that it could reduce a usually highly technically process to a fairly simple and straightforward exercise.

Table 3 is an example of the data elements that were included in the questionnaire and indicated how most of the form layout tables were presented for data entry purposes. All of these tables on the survey questionnaire would be an entity in the E-R model.

Table 3: Example of the Data Entry Table: Survey Instrument

Total number of LEARNERS enrolled at the school in 2002 according to grade and gender																	
Gender	GRADES																
	Pre-Gr R	Gr R	1	2	3	4	5	6	7	8	9	10	11	12	Post Matric	LSEN	Total
M		2					7		9								0
F							16							28			0
Total																	#

Table 3, for example represented the entity ‘learner enrolment’ in the E-R model. The data model for this entity was mapped from the questionnaire according to the following sensemaking process. The attributes (school_id, gender, grade and total) of the entity learner enrolment in the data model (Figure 3) have been derived from Table 3 in the following way:

1. The attribute ‘*school_id*’ is a reference number of a school that is used as a unique identifier to link this entity with other entities.
2. The next step is to map the row labels to the data model. In this example, the row label ‘gender’ becomes the ‘*gender_id*’ attribute in the data model.
3. The next step is to map the column label for grades as the attribute ‘*grade_id*’ in the data model.

4. The grid entries (the boxes of the survey questionnaire) become the ‘answer’ attribute in the data model.

This sensemaking mapping exercise resulted in a data model as presented in Figure 3.

A few other technical steps that are not recorded here resulted in the data model as presented in Figure 3.

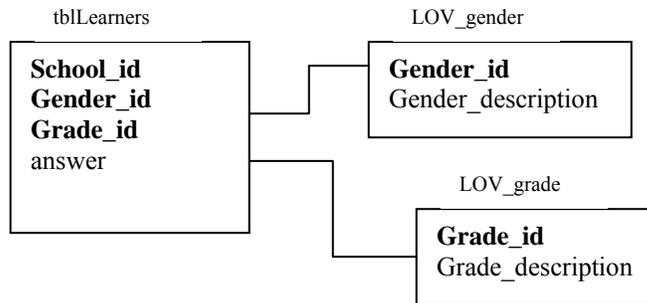


Figure 3: Entity-Relationship Diagram

The above figure is based on survey instrument in Table 3. From this diagram the eventual database is developed.

3.5.4 Output

The practical application development phase produced a number of outputs. A particular output in this phase, for example the completed data model, which was used as the input for the development of the database. The eventual output of this phase was the computerized application that was used as the data-capturing tool (input) in the next phase.

3.5.5 Key actors

A close collaboration existed between the developers and the users as was indicated throughout the development of the application. The development of the application emphasized the diversity of skills utilized in the process.

3.5.6 Concluding remarks on the application development phase

The web application facilitated much easier and more efficient data analyses. The development of a web-based computerized information system, supporting information processes and their use and development were analysed in terms of sensemaking, improvisation and bricolage. The sensemaking perspective as an ongoing process is illustrated when the output (developed application) of this phase is used as the input (capturing of the data) in the next phase. Ciborra (1998:16) states that one cannot continue with outdated methods when developing information systems – innovation is a necessary facet of ISD.

Current technologies (the Internet, World Wide Web and emergence of global IT infrastructures) demand that one being less formal and structured and to revisit our traditional, rational, structured, systematic methodologies. Information systems development needs to take these web-technology changes and trends into consideration. The literature on web-design and development of web-based information systems seem to propagate that the development of web-based systems is different from development of traditional IT-based systems (Barry, n.d.; Barry & Brown, 2003; Bødker & Carstensen, 2004; Carstensen & Vogelsang, 2001; Ciborra, 1998; Koufaris, Isakowitz & Bieber, 1999). Carstensen and Vogelsang (2001: 537) argue that

“web-based information systems are, however, different from traditional information systems in the sense that they require new approaches and often are the results of grass-root efforts”. They further state that “the development of information systems of all kinds has been heavily influenced by the introduction of the web-technology as a platform for many different types of systems. This implies major changes in the approaches, organisation principles, methodologies, tools, etc” (p. 546). Furthermore the interaction and collaboration between users and developers is in essence changing and the need for new skills is increasing. The Practice-in-Action could be a flexible method in order to adhere to the emerging development requirements and the dynamic environments.

3.6 PHASE THREE: INFORMATION SYSTEM DATA PROCESSING

3.6.1 Introduction

The information data processing phase is the third phase according to the information system model. In describing the data processing phase it was important to identify the operations and functions of the role players and determining the inputs to the information system, and the processes that transform these inputs into outputs. Important activities during this information system phase were the data capturing and data cleaning processes. The data cleaning processes, namely data verification and data validation transformed the data into quality data.

3.6.2 Input : Data entry

The computer application developed in the previous section was used as the input for this phase. Through the data capturing process data was entered into the system.

The nature of the data capturing process was determined by the accuracy that was required and the urgency of the data. The following data capturing solutions were in place in the Western Cape Education Department:

- Optical Character Recognition (OCR). This highly technical process is much faster way of capturing data and enabled the WCED to have data available in a relative shorter period in comparison when other conventional methods were being used. The preparation of the survey questionnaire was the key to success for the OCR process. Furthermore to capture these high volumes of data through this process the accuracy rate substantially increased.

All the surveys questionnaires of the WCED that were completed by the institutions were captured according to this method.

- On-line data capturing was another method used by WCED to capture data. Given that most of the transactional systems of the department were web-based applications, it enabled them to enter data on-line and this shortened the data capturing process considerably.

3.6.3 Transformation processing

Data verification and validation processes were important to change the raw data to quality data. It was important to prevent errors from occurring in the data and this verification and validation processes of the data contributed to the quality of the data.

3.6.3.1 Sensemaking

The verification and validation processes were based on the sensemaking perspective. The users who performed these processes needed to have knowledge of the education system and had to understand the purpose of captured data.

Verification ensured the prevention of errors occurring when data was captured. Verification entailed checking that data had been entered correctly from the survey form into the system correctly. The verification methods used were

- Checking that the data entered from the survey questionnaire corresponded with the data in the database (proof reading).
- The using of computer programmes (queries) to compare the total of data tables that should balance.

Validation was used to check that the data entered had meaning, for example, that data was within a range of values.

3.6.4 Output of the phase

The output of this phase rendered quality data that has been verified by the users. The data sets that emerged from the information systems would be used in the next section for analysis purposes. These data sets will be described in detail in the next chapter.

3.6.4.1 Quality of Data

One of the first steps in creating an analytical solution is to understand the data and transform it into a relevant format for further analysis. For the data to be trustworthy in the decision-making process, it should adhere to the characteristics of completeness, relevance, accuracy, timeliness and accessibility.

Relevance

Relevance was viewed as referring to the ability of data to meet current and potential user needs. It was important for the WCED to collect the relevant data so that important educational development could take place. The relevant data should be available for developing of education indicators, making high level decisions and determining critical success factors. Relevance, as an essence of data, refers to the fact that there is no point in collecting any data, unless one puts it to some meaningful use in solving policy decisions.

Completeness

Data is complete when data values are present for all records, occurrences or logical entities that require them within the database. Completeness as a constituent of data was important in the decision making process.

Accuracy

The degree of accuracy required of the data depends on the purpose to which the data will be put. Accuracy is important and should be reflected in the data capturing methods, data

validation methods and the data verification processes, possibly at the expense of time and money.

Timeliness

Data is timely when it is current or up to date as defined by the owner of the data. Data must be on time and available when it is required, otherwise the credibility of the information system diminishes. In this process to maintain the timeliness of data, cost and accuracy could sometimes be impacted. An understanding of the timeliness characteristics of the system contributes to the selection of appropriate data collection techniques.

Accessibility

Data should be accessible to the user. The user should know what data are available, where to find the data and how to retrieve the data. Metadata and data dictionary are important concepts to enhance the accessibility of data within an organization.

3.6.5 Key actors

The key role players during this phase were the data capturers and the users.

3.7 Concluding remarks on the information systems development

Part one of the study (chapters two and three) described the first three phases of ISD. The object of this part was information systems development (ISD). The objects of ISD are the education management information system (EMIS), the key actors and the quality data emerging from EMIS practice. In this section the processes, procedures and the key actors of

ISD were explored and provided a framework for developing and implementing an ISD in a large education department.

The research has explored an empirical method for studying information systems development as a work practice (Practice-in-Action) in a particular context (education in the Western Cape). This study of EMIS in WCED illustrated how improvisation, sensemaking and bricolage could play a vital role in ISD in a particular situation (context). This research approach indicates that although information systems development (ISD) reveals certain general characteristics, it is embedded in a particular reality, always as a particular matter in specific situations, within particular circumstances. This practical approach enlightens the reader about the essential role of improvisation, bricolage and sensemaking in education management information systems development.

The adoption of the Practice-in-Action approach was made possible through the application of the ISD model in a particular context. The ISD processes were described in this study according to the in-house developed method of Practice-in-Action. This practice lends support to the “alternative viewpoint that information systems development is a creative, problem-solving activity much more akin to a craft than an engineering discipline” (Lang, 2006:7).

The notion of “particularity” enabled improvisation, bricolage and sensemaking as systems development practices that are grounded “not on mere compliance with prescribed methods but rather more on the intuition of experienced developers [and users] as to how to act within the unique constraints of individual projects” (Lang, 2006: 7).

The contextual understanding of ISD that allows for sensemaking, bricolage and improvisation was emphasised in the study. ISD is not only a technical matter, but with sensemaking, bricolage and improvisation it might play a useful role in sensitising ISD to social, political, educational and cultural issues. According to Lang (2006: 6) “within the field of information systems development, ‘hard’ approaches have historically dominated. According to this philosophy, design is rational and objective, and design processes can be modelled systematically, thus giving rise to suggestions that information systems can be ‘engineered’ much like other structures in the man-made world”. The Practice-in-Action espouses the alternative approach that ISD is a

“deviation from the established ways of informing ISD (through methods, models, and guidelines) and localism and [particularity] contrasts with the universalising tendencies of ISD research by always attending to the local practices and specificities through which particular systems come into being” (Henriksen, 2003: 3).

This is the kind of situation that presents itself in any organisation, because nothing is stable and predictable. To be able to improvise in this manner the developer (as the sensemaker/improviser and the bricoleur) needs to have the authority to act when the need arises.

The next two chapters (four and five) form the second part of the thesis and provide an overview and contextual understanding of data analysis within information systems development. It describes phase four (Data Analysis and Reporting) of the developed ISD model (refer to Table 2) and is statistical, analysing various data sets to elicit knowledge and to demonstrate their usefulness for improving educational management. This part attempts to locate Data Analysis and Reporting as the last phase of the conceptual framework model within the context of this research in particular and EMIS in general. The structure of the two

chapters (four and five) follows the systems concept of input-transformation-output through the application of quantitative methods. Data analysis is an essential part of an EMIS and needs quality information to be useful. This means that an ISD model should be accompanied by data analysis practices. In an EMIS overview for UNESCO, Wako (2003:28), amongst other *raison d'être* practices, outlines data analysis as “looking more closely at the data, and in various ways, in order to extract information useful for planning and decision-making: ascertaining trends, comparing provinces, districts, urban and rural schools, boys and girls”. Data analysis practices are important to serve national educational policy considerations.

The next part is the final phase of the ISD model and uses the data sets produced in this section as input for data analysis. Two quantitative methods will be applied in chapters five and six in order to make sense of the data sets that emerged from the information systems development process.

CHAPTER FOUR

USING QUANTITATIVE METHODS TO MEASURE THE IMPACT OF SELECTED VARIABLES ON STUDENT ACHIEVEMENT

4.1 Introduction

Data analysis forms an integral part of the ISD process and to illustrate the possibilities thereof, quantitative methods were applied on data sets of EMIS and other information systems to complete the model. Given this and the context of the particular situation, the study utilised bricolage design (using existing data sources) in order to make sense (analysing the data sets) of the education system. Using this process model the study illustrates that bricolage and sensemaking exist also in the data analysis phase as an integral part of information system development process.

This chapter is divided into two distinct parts. The first part reviews the literature on the education production function in South Africa and compares it with similar studies in other countries. The second part is an analysis of the production function using data sets from different information systems. The main purpose of this chapter is to investigate the relationship between resource variables and student achievement across 319 secondary schools in the Western Cape in 2002. Different quantitative methods were used to determine whether these variables contribute to the expected Grade 12 results in the Western Cape. Basic

statistical techniques were used to determine which variables should be included in a multiple regression analysis explaining student achievement.

4.2 Different data sets as input for the application of statistical techniques

The input to this phase is the quality data that was produced in the data processing phase (Chapter 3). The data analysis practices in particular the application of quantitative methods were based on these data sets.

4.2.1 Sensemaking as an ongoing process

It is evident that ISD is a continuous process and the output of one phase becomes the input to another process. Sensemaking is therefore a continual process of understanding the input in the context of the specific development phase. Therefore it is critical that the foregoing processes be based on quality inputs, because it affects all the work that will follow. The participants in these processes have to understand and make sense of their work practice to realise the significance of the data for decision-making.

The output (the quality data) of the previous phase will be used as input for the data analysis in this phase, illustrating the continual process of ISD. According to Weick (1995) sensemaking is an “ongoing” process and has no definitive starting point.

4.2.2 Bricolage , sensemaking and information systems development

The previous chapter (Chapter 3) was a study of EMIS that shows how improvisation, sensemaking and bricolage could play a vital role in ISD in a particular situation. To apply statistical methods in a meaningful manner more data sets are required than those available in EMIS. Other information systems exist that are important data sources for analysis purposes. These sources could complement EMIS to develop a more comprehensive management information system that is effective and efficient in providing complete, relevant, accurate, timely and accessible information for reporting, analysis and decision-making purposes. By applying the bricolage approach, the existing data sources were enhanced by integrating data from Census, PERSAL and examinations with EMIS. Thus, the first step towards a good management information system is to have operational systems in place that provide the necessary data sets.

Integration as bricolage design

All the data sets mentioned below will remain fragmented islands of data and will exist in isolation of each other if it is not linked. This unintended fragmentation could be avoided by transforming the data into relevant format for further analysis. From a bricolage design perspective, such existing data from other information systems will be used in this study to create an overarching information system. The bricolage approach provides a more comprehensive information system through linking, modifying and reorganizing existing datasets to fulfill new purposes. This is done through **data integration**. Organisation-wide information integration is recognised as a significant part of the solution for the problem of fragmentation and disparate databases. Hence, the assumption is that in the Education

Department there is a need for an information integration solution that will enable education management and policy makers to analyse data that have been collected for different functions and purposes.

Data integration generally means linking different data sources through the use of a common field across a collection of data sources. Without creating new information systems the study integrates existing systems to create a more comprehensive information system.

Integration as sensemaking

This bricolage design approach outlined here links closely with the sensemaking dimension which Weick (1995) refers to as “enactiveness” (This explains how entities get there in the first place and that people often produce part of the environment they face). Integration from a sensemaking perspective is to create a new environment or reality (information system) that makes quality data available for data analysis. The function of EMIS is therefore more than just data collecting, but rather a system where all data in the government are managed in an integrated process within and across Departments so that it is available for strategic decision-making analysis. From this integrated information system, the different variables emerge that are used to apply the statistical techniques utilised in this chapter.

It is an appropriate point in the chapter to specify these existing information systems from where these data sets come that are used as input in this chapter and contribute to the bricolage design approach. These information systems offer the following data sets:

Education Management Information System (EMIS)

This was the information system that was explored in this study as a work practice (Practice-in-Action) in a particular context (education in the Western Cape). The essence of EMIS is that education departments annually collect data from all schools through a comprehensive survey. The National Education Department in collaboration with the Provincial Education Departments has compiled a survey questionnaire to ensure that the relevant data sets are collected. This will facilitate, among others, the creation of education indicators, policy planning and implementation; decision-making; and monitoring and evaluation of the education system. The Annual School Survey (ASS) is a comprehensive survey of all public and independent schools in the Western Cape. The survey was designed to provide comparable information on public and private sectors, as well as trend data over time. Information included in this survey is: general description of schools including name, address (physical and postal), phone numbers, e-mail; data on learners including learner ages per grade, home language, medium of instruction, learner subjects; information on educators.

The tables with the relevant data elements necessary in an Education Department that form part of the data sets for analysis purposes in this study include:

1. Learner information
 - a. Enrolment
 - b. Ages
 - c. Home Language
 - d. Language of Instruction
 - e. Subject Enrolment
 - f. Repeaters

g. Failures

2. Staff information

a. Educators

b. Non-teaching staff

c. Support Staff

3. Physical Infrastructure

An important function of EMIS should be to compile management information systems with data sets not only of surveys but with relevant data sets within the Education Department and across other Departments, such as Health, Social services, Housing, Census, and more. The origin of the variables that are used in the analysis are indicated under the column heading “Data Sources” in Table 4.

Census 1996 and 2001

Census (Statistics South Africa) is a national survey for the entire population that takes place every few years. Census data sets have been used in this research to calculate a poverty index for suburbs and to calculate educational attainment data sets, such as highest educational level (grade) obtained per person and children ages 15 to 19 within a suburb. Data sets from census 1996 and 2001 have been used. The particular year of the census data set is indicated in the section where the data set is used. The poverty index that was used in this chapter was created based on data sets obtained from the Census 1996 (refer to Appendix B for calculation formula).

PERSAL (personnel salary system)

This is a national operational system to manage the salary related issues, such as leave, salary, housing subsidy, qualifications, years of experience, nature of appointment, etc. for all staff. The experience and qualifications of teachers are data sets obtained from this source. Table 4 indicates the variables that were obtained from this data source.

Examinations

The examinations system is a national operational database used to manage the grade 12 examination. The aggregate mark of grade 12 indicated as the dependent variable in Table 4 has been obtained from the examinations data set.

4.2.3 The variables used as input for quantitative analysis: An exploration from a sensemaking perspective

Table 4 summarises the key variables that were used in the analysis⁹. These are input variables that have been produced through ISD and emerged from the data sets of information systems such as PERSAL, examinations, EMIS and Census 1996. The utilisation of these key variables that stem from the ISD indicates the ongoing dimension of sensemaking (Weick, 1995).

⁹ See Appendix B for a detailed explanation of these variables.

Table 4: Variable List and Description

Variable	Code	Description	Data Source
Dependent Variable			
Grade 12 Aggregate	Aggregate	Calculated by averaging the aggregate of the learners of a particular school.	Examination
Independent Variables			
Poverty Index of the Community	Poverty Index	An overall rating of each community's socio-economic status.	Census 1996
Average age of teachers	Age	Average age of all teachers at a particular school.	PERSAL
Average experience of teachers	Experience	Average teaching experience of all teachers at a particular school.	PERSAL
Attainment rate	Attainment rate	Learners enrolled in grade 8 in 1998, expressed as a percentage of grade 12 learners in 2002 of a particular school.	EMIS
Average qualifications of teachers	REQV	Average qualifications of all teachers at a particular school.	PERSAL
REQV 14	REQV 14	Number of teachers with REQV 14, expressed as a percentage of number of teachers at a particular school.	PERSAL
REQV 15	REQV 15	Number of teachers with REQV 15, expressed as a percentage of number of teachers at a particular school.	PERSAL
REQV 16	REQV 16	Number of teachers with REQV 16, expressed as a percentage of number of teachers at a particular school.	PERSAL
REQV 17	REQV 17	Number of teachers with REQV 17, expressed as a percentage of number of teachers at a particular school.	PERSAL

Number of teachers paid by the school	SGB teachers	Number of teachers who are privately paid by the school's governing body.	EMIS
Number of teachers paid by the state	State teachers	Number of teachers paid by the state per 1000 learners at the school.	EMIS
Ex Cape Education Department	CED	Dummy variable indication previous education department of school (previously advantaged schools).	EMIS
Ex House of Representatives	HOR	Dummy variable indication previous education department of school (predominantly from the coloured population).	EMIS
Ex Department of Training	DET	Dummy variable indication previous education department of school (black learners under the previous government before 1994).	EMIS
Number of learners who repeated grade 11	Repeaters	Number of grade 11 repeaters in 2002, expressed as a percentage of grade 11 enrolment in 2001.	EMIS
Dropout rate	Dropout rate	Dropout rate of school.	EMIS
Metropole central	Metropole central	Dummy variable for a specific education district of a particular school.	EMIS

4.3 Transformation of data by using statistical methods to analyse the resource variables that influence student achievement

To make sense of the data sets that emerged from EMIS (application development, data collection and data cleaning processes) it was integrated with the other data resources (PERSAL, Examinations and Census 1996). This chapter reports on quantitative techniques used to facilitate the process of knowledge discovery in large data sets. The research further

attempts to show how the application of quantitative methods and statistical analysis could be used to generate insights into organisational sensemaking. Making sense of a body of data is common activity in any kind of analysis. Below is a brief exposition of how statistical methods were applied to this end.

4.3.1 Scatterplots to explore the bivariate relationship between dependent and independent variables

A scatterplot is a useful summary of a set of bivariate data. It gives a good visual picture of the relationship between the two variables, and aids the interpretation of the correlation coefficient (shown as r in each of the scatterplots presented here) or regression model. Each unit (a school) contributes one point to the scatterplot, on which points are plotted but not joined. The resulting pattern indicates the type and strength of the relationship between the two variables (Scatterplot 2004:1). School performance is the result of a multi-dimensional effect of many variables. The bivariate relationship of grade 12 aggregate and other key variables presented by the following scatterplots show the significance of the relationship as well as non-linearities that might exist in the data.

Repeaters

Figure 4 reflects a scatterplot of the bivariate relationship between the repeater rate for grade 11 learners (refer to Appendix B for calculation) and the aggregate grade 12 results.

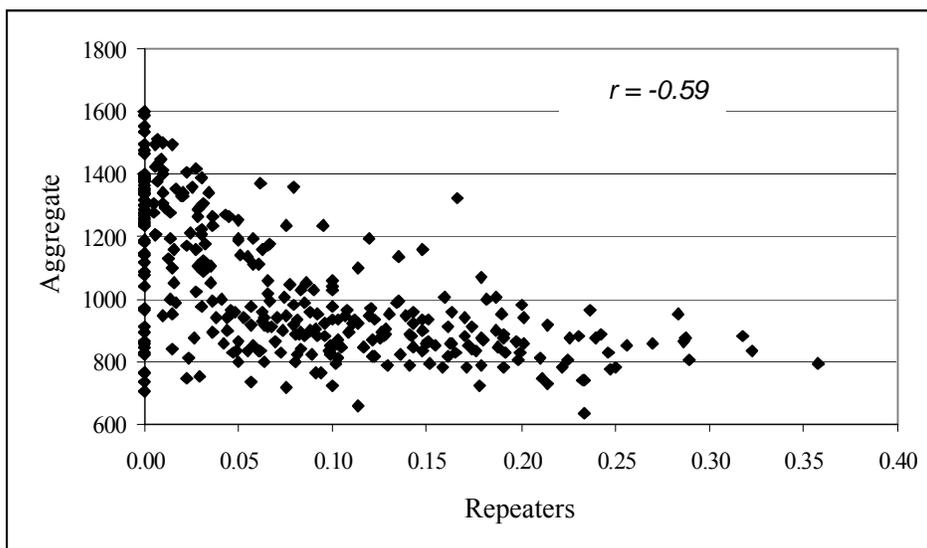


Figure 4: Aggregate vs. Number of Learners who Repeated Grade 11 (Repeaters)

Figure 4 displays a negative relationship between the aggregate matric mark and the repeater rate for grade 11 learners for 319 schools in the Western Cape. This is an indication that weeding¹⁰ (also called skimming) probably played little or no significant role in the grade 12 pass rate. It must be noted here that this only relates to the holding-back of learners from the previous grade (grade 11) to maximize the grade 12-pass rate. The study however, does not deal with the possibility that learners may be pushed out altogether. That would require not only looking at the repetition rate in grade 11, but also the dropout between 11 and 12. There is some evidence (refer to the high dropout rate between grades 10 and 12 as reported in the

¹⁰ Weeding is the process whereby schools only allow the best learners to be promoted to Grade 12 in order to achieve better matriculation results.

next Chapter) that the repetition is being pushed back to grade 10 as a reaction to the pressure to have good matric pass rates.

Poverty Index

The overall rating of each community's socio-economic status is based on the following indicators: (i) income, (ii) educational attainment, (iii) unemployment, and (iv) lack of water supply (Refer to Appendix B for a detailed explanation on the calculation of the poverty index). An index of 0 indicates a rich community whereas an index of 1 indicates a poor community. Figure 5 displays the scatterplot for the relationship between grade 12 aggregate marks and the poverty index of the community.

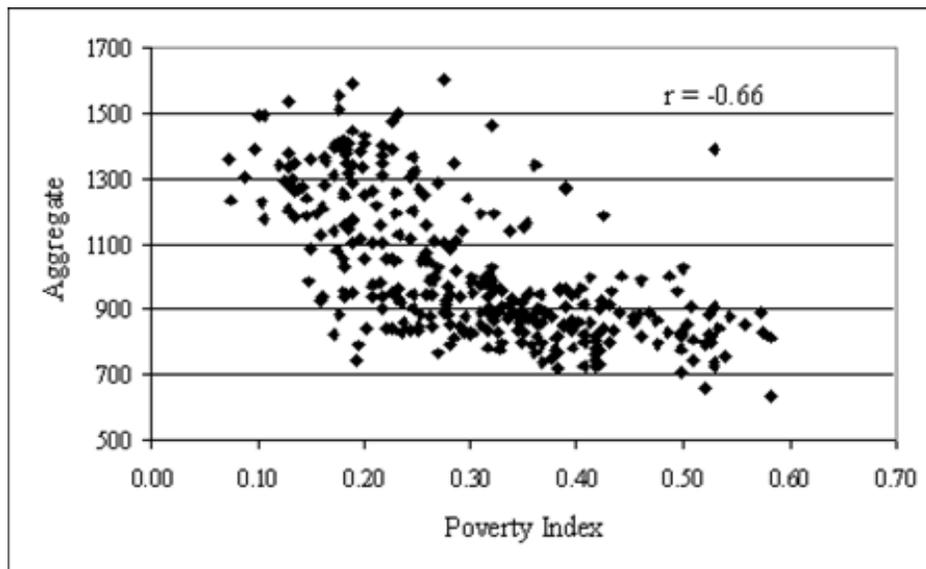


Figure 5: Aggregate vs. poverty index of community

Figure 5 shows a significant positive correlation between school performance and poverty. The higher the socio-economic status (SES) of a school, the better it performs.

SGB teachers per 1000 learners

Figure 6 displays the relationship between the teachers paid by the school's governing body per 1 000 learners and the grade 12 aggregate marks.

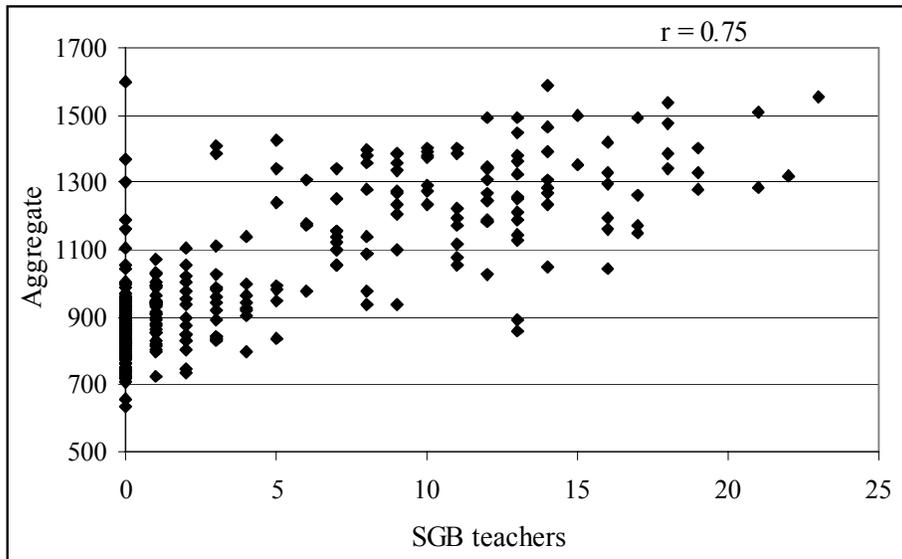


Figure 6: Aggregate vs. Number of Teachers per 1 000 Learners Paid by the School (SGB Teachers)

Figure 6 displays a positive relationship between aggregate and SGB teachers (teachers paid by the school). The more extra teachers per 1 000 learners a school can afford, the better the school performs. However, it may only reflect the correlation between aggregate and the socio-economic status of the school, since schools with higher socio-economic indices also tend to have more privately paid teachers. Note also the very large differences amongst schools with no SGB teachers, as depicted on the vertical axis.

Average qualifications of Teachers

Refer to Appendix A for the definition of requirements of the relative education qualification value (REQV) and to Appendix B for the calculation of average qualifications of teachers.

Figure 7 displays the relationship between REQV and the aggregate of grade 12 marks.

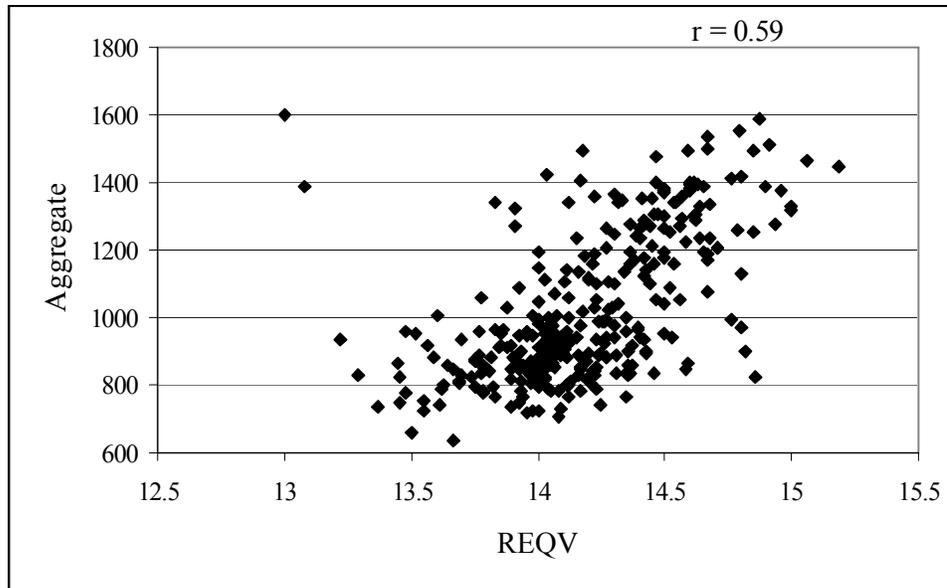


Figure 7: Aggregate vs. Average Qualifications of Teachers (REQV)

Figure 7 shows a strong positive correlation between the aggregate mark and the average qualifications of teachers. Thus the higher the qualifications of teachers, the better the grade 12 results. Once again, this may be linked to the role of socio-economic status, as richer schools usually have better qualified teachers.

Teachers paid by the state

The number of teachers paid by the state per thousand learners indicates the educator-learner ratio. Figure 8 shows the relationship between the teachers paid by the state per 1 000 learners and the aggregate marks of the grade 12.

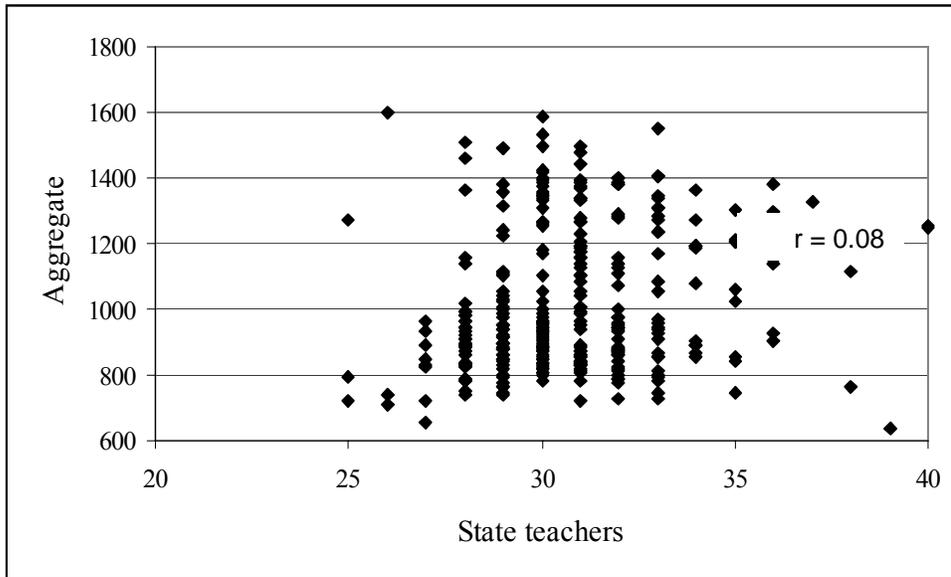


Figure 8: Aggregate vs. Number of Teachers per 1 000 Learners Paid by the State (State Teachers)

Figure 8 shows that grade 12 results are not significantly correlated with the number of teachers per 1 000 learners paid by the state.

Experience of Teachers

Figure 9 displays the relationship between experience of teachers and the aggregate marks of grade 12.

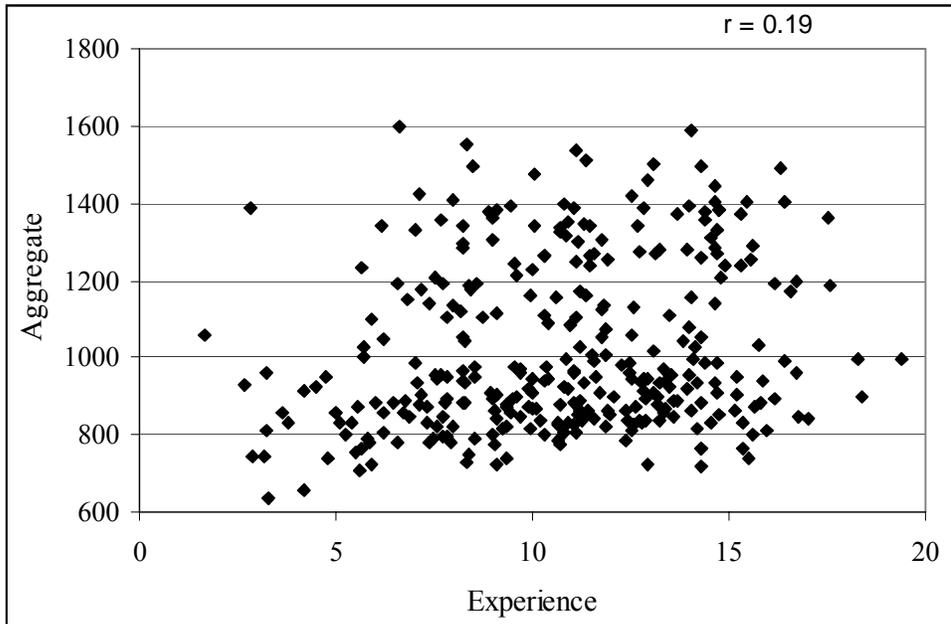


Figure 9: Aggregate vs. Average Experience of Teachers (Experience)

According to Figure 9, there is a weak relationship between the grade 12 results and the average experience of teachers.

4.3.2 Correlation Coefficient (r)

A number of correlations were compiled between grade 12 and other primary variables. Table 5 summarises the correlations obtained from these calculations.

Table 5: Correlation between Aggregate Grade 12 Results and Other Variables

Variables	Correlation coefficient
Average qualifications of teachers	.587**
Dropout rate per school	.351**
Percentage of learners who failed in grade 11	-.572**
Poverty index of the community	-.656**
Number of teachers paid by the school per 1 000 learners	.755**
ex-Cape Education Department	.790**
Average age of teachers	.621**
Average experience of teachers	.193**

*** Correlation is significant at the 0.01 level*

Table 5 shows the correlation coefficients between aggregate grade 12 results and other variables listed in the Table 4. It is clear that there is a significant negative correlation between the aggregate grade 12 results and poverty of the community. The coefficient has the expected sign. One must remember that if the socio-economic status index is lower, the situation is better (1 is poor and 0 is rich). The correlation coefficient between aggregate and CED is statistically significant and higher than the coefficients of the other variables. The number of teachers paid by the school is positively associated with grade 12 results. These are simple bivariate relationships and many other intervening variables that could affect the results, need to be considered. A multiple regression analysis is the technique used to consider the simultaneous influence of more than one variable.

4.3.3 Multiple Regression Analysis

There are many ways that analyses could affect sensemaking, this study considers the production functions as an approach to analyse the data and construct meaning and knowledge from the data.

4.3.3.1 Introduction

The approach in this section is to use multiple regression analysis to determine the different variables that have an influence on the grade 12 results of schools in the Western Cape. Ordinary least squares (OLS) regression analysis was conducted to determine the relationship between grade 12 results and school and teacher characteristics. The number of students who wrote the grade 12 examination in each school was used as weights in the analysis. The predictors included teacher related variables (qualification, experience, and age), school variables (teachers per 1000 learners paid by the state, teachers per 1000 learners paid by the school) and a socio-economic variable as a proxy for poverty. The dependent variable was the average aggregate of grade 12 results.

The theoretical framework is an educational production function¹¹ where grade 12 results is a function of school inputs, the quality of such inputs and socio-economic status. This allows the research to ascertain how much of the differences in school performance are explained by

¹¹ A mathematical expression of the relationship between inputs and outputs as defined by Levin (1993). Levin (1993: 2) cited David Monk who in his 1990 book, *Educational Finance: An Economic Approach*, uses the production function as the basic element in studying productivity in schools. According to Levin, Monk defines a production function as a model, which links conceptually and mathematically, the outcomes, inputs, and the processes that transform the latter into the former in schools.

teacher and school characteristics. It is important to note that the grade 12 results and the teacher qualifications were individual level data from which school level averages were constructed, while the socio-economic status measures were school level averages.

Multiple regression is a quantitative method

“to account for (predict) the variance in an interval dependent, based on linear combinations of interval, dichotomous, or dummy independent variables.

Multiple regression can establish that a set of independent variables explains a proportion of the variance in a dependent variable at significance level (significance test of R^2), and can establish the relative predictive importance of the independent variables (comparing beta weights).” (Multiple Regression 2003:1)

With multiple regression more than one independent variable is included in the equation that acknowledges that the dependent variable is influenced by more than one dimension. We follow the same general multiple regression equation as Lewis-Beck (1995:53), where the dependent variable is seen as a linear function of more than one independent variable:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots b_kX_k + e$$

where Y is the dependent variable; $X_1, X_2, X_3, \dots, X_k$ are independent variables; a is the estimated intercept (constant); b_1, b_2, b_3, b_k are estimated partial slopes (coefficients); and e is the error term.

4.3.3.2 Literature Review of Education Production Function Analysis in the USA

There is a vast body of literature where education production functions are used in empirical studies to evaluate the effect of input resources (school, family, community, etc.) on student achievement (Bohte, Polinard, Meier & Wrinkle, 1999; Clotfelter & Ladd, 1996; Coleman, 1966; Hanushek, 2002a,2002b; Hanushek & Raymond, 2002a, 2002b; Krueger, 1999; Ladd, 1996; Meier, Wrinkle & Polinard, 1999; Mendro et al., 1999:8; Meyer, 1997; Pritchett & Filmer, 1999; Rothstein, 2001).

There has always been a need to evaluate the contribution of the school towards the achievement of students. The basic premise of virtually all proposed accountability systems is that student performance should be the key element (Hanushek & Raymond, 2002a: 1). Clotfelter and Ladd (1996: 26) states that it would be unfair to compare schools simply on the basis of their students' scores and contends that using absolute test scores, as a measure for school performance is an imperfect and limited strategy. Production functions try to consider this by controlling for socio-economic status and the level and quantity of inputs.

4.3.3.3 Education production function analysis: A South African perspective

4.3.3.3.1 Empirical strategy and theoretical framework

In economic terminology, efficiency in education could be measured according to the nature of the relationship between inputs and outputs. Increasingly the education production function approach, a mathematical expression of the relationship between inputs and outputs, has

become an important methodology for determining the effect of various resources (inputs) on education performance (outcomes). This section will review the literature on the education production function in South Africa and compare it with similar studies in other countries. School outcomes are often used to determine the quality and the efficiency of the education system. Standardised test scores are regarded as the best indicators to measure school performance. In South Africa, grade 12 results were the only available measure in the past. Assessment results should help to identify the effects of educational inputs on student achievement. The analysis of these data sets could provide decision-makers with information to develop a framework for addressing future educational problems.

The growing literature on student achievement and school performance in South Africa is of particular importance for this study (Archer, 1995; Case & Deaton, 1999; Case & Yogo, 1999; Chetty, 1992; Crouch & Mabogoane, 1998a; Crouch & Mabogoane, 1998b; Deaton & Case, 1996; Fiske & Ladd, 2002; Fiske & Ladd, 2003; Fuller & Prema, 1994; Gustafsson, 2005; Hosking, 1994; Van der Berg, 2005a; Van der Berg & Burger, 2003; Van der Berg & Louw, 2006). Most of these studies analysed the contribution of education inputs to effective schooling is concluded that resources appear to matter less than the efficiency of their application. This supports the aim of the present study that intends to explore the importance of quantitative methods for data analysis in order to improve management efficiency with regard to the input of resources.

Research on school performance in South Africa has focused on the use of matriculation results as a measure of performance. This has brought some illuminating insights, amongst others, poverty was found to be an important predictor of school performance. However, a

large part of differences in school performance cannot be explained by either socio-economic status or resource inputs. This phenomenon requires further analysis, as most researchers have ascribed issues of efficiency within schools to management.

Crouch and Mabogoane (1998a) pointed out that educational reform in South Africa needs more than just a redistribution of physical inputs into education. It requires a redress of educational output, or cognitive outcomes. They argue that the unexplained variance in terms of school management and the culture of learning is key to efficiency improvement.

The approach of Crouch and Mabogoane (1998a) encapsulated the essentials (i.e. outcomes, inputs, and processes) of the production function when they explored “how resources are deployed to achieve a high output/input ratio, as part of changing from the formerly segregated and unequal education system to a unified one” (p.2). They noted: “improvement in ‘production’ and management processes taking place within schools and the Department is crucial to achieve equalisation of quality education in a new education system” (Crouch & Mabogoane, 1998a: 2).

Crouch and Mabogoane (1998a) reviewed the impact of inputs on learners’ cognitive development based on studies carried out in developed and developing countries. Though there is a debate, most of the studies suggested that the impact of traditional material inputs is limited in developed countries. However, studies done in the developing world seem to suggest that schools have greater impact on learners, and therefore inputs seem to matter more. Crouch and Mabogoane, however, specifically focused on the studies carried out in South Africa. They indicated that earlier quantitative studies conducted by economists in South

Africa, on what makes for effective schools, tended to emphasise inputs into education rather than the management of the “production” processes.

According to Crouch and Mabogoane (1998a), a further difficulty with production function studies in South Africa is the use of data that was gathered for purposes other than educational analysis. The fundamental focus of their research was:

- to rely on databases that include true, solid, and linked measures of cognitive achievement, inherent cognitive ability as a background variable, school-level resources and student background and,
- to concentrate on understanding the unexplained variance in terms of school management and culture of learning processes as key to efficiency improvement.

Another contributing factor in the analysis of outcomes is the level at which it should be measured. This depends largely on the availability of the data. The data used in their analysis have been collected at school level. The basis of their method was to rank schools according to:

- an index of absolute achievement. The six components used to produce the weighted index of absolute achievement were Matric passes, Matric exemption, English passes, Maths passes, sports and culture.
- a relative index of achievement. In order to correct for resource usage and historical advantage, they regressed each of the six components on the various resources and the advantage variables and stored the residuals. The residuals for each of the models were summed to construct a relative achievement variable R.

In this regression analysis Crouch and Mabogoane used different urban areas as a proxy for socio-economic status and availability of deputy principals and heads of departments as proxies for management. The same weights as those used for the absolute achievement were used in constructing the Relative Achievement Index. The socio-economic indicator could enable them to identify schools that were successful in spite of adverse conditions.

The dependent variable in the production function analysis that is of particular interest to this study is the matriculation pass percentage (Crouch & Mabogoane, 1998a: 9). The independent variables in the model that are hypothesised to influence the matric pass percent are dummy variables for formal townships, urban high-rise and suburbs with low density flats, resources, head of departments. Their framework of analysis is a standard linear production function:

$$Y = 0.539 + 2.0E-03 * X_1 + 3.7E-03 * X_2 + 3.9E-03 * X_3 + 4.2E-05 * X_4 + -1.0E-02 * X_5$$

(4,0) (2.8) (6.7) (3.1) (2.2)

where Y is the dependent variable matric pass percentage and the independent variables are X₁ = formal townships, X₂ = urban high-rise and suburbs, X₃ = low density flats, X₄ = resources and X₅ = head of department. The t-values are in brackets and the R² of 0.479, according to them is higher than in other studies, but not enough as a basis for making recommendations on social policy.

Based on this method they could construct an absolute index and a relative index for each school. By comparing the actual rating (absolute index) with the predicted rating (relative index) they could determine over-achiever schools relative to their resources.

In another study, Crouch and Mabogoane (1998b) used the production function to rigorously control for the influence of resources and poverty on a school's performance. According to them over the previous few years, as South Africa attempted to transform its education system, much of the intellectual and bureaucratic effort had gone to the redistribution or allocation of quantitative resources. They specifically mentioned:

- the school funding norms whose emphasis was mostly the redistribution of material and financial resources
- educator re-deployment, based on quantitative norms relating to learner: educator ratios.
- education budgets, and
- building of schools in areas which previously have not been well-provided

In this study the education production function was again employed to show that the matric pass rate in Gauteng was a function of the dummy variable township, average qualification of teachers, media centre resources, average post of educators, computers, poverty index, and deputy principals:

$$Y = -127.1 + -25.5*X_1 + 16.3*X_2 + -7.19*X_3 + -12.0*X_4 + .314*X_5 + -21.5*X_6 + 2.74*X_6$$

where Y is the dependent variable matric pass rate in Gauteng and the independent variables are X_1 = dummy variable for township, X_2 = average qualification of educators, X_3 = number of rooms devoted to media centres, X_4 = average post ranking of educators, X_5 = computers used for computers, X_6 = deputy principals.

The production function results enabled them to predict the matric pass rate based on the inputs and compare it with the schools' actual results. They emphasised the importance of the residuals when they stated:

“The remnant, namely the degree of performance that cannot be explained away via the resources and environment must have to do with the quality of management. This is important from a management point of view, because it gives us a method that allows us to develop, for every school a reasonable performance expectation of its management” (Crouch & Mabogoane, 1998b: 2).

They asserted that for the first time data from four separate databases have been integrated, namely the EMIS, School Register of Needs (SRN), EXAMS and SOCIO-ECONOMIC DATABASE to examine the correlation between all these factors and to try to elicit the effects of these factors on learning. This links to the bricolage design approach described in this study where data sets from different data sources were integrated to create a new organisation-wide information system.

In their description of the analysis, they pointed out that poverty has a greater influence on matric pass rate than resources, “in the sense that the t-values and the coefficients for the township and poverty variables are so much higher than the resources variables' coefficients and t-values: being in the poorest township tends to imply a pass rate that is 47 (25.5 + 21.5) points lower than it would be in the best of suburbs, even supposing equality of resources” (Crouch & Mabogoane, 1998b: 11). Resources therefore did not have the desired expected outcomes, in the light of the intellectual attention and bureaucratic effort that has gone into the re-distribution or allocation of quantitative resources.

Crouch and Mabogoane (1998b: 12) emphasised the importance of teacher qualification and based on the production function results (one year training is associated with 16 point increase in the pass rate) suggested that:

“all of the disadvantage of being in a poor township school could be overcome via whatever knowledge and behaviours teachers with 3 more years of training exhibit over other teachers (since $3 * 16 = 48$ approx. = $25.5 + 21.5$)”.

The estimate of teacher qualification variable in their research showed that the general opinion of “paper qualifications” that do not count indeed was not correct.

Furthermore, Crouch and Mabogoane emphatically underlined the fact that it

“is difficult to fight the negative effect of poverty with mere resources, unless they are well-managed, and unless one pays more attention to the qualitative nature of the resources, but importantly, if one pays attention to whatever qualitative behaviours are associated with educators with high REQVs - a term used in South Africa to indicate level of teacher qualification” (Crouch & Mabogoane, 1998b: 12).

Crouch and Mabogoane (1998b) concluded by using the production function results to predict the likely performance of a school given its resources and its environment, which then enabled them to compare the predicted or expected performance with the actual performance. According to this method they could establish how much a school over- or under-performed relative to expectation. Their finding showed that just comparing schools based on actual school performance is not a fair reflection of schools performance, because in this instance no poor school is amongst the good performers. However, when the predicted performance is taken into account, 43% of the over-performers are township (poor) schools. They therefore underscored the fact that by using only the actual performance as a measure to rate schools, significant inequalities in educational achievement were hidden.

Van der Berg and Burger (2003) have similar questions when they stating “to what extent persistence in school performance inequalities is attributable to resource allocation and to what extent it is attributable to differences in efficiency of resource use” (Van der Berg & Burger, 2003:1).

This afore-mentioned phenomenon concurs with the work of Crouch and Mabogoane (1998a; 1998b) that used the education production function in the effort to identify the influence of input resources and poverty on a school’s performance. Van der Berg and Burger (2003) attempted to provide a systematic and quantitative perspective of school level performance in their analysis of resource allocation and efficiency as determinants of school performance. The difficulty to obtain school level data for analysis purposes and to measure school performance was also pointed out in their research. The effective application of the production function for analysis necessitates the availability of comprehensive, accurate, relevant and complete data sets. Van der Berg and Burger also highlighted the nature and the extent of the residual that in their analysis will be interpreted as a proxy for the omitted variable, namely efficiency.

Van der Berg and Burger (2003) constructed five regression models based on the school level data sets of six provinces in the country to consider the empirical relationships of school performance with socio-economic indicators and resource investments on a national level. The coefficient of determination in all five equations was 0.55, which corresponds with the results obtained by other researchers in South Africa.

The first regression showed that the socio-economic background of students, as measured by matriculation pass rates, influences the performance of schools by the availability of teachers, teacher quality, and the racial composition of schools. The equation showed that predominantly black schools still appeared to be functioning comparatively poorly, even after controlling for their lower socio-economic background (Van der Berg & Burger, 2003:17). The second regression equation illustrated that the Western Cape outperforms equivalent schools nationally by almost 8 percentage points (Van der Berg & Burger, 2003:18).

In the third equation the use of interaction variables indicated that predominantly coloured schools in the Western Cape outperformed the national average for similar schools by 10.1 percentage points. Formerly white schools in the Western Cape outperformed the national average for similar schools by 6.7 percentage points (Van der Berg & Burger, 2003:18).

Regressions four and five demonstrated the impact of interaction effects between Western Cape location and measures of socio-economic background and teacher inputs. The impact of the quality and availability of teachers seemed to have a greater effect in the Western Cape than nationally (Van der Berg & Burger, 2003:17).

In this part of their analysis, Van der Berg and Burger also performed a production function for the schools in the Western Cape only, based on the same dataset as the one used for the equations described in the afore-mentioned paragraphs. The coefficient of determination is this time much higher (0.67) than in the equivalent equation for the national schools which was 0.55. This implies that about two-third of the variation in performance between schools in the province can be explained by variables reflecting socio-economic status, teacher resources and racial composition (Van der Berg & Burger, 2003:18).

In the same research paper using another more complete and detailed dataset with more available variables, Van der Berg and Burger (2003:20) examined differentials in the school performance of 247 high schools in the Western Cape. The aim was to study the correlation between performance and both resource allocation and the efficient use of allocated resources, controlling for socio-economic background throughout.

Van der Berg and Burger (2003), like in many other studies (Cowley & Easton, 2004; Crouch & Mabogoane, 1998a; Mendro et al. 1999), constructed an overall and a mathematical performance index, because they regarded the pass-fail measures as too limited and restrictive. In examining the determinants of overall and mathematical school performance they used regression equations that included five variables, namely percentage of school pupils that were white, school fees per pupil, a poverty index for the area surrounding the school, pupil-teacher ratio, and an index measuring the average age of the teachers at the school. All the variables were highly significant in both the models except for average age of teachers and teacher qualification that were excluded from the equation for mathematical performance. The coefficient of determination in both equations was above 0.70, indicating that these variables influenced overall and mathematical performance in the Western Cape significantly. When they examined the overall performance for predominantly white schools in the Western Cape, they included the poverty index, learner-educator ratio and the index for average teacher age in the analysis. Their findings indicated that the overall performance was explained by these variables, based on the high coefficient of determination (0.75). When they regressed the poverty index and learner-educator ratio on overall performance for predominantly coloured and black schools only, they discovered that the coefficient of determination was drastically lower than for predominantly white schools. Van der Berg and Burger (2003: 26) concluded

by pointing out that resource allocation variables fail to contribute substantially to the explanation of results obtained by black and coloured schools. Their explanation of the low explanatory power of this model is that this could be ascribed to the substantial role of omitted variables, such as management.

Fiske and Ladd (2003:38) used data for 277 secondary schools in the Western Cape. Their findings indicated that teacher quality (average qualification and percentage of under qualified teachers), availability of teachers (learner-educator ratio and school governing body teachers) and school resources do affect matriculation pass rates and that former departments remain important when predicting school performance. They used community poverty and school fees to control for differences in learner backgrounds that may affect school performance. The coefficient of determination in both the regression equations is above .60, an indication that about 60% of the variation in the matriculation pass rate between schools in the Western Cape can be explained by variables reflecting socio-economic status, teacher quality, teacher availability, school resources and former government departments. These findings are in line with the results of Van der Berg and Burger (2003).

Case and Deaton (1999) used a production function to examine the effects of learner-educator ratios and school facilities on educational outcomes in South Africa immediately before the end of apartheid. The social, political and economic circumstances of the period just before 1994 completely differed from that of the current situation. The focus of their research was on the learner-educator ratio as a significant predictor of school performance. The research however was done in the period when huge discrepancies existed between race groups, based on the policies of the apartheid government before 1994 in terms provisioning of staff and

school facilities. It is important to note that in later research of this kind (Crouch & Mabogoane, 1998a; Crouch & Mabogoane, 1998b; Fiske & Ladd, 2002; Van der Berg & Burger, 2003) the learner-educator ratio did not appear in any equation as a significant determinant of school performance. The obvious reason could be that a small variation in learner-educator ratio between schools now exists which has no significant influence on learner outcomes in the country. One of the limitations in the research of Case and Deaton (1999) was the use of data that was gathered for other purposes. Another factor was that they conducted their research in the apartheid period when no integrated datasets existed between the many different government systems.

Learner-educator ratio is an explanatory variable in all the regressions by Case and Deaton (1999). Among the other controls are age, urbanization, gender, and various measures of family background. In their paper the analysis of educational attainment (measured as years completed) was used as the dependent variable, educator-learner as an explanatory variable, age was entered as a series of dummies, a dummy variable for gender, various household socio-economic characteristics (e.g. education of the household head, household size, household resources), as well as dummy variables for urbanization and provinces (Case & Deaton, 1999: 1069). The production function results in the research of Case and Deaton (1999: 1071) indicated a huge difference in the educational attainment between Whites and Blacks. The White population showed an education attainment rate of almost 100% for all ages whereas the Black population indicated that the average Black child obtains only two-thirds of a year education for each additional year of age. The learner-educator estimate for Blacks (-1.82) indicated that reducing the average Black class size from 40 to 30 learners per educator would raise the average educational attainment by 0.52 years, equivalent to adding

about ten months to age. Case and Deaton used interaction variables in their research all of which were significant in the regression. The interaction variable between learner-educator ratio and age showed that when the learner-educator ratio was halved, from 40 to 20, it would have little effect on attainment at age six, but would add three-quarters of a year to educational attainment by age ten, and a year and a half by age fifteen. An interesting contrast in the work of Case and Deaton (1999:1073) was in the effects of resources. Household resources had no effect on the educational attainment of Whites but a marked positive effect on the education of Blacks. The same appeared to be true for learner-educator ratio that had no obvious effect on educational attainment of White children. However, for the Black learners, where learner-educator ratios in primary and secondary schools were more than twice as big, higher learner-educator ratios had a strong and significant negative effect on attainment (Case & Deaton 1999: 1073).

4.3.3.3.2 Comparative Summary of literature review

Data description

In examining the determinants of school performance, using the education production function as theoretical framework, most researchers generally encountered the same kind of data problems. Translating learner performance into a fair measure of school performance requires taking into account a wide range of variables including socio-economic background of learners, the quality of teachers, availability of teachers, ability of students, different resources and facilities and management skills. Data representing these variables is not often readily available in South Africa for different reasons, namely:

- The data sets are not always integrated, so that the variables available in one data set are not always accessible from other data sets.
- The key input variables that are required for regression analysis purposes are not always available. Sometimes data is required at an individual level, but is only available at school level or even only at magisterial district level.
- Researchers are restricted to a few available data sets that narrow their findings. Consequently, rough proxies are often used for key input variables.
- Researchers often have to use data that was not gathered for educational purposes.

The central objective of this study is a bricolage and sensemaking design approach. The bricolage design approach is recognized and explained as the process of linking datasets from different sources to compile one comprehensive dataset. Based on this bricolage design principle, the study recognises the important contribution of exponents of production function studies undertaken in South Africa in the past decade.

Production function results as a measure of ranking schools.

Researchers in South Africa concurred that school performance cannot be ranked purely on their actual results. Based on the production function results it can be determined how well schools actually perform in relation to how well the statistical model predicts the school should perform based on the school's inputs and socio-economic status. The importance of using the residual in rating schools is emphasised in some studies in the country. Using this methodology, under-achiever or over-achiever schools can be identified. In comparative empirical research in other countries the same procedure is followed to rank schools.

Cowley and Easton (2004) used the production function results in their ranking of British Columbia's Secondary schools. Their basic approach in ranking a school could be

summarised as a measure of the success with which each school took into account the socio-economic characteristics of the student body. They used the formula derived from the regression analysis to predict the *overall rating* for each school. They then reported the difference (*actual rating vs. predicted rating based on the parents' average education*) between the actual *overall rating* and this predicted value in each school's results table. They then quoted specific examples to indicate which schools were more successful than others based on the regression results (cf. Cowley & Easton 2004:13).

Another example of this method is the Texas Educational Excellence Project, a program at Texas A&M University, which was dedicated to original research on the quality and equity of education in Texas. Researchers of the project used the Texas Assessment of Academic Skills (TAAS) as the dependent variable to evaluate school performance. The Texas Assessment of Academic Skills (TAAS) was a state-administered basic skills test in reading and mathematics at grades 3-8, and 10, in writing at grades 4, 8, and 10, and in science and social studies at grade 8.

In examining the Black student improvement on the TAAS examination Meier, Wrinkle and Polinard (1999) in the Texas Educational Excellence Project used the production function methodology to identify school districts that did a better job of educating black students. In the analysis of the relationship between student performance and economically disadvantaged students, black poverty, the education level of black families, class size, teacher experience, teacher qualification and other factors Meier, Wrinkle and Polinard (1999) compared how well districts actually performed to how well the statistical model predicted they should perform.

Bohte, Polinard, Meier and Wrinkle (1999) in the Texas Educational Excellence Project also followed the same approach to identify the top school districts in Texas for Latino students. This study created an education production function for Latino student passes in the TAAS examination, to compare *actual* performance with the *predicted* performance to identify districts that performed better than expected. Results in the education production function made it possible to identify school districts that excelled in teaching basic reading and mathematics skill to Latino students. The education production function controlled for per learner operating expenditure, teacher salaries, teacher experience, teacher certification, class size, gifted classes, percentage poverty background, percentage foreign born and learner attendance that might affect the learner performance of Latino students on the TAAS examination (Bohte, Polinard, Meier & Wrinkle, 1999).

Predictors of school performance

From the above-mentioned studies it appears that the explanatory variables to predict school performance can be summarised as follows:

- 1) School performance and Teacher Quality. Most researchers explored the extent to which differences in school performance can be explained by differences in teacher quality as often indicated by teacher qualifications, salary, age and experience. Findings in general suggested that teacher quality is a powerful predictor of learner achievement. Crouch and Mabogoane (1998b) suggested that average teacher qualification (REQV) is strongly and positively associated with matriculation pass rate. Their assumption is that “an increase in the education of educators equivalent to one REQV (roughly one year) appears to be associated with an increase on pass rates of about 16 or so points” (Crouch & Mabogoane, 1998b: 3).

These results are very similar to research conducted in Chile (Vegas, 2002) about student performance and teacher and school characteristics. He conducted regressions of average student test scores on sector, teacher, and school characteristics, controlling for average student socio-economic background. His findings indicated that percentage of teachers with university education was positively related to student outcomes (Vegas, 2002: 22).

2) Teacher availability as measured by learner-educator ratio. Prior to the dismantling of apartheid, huge discrepancies existed in class size among different race groups. The relationship between school performance and learner-educator ratio was regarded as a significant predictor of student outcomes. In the post apartheid period, with the equitable re-distribution of resources, research indicates that class size is not really a significant predictor of learner outcomes. In his analysis of project STAR (Student/Teacher Achievement Ratio) an experiment on class size in Tennessee, United States, Alan Krueger's (1999: 497) main conclusions were:

- On average, performance on standardised tests increased by four percentage points the first year students attend small classes;
- the test score of students in small classes expanded by about one percentage point per year in subsequent years;
- teacher aides and measured teacher characteristics had little effect;
- class size had a larger effect for minority students and those on free lunch (poor students).

3) Socio-economic background of the learner is highly correlated with learner achievement (Coleman, 1966). The key indicators measuring the socio-economic status of learners,

such as education of parents and family income, are not available. Researchers resorted to using proxies, such as school fees, race, or place of residence as control variables.

Ladd (1996:2) states that:

“Translating student performance into a fair measure of school performance requires taking into account wide variations in the socio-economic backgrounds of students in different schools. Equality of Educational opportunity, the landmark 1966 report on America’s schools by James Coleman, showed that socio-economic characteristics of students explain far and away the largest share of measured differences in educational achievement.”

Mendro et al. (1999:8) in A Case Study of the Dallas Public School School-Based Performance Award Program state that:

“An important consideration in the Dallas system is the fairness of the performance ranking system to schools. The philosophy of the program is that schools should not be held accountable for the characteristics that students bring to school, which may be related to or proxies for factors that influence learning. The Dallas system attempts to control for a variety of student and school-level characteristics, so that the remaining “gain” (or lack thereof) can more fairly be attributed to the school’s efforts. Controls at the student level include ethnic/language status, socio-economic status (measured directly by participation in free and reduced lunch, and imputed from block-level census data), and gender.”

State policymakers in South Carolina followed the same approach in their recognition and reward programs by taking into account the socioeconomic background of their students as measured by eligibility for free and reduced price lunches (Ladd, 1996:3).

The control variable that measured poverty, in the research on Black and Latino student improvements on the TAAS examination, in order to rank education districts, was the percentage of students that qualified for free or reduced-price meals in school lunch programs (Meier et al., 1999, Bohte et al., 1999).

Chile has one of the oldest and largest voucher programs in the world and a great deal of research has centred on the question of whether private voucher schools are more effective than public schools (Vegas, 2002). Vegas states that “under the voucher system, families can choose to send their children to free subsidised schools, either municipal or private, or they can choose fee-paying private schools if they can afford the tuition fees” (Vegas, 2002:4). In his method to explore the extent to which differences in student outcomes among schools can be explained by differences in the students they serve, in the characteristics of the teachers they employ, and by differences in their management strategies, Vegas conducted “weighted least squares (WLS) regressions of average student test scores on sector, teacher, and school characteristics, controlling for average student socioeconomic background” (Vegas, 2002: 8). His principal measure of average student socioeconomic background by school consisted of an “index of vulnerability”. The measures in the “index of vulnerability” include students’ weight, height, medical needs as well as the mothers’ education. Vegas is clear in his findings about the effects of students’ socioeconomic background on student outcomes when he explicitly stated that “controlling for socioeconomic background reduces the estimated

differences in achievement among public and all other schools substantially. In fact, after controlling for student socioeconomic background, the advantage of the private voucher/shared financing schools disappears". (Vegas, 2002: 22).

In their assessment of the effectiveness and efficiency of private schools in Chile's voucher system, McEwan and Carnoy (1999) employed the framework of education production functions, positing that academic achievement is produced by family and school inputs. They conducted regression analysis of the relationship between two dependent variables (Mathematics and Spanish) and independent variables, such as teacher characteristics, school resources, socio-economic status and family background and school location, for the years 1990 to 1996. The findings of the achievement regression results (McEwan & Carnoy, 1999:31) indicated that the signs and significance of the background variables corresponded to expectations and that the coefficients on index of school's socio-economic status were negative and significant, indicating that increasingly disadvantage schools had lower mean achievement. They concluded, supporting the findings of the above-mentioned researchers, that adjusting for parental background and location considerably altered the relative achievement differences among school types.

Bravo, Contreras and Sanhueza (1999) examined the results in terms of educational achievement, inequality and performance gap between private and public schools in Chile. Evidence from the regression analysis indicates that the private subsidized schools diminish their superiority with respect to the public schools when control variables are considered in the analysis, including socio-economic level, geographical characteristics and school characteristics.

Unexplained variation in school performance

The unavailability of key indicators, such as school management and learner ability, places a limitation on the results of the production function analysis. Van der Berg and Burger (2003) as well as Crouch and Mabogoane (1998a, 1998b) ascribed a great deal of the unexplained variation in school performance in South Africa to management. Crouch and Mabogoane (1998b: 2) stated: “The remnant, namely the degree of performance that cannot be explained [according to them up to 30%] away via resources and environment must have to do with the quality of management”. Van der Berg and Burger (2003:1) explore the same question by asking to what extent persistence in school performance inequalities was attributable to resource allocation and to what extent it was attributable to differences in efficiency of resource utilisation.

In similar studies in other countries, researchers have come to the same conclusions with regard to the impact of management on school performance. In his research on school choice, student performance and teacher and school characteristics, Vegas (2002) emphasised the importance of school management for student outcomes. His findings suggested that the way schools were managed was more strongly related to student outcomes than were observable teacher characteristics such as education, experience, and teachers’ high school grades. He was however, quick to acknowledge that future research should further our understanding of the characteristics of school management that are related to student learning (Vegas, 2002: 25).

Performance Measures

Standardised test scores are regarded as the best indicators to measure school performance and in South Africa the grade 12 results were the only available measure in the past. The analysis of assessment results should help to identify the effects of educational inputs on learner achievement and to put mechanisms in place to correct it. Researchers, however recognised the limitations in pass-fail measures, such as the matriculation results and many of them created a composite index that included other indicators to reflect a more comprehensive description of the schools performance. Van der Berg and Burger (2003: 21) used the Swedish points system, commonly used to calculate entry eligibility for South African Universities, to construct an overall and mathematical performance index. Crouch and Mabogoane (1998a) constructed a weighted index of absolute achievement and a relative index. By comparing the actual rating (absolute index) with the predicted rating (relative index) they could identify over-achievers and under-achiever schools relative to their resources.

Cowley and Easton (2004) have used a similar method to create a Report Card on British Columbia Secondary Schools. The underlying principle of their report card is that just as teachers combine test scores, homework, and class participation to rate a student, they have combined eight indicators to produce an overall school rating. The weightings used in their calculation were Average exam mark – 20%, Percentage of exams failed – 20%, School vs. exam mark difference – 10%, English 12 gender gap – 5%, Mat 12 gender gap – 5%, Exams taken per student – 20%, Graduation rate – 10%, and Composite dropout rate – 10% (Cowley & Easton, 2004: 53).

The Dallas Public School-Based Performance Award Program is based on a school effective index (Mendro et al., 1999).

“The Dallas School Effectiveness Index uses three basic types of measures: student test results, “school-wide” attendance, drop-out, and/or promotion rates, and at the high school level, participation rates in activities such as AP classes and college entrance tests” (Mendro et al., 1999:5). “For a K-4 school, the two non-test indicators, attendance and promotion rate, are given 30% of the weight, and the test-based indicators, 70%. For a high school, test scores get approximately 80% of the weight, with percentages of students taking SAT, PSAT, and AP about 16% and graduation rate 4%” (Mendro et al.,1999: 7).

Concluding remarks

The literature on education production methodology offers useful insights into the relationship between school performance and resource inputs, such as socio-economic background, school and teacher characteristics. It is important to note that literature findings indicate that teacher and school characteristics do affect school performance, but that a great deal of the unexplained variance remains important in determining learner achievement and school performance. The education production function estimates generally indicate that actual pass rates are a poor measure of school performance and that socio-economic variables are important in explaining student outcomes.

Findings of various researchers (Crouch & Mabogoane, 1998a; Crouch & Mabogoane, 1998b; Fiske & Ladd, 2003; Van der Berg & Burger, 2003) indicate that teacher qualification is positively related to school performance. Crouch and Mabogoane (1998b) showed that teacher qualifications are statistically significant. Their estimate suggested that a year of training is

associated with a 16 point increase in the pass rate. When examining the determinants of overall and mathematical performance in the Western Cape, Van der Berg & Burger (2003: 27) argued that their “findings are in line with the results of a comparable study by Crouch & Mabogoane (1998b), who found that matriculation results had a strong positive correlation with the qualifications of teachers”. When comparing schools in the Western Cape with the National data, the quality of teachers seem to have a greater effect in the Western Cape (Van der Berg & Burger, 2003: 18).

The literature review was used to inform this study of what the existing research says about education production functions in South Africa. The research review of the education production function literature in this study helped to identify what is already known, as was discussed in the comparative summary. The literature review indicated what was already investigated and identified the focus for further research. A benefit of the literature review was that it provided the context within which to interpret and report the findings of this study and to compare it with the results reported by these other studies. The literature review served as a precursor for the empirical education production function research that follows and set it within a conceptual and theoretical context or framework. There is an ongoing debate and critique of the production function studies (Subotnik & Walberg, 2006). Although this is an important debate, this issue falls outside the scope of the research for this thesis.

4.3.3.4 Research findings: Analysis of education production function Results based on input variables of schools in the Western Cape

The following discussion is based on the estimates of the production function for Western Cape schools. There are many ways that data analyses could affect sensemaking, here only the production function is considered.

4.3.3.4.1 Analysis of estimates

Equation 1 in Table 6 suggests that average qualifications of teachers (REQV), dropout rate per school (dropout rate), percentage of learners who repeated grade 11 (repeaters), poverty of the community (poverty) and the number of teachers per 1000 learners paid by the school (SGB teachers) significantly influence the 2002 grade 12 results. The R^2 or coefficient of determination is an indication of the proportion of variation in the aggregate grade 12 results (aggregate mark) between schools explained by the above-mentioned five independent variables. According to equation 1 these five independent variables account for 75% of the variance in the grade 12 results in the Western Cape. These estimates tell us a good deal about what affects grade 12 results in the Western Cape Education Department, although some of these variables could probably be seen as proxies for other factors, such as school management.

It is important to note that both the variables poverty and SGB teachers are included in the equation and they are still both statistically significant.¹² This means that one is not simply a proxy for the other, but each one has its independent effects. In that case, the effect of the SGB teachers could have three sources. The production function approach, like all others, has limitations in proving causality and identifying causal channels. From the analysis, it transpired that the positive effect of the SGB teachers could have the following implications:

It could be either a pure learner-educator ratio effect (the lower the learner-educator ratio, the better the performance), or it could indicate that the SGB teachers tend to be more productive than the teachers paid for by the state (state teachers), perhaps because they know that their (lack of) performance could more easily result in dismissal, or it could be another factor measuring socio-economic status. Schools that can afford to hire extra teachers get the desired return on investment that results in either lower student-teacher ratios or better quality teachers. Knowledge about the importance of these two factors with regard to SGB teachers in the determination of student outcomes remains tentative, and point to the limitations of quantitative analysis of large datasets of this nature.

¹² Tolerances were examined to determine whether collinearity was a problem. Tolerance statistics were all within acceptable ranges “Tolerance is a statistic used to determine how much the independent variables are linearly related to one another (multicollinear). The proportion of a variable's variance not accounted for by other independent variables in the equation. A variable with very low tolerance contributes little information to a model, and can cause computational problems. It is calculated as 1 minus R squared for an independent variable when it is predicted by the other independent variables already included in the analysis” (SPSS 9.0 for Windows, 1998).

When we added a dummy variable¹³ for ex-department in equation 2 (Table 6) the R² increased significantly and the coefficient was large and significant, showing that schools in the former Cape Education Department (CED) have a 181-mark advantage over schools that formerly belonged to other departments, even when holding other variables constant. In fact, the R² indicates that these variables in equation 2 account for 79% of variation in grade 12 results in the secondary schools in the Western Cape.

To verify this effect, we then used the CED as reference and included other former departments in equation 3. This time it shows that a school in the former House of Representatives (HOR) has a 181-mark disadvantage over schools in the former Cape Education Department and a school in the former Department of Education and Training (DET) has a disadvantage of 179 marks, that is, the difference between HOR and DET schools is minimal. Moreover, the regression findings suggest that CED schools outperform non-CED schools. The results of African and Coloured learners in ex-HOR and ex-DET schools continue to reflect the imbalances of the past. At this stage of the analysis it was not clear what were the factors within the CED schools that contributed to the better student performance. However, this is an investigation that is pursued later in this chapter.

¹³ “Dummy variables are a way of adding the values of a nominal or ordinal variable to a regression equation. Each value of the categorical independent except one is entered as a dichotomy (for example, CED = 1 if school is CED, otherwise 0; HOR = 1 if school is HOR, otherwise 0; etc.). One category must be left out to prevent perfect multicollinearity in the model. For instance, for the nominal variable ‘previous Education Department’ we may create a set of dummy variables called CED and DET, leaving out HOR. (Refer to the digital document on multiple regression”. [Online] Available url: <http://www2.chass.ncsu.edu/garson/pa765/regress.htm>)

Table 6: Multiple Regression Estimates

	MEAN	Equation1	Equation2	Equation3	Equation4	Equation5	Equation6	Equation7
REQV ¹⁴	4.18	137.66 (5.202)	102.88 (4.153)	102.44 (4.136)	58.17 (1.934)	102.55 (4.132)	101.01 (4.078)	46.66 (1.52)
Dropout rate	0.07	-238.71 (-3.323)	-193.93 (-2.923)	-186.81 (-2.799)	-187.31 (-2.846)	-194.02 (-2.921)	-186.26 (-2.802)	-174.14 (-2.83)
Repeaters	0.08	-477.51 (-5.202)	-357.61 (-4.165)	-359.56 (-4.054)	-350.98 (-4.122)	-359.65 (-4.175)	-340.82 (-3.935)	-324.27 (-3.775)
Poverty Index	0.30	-425.34 (-5.770)	-313.21 (-4.513)	-306.39 (-3.702)	-333.50 (-4.816)	-302.54 (-4.060)	-284.20 (-3.923)	-288.74 (-3.646)
SGB teachers	4.79	13.46 (8.662)	5.60 (3.173)	5.86 (3.415)	4.54 (2.524)	5.49 (3.076)	12.48 (2.356)	15.01 (2.801)
CED	0.36		181.22 (7.602)		-357.90 (-2.311)	196.96 (4.256)	202.78 (7.118)	-421.85 (-2.590)
HOR	0.47			-181.50 (-5.139)				
DET	0.14			-179.90 (-7.900)				
Interaction of REQV & CED	1.64				125.99 (2.571)			149.97 (2.914)
Interaction of SGB teachers & CED	1.72						-7.54 (-1.377)	-11.73 (-2.066)
Interaction of Repeaters & CED	0.03							
Interaction Poverty Index & CED	0.11					-70.63 (-0.397)		-23.05 (-0.124)
R ²		0.75	0.79	0.79	0.79	0.79	0.79	0.79
Adjusted R ²		0.74	0.78	0.79	0.79	0.78	0.78	0.79
N		319	319	319	319	319	319	319
Constant		573.09	647.17	825.14	835.03	645.11	635.84	852.50
MEAN OF AGGREGATE (Dependent Variable)	1021.93							
Predicted Aggregate		1030.44	1033.10	1037.18	1036.45	1030.24	1052.30	1066.03

¹⁴ See appendix B

4.3.3.4.2 Interpretation of the interaction between the dummy variable CED and the average teacher qualification (equation 4)

Interaction effects

To determine if the input variables have a significantly different impact on the grade 12 results within the former CED schools, the approach was to use interaction terms as a technique in multiple regression to determine the joint effect of two variables (e.g. CED and REQV) on the grade 12 results. It can be written as:

$$Y = a + b_1X_1 + b_2X_2 + b_3 X_1 X_2$$

The multiplicative term, $X_1 \text{ times } X_2$, is a new variable representing the contribution of X_1 on X_2 . The slope, b_3 , aims to measure the *interaction* effect (Jaccard, 1990:5). The following explanation is largely based on the interaction interpretation of Allison (1999:169-170). Consider the results in Table 6, equation 4. Equation 4 shows the regression of aggregate marks that includes years of teacher qualifications, a dummy for CED (CED = 1, non-CED = 0), and the product term of CED and REQV. The product term is statistically significant, so we may conclude that there is an interaction between CED and REQV. The main effect of REQV (58 marks) is the effect of teacher qualification when CED = 0, that is when the school is not CED. So among those schools that are not CED, each additional year of qualification produces an increase of 58 marks in the aggregate. The coefficient for the product term (125) is the additional effect of qualifications when the school is CED, so the effect of qualifications for CED schools is $58.173 + 125.986 = 184.159$: Each additional year of teacher training brings 184 more aggregate marks, thus considerably more than in other schools.

We can also interpret the interaction with respect to how the effect of CED varies with teacher qualifications (refer to the calculation of REQV in appendix B). The coefficient for CED (-357) indicates that among schools with teachers with no qualifications, those who are CED score about 357 marks less than those who are non-CED. Thus when one adds a dummy variable for the CED, and an interaction term for the CED times the qualifications of teachers (REQV), then the coefficient on the CED turns negative. It means that ex-CED schools do worse than average if one does not take into account how they use qualified teachers.

The predicted average for any of these equations is still at about the actual average of the aggregate (see Table 6). That is, multiplying the average of the predictor variables with the coefficients plus the intercept, gives back approximately the average aggregate.

Simulated scenarios of the interaction of CED and the average teacher qualification (REQV)

The following equations (Table 7) based on equation 4 (Table 6) further emphasise and explain the interaction between the dummy variable CED and REQV.

The following formula was used to determine the predicted aggregate for each of the equations in Table 7:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6 X_1 X_2$$

where Y is predicted aggregate, a is the intercept, X_1 is average qualification of teachers, X_2 is dropout rate, X_3 is the number of learners who repeated grade 11, X_4 is the poverty index of the community, X_5 is dummy for ex CED, and X_1X_2 is an interaction term for CED and REQV.

The actual average 0.36 in Table 7 (column 1) means that 36% of the schools in the sample belonged to CED. Note that the predicted aggregate in Table 7 (column 1) is approximately the same as the actual average aggregate (see Table 6).

Column 2 (Table 7) takes the argument one step further. The CED value was set to 1, turning this school into a CED school, but retaining all other variables at their average values. Note the predicted aggregate is now 1127. In column 3 the average teacher qualification was arbitrarily set to 1 below the actual mean (see Table 6). The predicted aggregate is now 943, meaning that being CED by itself is not what brings the improvement. It is being CED and having well qualified teachers that make the big difference. See how low the predicted aggregate is for a CED school with poorly qualified teachers (column 3).

In column 4 the school was turned into a non-CED school by changing the CED variable to a zero. Compare how low the predicted average (900) is for a non-CED school with poorly qualified teachers. In column 5 the school is modified to simulate a non-CED school with well qualified teachers. Being non-CED but having very well qualified teachers (way above the actual average) could compensate for being non-CED. The average was still approximately the same as the actual aggregate (1016).

The last scenario, column 6 (Table 7) simulates a CED school with very well qualified teachers. By putting a 1 in for the ex-CED the school was turned into a CED school. Being

CED and having really well qualified teachers make a statistical significant difference, because there appears to be something about how CED uses well-qualified teachers¹⁵.

When the three key variables, namely average teacher qualification, CED, and teacher qualification interacted with CED were applied, the following inferences were made:

- It appears that the reason why CED schools do better as shown in equation 2 (Table 6), lies in how much better highly qualified teachers are used in the ex-CED schools.
- Similarly, it indicates that while the average qualification of teachers seem to make a difference of 58 marks (row one, equation 4 in Table 6), the average qualification of teachers in the ex-CED schools makes even more of a difference (125 marks). Furthermore, the findings seem to indicate that the greater effectiveness with which highly qualified teachers are utilised in the ex-CED schools accounts for the better performance in those schools. This could be attributed to:
 - ex-CED schools knowing how to make better use of their teachers or,
 - the fact that teachers in CED schools have a different type of qualification, or studied at better resourced teacher colleges or universities, than other teachers.

Again, issues such as teachers who have the same numerical REQV but appear to have a better quality training, or the assumption that CED schools better utilise well-qualified teachers, are areas for future research, where the present dataset needs to be supplemented by qualitative research.

- Equation 4 (Table 6) revealed that equations 2 and 3 (Table 6) could be misleading, because equation 2 (Table 6) infers that being CED is "good" in some vague way, and

¹⁵ Or it may be that the REQV itself is not a sufficient indicator of quality, that it is the actual type of college or university the teacher went to, not just the years, which makes a difference. This research cannot answer that question.

equation 3 does not dispel that idea. But according to equation 4 (Table 6) it appears that the better performance of the CED schools could be ascribed to:

- teachers who have the same numerical REQV, but who appear to have a better training, or
- because CED schools better utilise well-qualified teachers.

Crouch and Mabogoane (1998b) argued along similar lines with regard to teacher qualification when they gave a description of the analysis of their production function. They emphatically underlined the fact that it “is difficult to fight the negative effect of poverty with mere resources, unless they are well-managed, and unless one pays more attention to the qualitative nature of the resources (even as rudely measured as REQV), but more importantly if one pays attention to whatever qualitative behaviours are associated with educators with high REQVs” (Crouch & Mabogoane, 1998b: 13).

The interpretation of equation 6 in Table 6 is the opposite of equation 4. The main effect of SGB teachers (12) is the effect when the school is $CED = 0$, that is when the school is not CED. So among those schools that are not CED, each additional SGB teacher per 100 produces an increase of 12 marks in the aggregate. The coefficient for the product term (-7) is the additional effect of SGB teachers when the school is CED, so the effect of SGB teachers for CED schools is $12 + (-7) = 5$: Each additional SGB teacher brings 5 more aggregate marks. It reveals that in the CED schools having extra SGB teachers does not make a bigger difference than in the other schools. The interpretation in equation 5 is similar. Being poor and CED does not do anything special compared to either being poor or being CED. That is, being CED and being poor are factors that have independent effects.

Table 7: Simulated scenarios of the interaction of REQV and CED

	1	2	3	4	5	6	
Predictors of Performance	Actual Average	Simulation of a CED school	Simulation of a CED school with very low teacher qualifications	Simulation of a non-CED school with very low teacher qualifications	Simulation of a non-CED school with high teacher qualifications	Simulation of a CED school with high teacher qualifications	Coefficients
REQV	4.18	4.18	3.18	3.18	5.18	5.18	58.17
Dropout rate	0.07	0.07	0.07	0.07	0.07	0.07	-187.31
Repeaters	0.08	0.08	0.08	0.08	0.08	0.08	-350.98
Poverty Index	0.3	0.3	0.3	0.3	0.3	0.3	-333.5
SGB teachers	4.79	4.79	4.79	4.79	4.79	4.79	4.54
CED	0.36	1	1	0	0	1	-357.90
HOR	0.47	0.47	0.47	0.47	0.47	0.47	
DET	0.14	0.14	0.14	0.14	0.14	0.14	
Interaction of REQV & CED	5.10	14.18	13.18	0	0	15.18	125.99
Interaction of SGB teachers & CED	1.72	1.72	1.72	1.72	1.72	1.72	
Interaction of Repeaters & CED	0.03	0.03	0.03	0.03	0.03	0.03	
Interaction of Poverty Index & CED	0.11	0.11	0.11	0.11	0.11	0.11	
Predicted Average	1019.41	1127.42	943.26	900.49	1016.83	1311.58	
Constant							835.03

4.3.3.5 Application of production function results : A sensemaking perspective

The approach in this section is to compare actual scores with predicted scores. The education production function results are used to predict how well each school should do. The top 50 schools, ranked according to the best residual, of the 322 in the Western Cape are shown in the Table 8. The first column is the ranking of the school. The second column is the actual average aggregate. The third column is the predicted average aggregate and the third column is the residual.

$$\text{Residual} = \text{actual aggregate mark} - \text{predicted aggregate mark}$$

The residual column in the table has been used to produce a ranking of schools (column 1). Notice that the higher the residual the better the performance of the school. Schools have been divided into 5 quintiles (refer to column 5), according to the following method. The schools were sorted according to the poverty index of the community in descending order. The schools in this list of schools with the first 20% of the learners constitute quintile A, which includes the poorest schools. Remember a poverty index of 1 is poor and 0 is rich. The range in the list with the next 20% of learners forms quintile B; the range with the last 20%, also the lowest poverty index, is quintile E, which includes the most affluent schools.

Table 8: Top over-achieving schools

RANKING	ACTUAL AGGREGATE	PREDICTED AGGREGATE	RESIDUAL	QUINTILE
1	1601	872	729	D
2	1386	810	575	A
3	1361	1042	318	E
4	1590	1348	242	E
5	1004	773	231	A
6	1426	1200	226	E
7	1159	939	220	D
8	1499	1315	184	D
9	957	775	182	A
10	1342	1162	180	D
11	1027	848	179	A
12	1477	1305	172	D
13	1494	1322	172	E
14	1072	901	171	D
15	1371	1206	164	D
16	1535	1371	164	E
17	1552	1390	162	E
18	1404	1243	162	E
19	1494	1332	162	E
20	951	800	151	D
21	1111	962	149	D
22	1462	1318	144	C
23	1325	1182	143	D
24	1409	1267	142	E
25	1138	997	141	D
26	1493	1352	141	E
27	818	678	140	A
28	1108	970	139	D
29	1398	1265	134	E
30	1105	979	126	D
31	1004	879	125	C
32	1005	885	120	A
33	1509	1391	119	E
34	961	845	117	B
35	997	886	111	D
36	955	843	111	A
37	1344	1233	111	E
38	1381	1271	110	E
39	1402	1294	108	E
40	959	851	108	B
41	985	877	108	C
42	1272	1167	104	B
43	962	858	104	D
44	875	772	104	A
45	952	851	102	B
46	1171	1071	100	E
47	1340	1242	98	C
48	1392	1297	96	E
49	992	899	93	A
50	1363	1271	92	D

Table 9 displays the distribution of the top 50 schools according to poverty quintiles. A sensemaking view on the performance of schools suggests that education practitioners and policy makers have to make sense of what is happening in the education environment.

Table 9: Distribution of top 50 schools according to poverty quintiles

<i>QUINTILE</i>	<i>PERCENTAGE OF SCHOOLS</i>
<i>A</i>	<i>18%</i>
<i>B</i>	<i>8%</i>
<i>C</i>	<i>8%</i>
<i>D</i>	<i>32%</i>
<i>E</i>	<i>34%</i>

Table 9 indicates that the 50 top ranking schools include 18% of the poorest schools (quintile A). However, if we rank the schools on actual test scores then there is no poor school within the top 50 schools.

Figure 10 displays the relationship between the actual aggregate and the predicted aggregate marks of grade 12, depicting over-achieving and under-achieving schools.

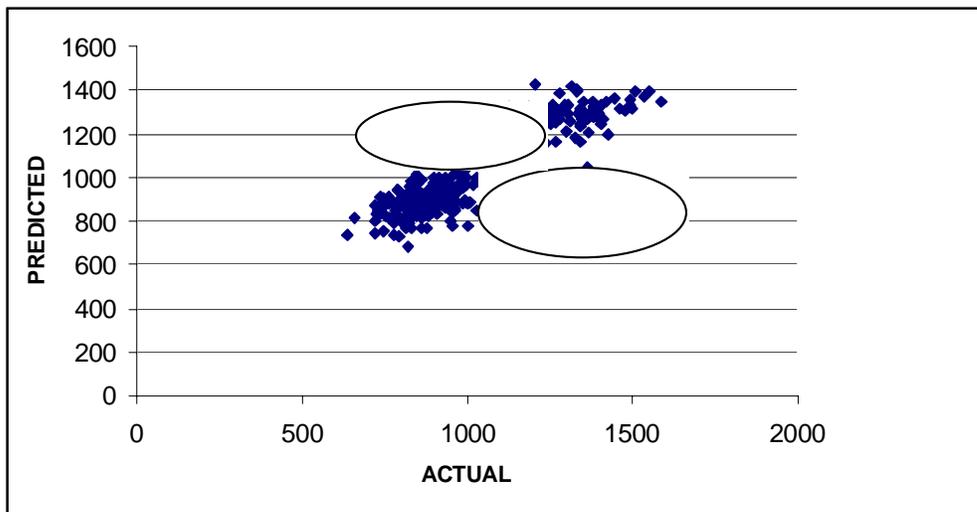


Figure 10: Over-achiever and under-achiever schools

The top circle in the Figure 10 indicates the over-achiever schools and the bottom circle the under-achiever schools. The under-achiever schools tend to be some of the richest schools. In a table with the 50 lowest ranking schools, according to the constructed index, 32% are of the richest quintile. A ranking of schools on actual test score indicates virtually no rich schools in the bottom 50. This indicates the importance of this kind of data analysis for making sense of school performance.

A further benefit about the ranking is that it is now possible to compare homogenous groups such as schools within the same suburbs or schools with the same poverty index. Table 10 shows schools within the same suburb.

Table 10: Comparison of schools within the same geographical area

RANKING	ACTUAL AGGREGATE	PREDICTED AGGREGATE	RESIDUAL
1	1004	773	231
2	951	800	151
3	955	843	111
4	832	769	64
5	881	825	56
6	880	832	48
7	885	843	42
8	859	818	41
9	845	820	25
10	855	836	19
11	795	812	-17
12	725	834	-109

A school can now evaluate its performance based on the position within the list. All it needs to know is what its relative position is without knowing the names of the other schools. However, one needs to be careful not to over-emphasize this analysis in the case of very small schools. A very small school's performance in any given year can be highly affected by random factors that should not be seen as an increase or decrease in value-added.

4.4 Output and results : Discussion and concluding remarks

Sensemaking is an output of the production function model and serves as a reasonable guide for policy action. For further research findings and discussion on the production function results refer to Chapter six. This chapter attempted to locate the data analysis with the context of the EMIS environment and explained the structure of it using the systems concept of input-transformation-output. Although this chapter concludes the first approach regarding the quantitative methods the Data Analysis and Reporting phase continues in the next chapter. The next chapter is therefore not a continuation of this chapter, but rather introduces the next approach of the use of quantitative methods.

CHAPTER FIVE

AN ANALYSIS OF LEARNER ENROLMENT AND THE FLOW OF LEARNERS THROUGH THE EDUCATION SYSTEM: ASSESSMENT OF INTERNAL EFFICIENCY IN EDUCATION

5.1. Introduction

This chapter describes the effects of dropouts and repeaters on the internal efficiency of the education system in the Western Cape. The basic principle of the method facilitates the estimation of the flow of a cohort of learners through a particular system in order to derive indicators of efficiency. The two techniques that are explored in this chapter to determine these effects are cohort analysis and educational attainment. For each one of these techniques a literature review and an empirical analysis are outlined.

The first technique, educational attainment, uses the 2001 census data. Using this technique the impact of socio-economic status on the educational enrolment and attainment in the Western Cape is explored. The second method, cohort analysis, assesses the impact of repetition and dropout measures on the internal efficiency of the education system in the Western Cape and uses school data of the years 2001 and 2002. The hypothesis of this method and how to employ it is explained in a literature review and the empirical application.

It is often not so easy and straightforward to determine how effective and efficient an education system is. Where there is no measurement of efficiency there is no idea if

applications of resources are appropriate. In a South African context this can be ill afforded given the budget constraints and the limited resources. Further, one needs to explore what feedback and control mechanisms are operational that evaluate how well the system is functioning given all the inputs of fiscal, material and human resources. All these efficiency issues are key to determine where the money should be spent to improve the system, but also to reach the areas where the needs are the highest.

Generally, economists define efficiency as the relationship between input and output (outcomes). Given this definition, the flow of learners through an education system could be used as such a quantifier to measure efficiency. The structure of the chapter will be outlined according to the systems concept of input-transformation-output through the application of the learner flow through models.

5.2. Different data sets as input for the application of flow through models

The input to this phase is the quality data that was produced in the data processing phase (Chapter 3). As stated before, ISD is a continuous process where the output of one phase becomes the input to another process. Thus the output (the quality data) phase 3 will be used as input for the data analysis in phase 4, indicating that ISD is a continual process or “ongoing” as described from a sensemaking perspective. However, the application of flow models to measure internal efficiency requires more data sets than those available in EMIS. Other information systems that exist are important data sources for analyses purposes. The incorporation of additional data sets to augment the existing EMIS information constitutes an exercise of bricolage. It is an appropriate point in the chapter to mention these existing

information systems from where these data sets come that will be used as input in this chapter.

These information systems offer the following data sets:

Education Management Information System

A comprehensive discussion regarding this system is already outlined in the previous chapter and therefore not repeated here. However, it is important to note that the datasets used in this chapter are mainly derived from this system. These distinct datasets will be described in the section coupled with the appropriate data analyses practices.

Census 2001

The census data sets of 2001 have been used in this research to calculate a poverty index for suburbs and to calculate educational attainment data sets, such as highest educational level (grade) obtained per person and children ages 15 to 19 within a suburb.

5.3. Transformation processing : Flow models to analyse data

Through the data analysis and the application of different quantitative techniques the study attempted to transform the data sets in such a manner that would make sense to decision makers and education management.

5.3.1. The effects of dropout and repeaters on internal efficiency of the education system

There is a vast literature on educational attainment in the world (Abagi & Odipo, 1997; Ahuja & Filmer, 1996; Arends, 2006; Byrnes, 1989; Cassidy, 2005; Cuandra, 1991; Cuadra & Fredriksen, 1992; Filmer, 1999; Filmer & Pritchett, 1998a, 1998b; Fredriksen, 1991; Grissom & Shepard, 1989; Holmes, 1989; UNESCO, 1998, 2004a, 2004b; Wako, n.d.,1995; Wolff, Schiefelbein & Valenzuela, 1994). The focus of this literature overview is on empirical studies assessing the effects of dropout and repeaters on internal efficiency.

5.3.1.1. Cohort Analysis

The basic guidelines of cohort analysis models for assessing educational internal efficiency have been developed by UNESCO (1998, 2004a, 2004b) and since been applied in various studies (Arends, 2006; Cassidy, 2005; Cuandra, 1991; Fredriksen, 1991). UNESCO (2004a) followed the essence of the system theory in describing how to measure internal educational efficiency. An education system is efficient if maximum output is obtained from a given input. UNESCO (2004a) quantified educational inputs as the buildings, teachers, teaching-materials, etc. which may be aggregated financially in terms of expenditure per learner-year. Educational output they defined as the number of learners who completed the cycle (graduates). UNESCO (2004a) indicated that such a definition is rather restricted since the dropouts have acquired some skills. They acknowledge that educational internal efficiency has limitations that must be recognised and respected. They pointed out that these limitations are related to the weaknesses of some of the key concepts used to define efficiency in education, such as:

- Inputs: The pupil-year is a non-monetary measure of input, which fails to take into account the concepts and findings of educational cost analysis.
 - Outputs: The fact that the output is equated with number of graduates makes for a very narrow view of the education process. Identifying grade repetition as wasteful is not entirely true.
 - Process: The concept of internal efficiency in education is applicable only to those educational processes that follow the age/grade-pattern of conventional formal schooling.
- UNESCO (1998, 2004b), in their document of technical guidelines, outlines the techniques for measuring internal efficiency and ‘wastage’ in education.

The three methods they emphasise when analysing the internal efficiency of an education system by looking at how a cohort of learners moves through a cycle of education are:

- **The true cohort method:** According to UNESCO (1998:45) the ideal way to obtain a precise assessment of educational wastage is through the use of the true cohort method. The true cohort method is a longitudinal study that monitors the progress of a selected cohort of learners through the educational cycle. This method however requires good and reliable school-records with information on individual learners over a number of different years. Data of this nature is not often available in developing countries. Thonstad (1980: 13) is more explicit when he described this method by assigning an identification number to each learner. In that case it would be possible to follow the path of each learner through the school system. Although such individualised information (Cassidy, 2005; Cuadra, 1991) is costly to introduce and to maintain it is much more possible in this technological age and necessary in the information society. It is a powerful method to monitor enrolment that is a main cost driver in

education for staff establishment, school funding and building of schools. The essence of the true cohort method, which is in fact a learner tracking system, requires that

- Each learner has a unique number. In a South African context it could mean the identity number.
- Each learner is registered at a central point.
- Each learner is verified against the national registration system based on a unique number.
- Transfer of a learner between schools (internal and external) is monitored centrally.
- Each learner promotion is noted via the central system.
- Each learner is captured centrally at the beginning of each year to build up a learner history. Optical character recognition, an optical mark reader and web-based system for on-line data capturing are technological innovations appropriate for such a function.

Thornstad (1981:13) suggested a simpler “school cohort coding system” where each learner is given a code number showing: the year of school entry, the age at school entry and the grade of school entry. In such a system, according to Thonstad (1980:13), all learners of the same age entering grade 1 in a given year would have the same code. Using such a system it is possible to trace a cohort of learners and to analyse the flow patterns operating in the school system as a whole.

- **The apparent cohort method:** This method is applied when there is no data on repeaters and is appropriate for countries applying automatic promotion.
- **The reconstructed cohort method:** The UNESCO (1998:45, 2004) technical guidelines also set out the use of the reconstructed cohort method, using the main flow-rates of promotion, repetition and dropout. The reconstruction of the learner-cohort flow makes it possible to derive the indicators of internal efficiency. These indicators (survival rates, coefficient of efficiency and the years input per graduate) are discussed in the technical

guidelines. This study uses this reconstructed cohort method to assess the impact of repetition and dropout and to measure the internal efficiency of the education system in the Western Cape. The hypothesis of this method and how to employ it will be explained later in the chapter.

In their exposition of the problem of school wastage, UNESCO (1998:11) defined wastage as learners “who do not complete their schooling in the prescribed number of years either because they dropout of school entirely or because they repeat one or more grades”. It is in terms of this statement that UNESCO (1998) investigated the impact of wastage on the internal efficiency of an education system. From a system theory point of view, the report stated that an education system is considered efficient if for a given input of resources (human, financial and material), it maximises the desired output, both in quality and quantity. In analysing and measuring school wastage, the report outlines the reconstructed cohort method, using the learner-years as input to obtain all the indicators of internal efficiency. However, the focus of the report is on the dropouts and repeaters.

In terms of dropping out of schools the UNESCO (1998: 14) report stated that “an obvious and blatant form of wastage involves learners who start school but dropout before they reach a level of sustainable literacy and numeracy”. Analysis of the dropout rates showed that it is virtually non-existent in developed countries, but it is a major problem in less developed countries, especially in the early grades. In fact the report indicated that much of the dropping out of school in these developing countries occurs between the first and second grades. It further pointed out that boys persist in school at slightly higher rates than girls do in most developing countries. A particular concern that is addressed in this report is what happened to

these dropouts. The result of the analysis of school wastage indicated that rural schools often have higher dropout rates than urban schools. The UNESCO report (1998:15) referred to an interesting tracer study in China that showed that most rural school dropouts worked on farms compared to urban dropouts who were in fulltime or part-time employment.

Findings of the UNESCO report (1998:18) suggested that repetition, like dropouts, tends to be more prevalent in the first and final grades of the primary school cycle. Furthermore, the UNESCO report (1998:16) emphasised that, contrary to popular beliefs, repeating a grade does not help students gain ground academically and has a negative impact on social adjustment and self-esteem. These views concur with the findings of Holmes (1989), Byrnes (1989), and Grissom and Shepard (1989) that repeating a grade does not necessarily improve a learner's performance.

Thonstad (1980) followed a similar approach as UNESCO (1998, 2004a, 2004b) when he compiled a manual for analysing and projecting school enrolment in developing countries. The main part of the manual is based on the reconstructed cohort method or, as Thonstad referred to it, the Grade Transition Model (GTM). The analytical exposition of the grade transition model outlines the reasons why it is possible to approximate the school history of the learners entering primary education in a given year. Thonstad illustrated the method, using transition rates (dropout, promotion and repetition) between school years 1975 and 1976 for Upper Volta. Through the reconstructed method it is then possible to calculate the indicators of efficiency. These indicators of internal efficiency will be dealt with in the empirical application of the reconstructed cohort method and therefore are not repeated here.

In order to improve the quality of education, Wako (1995) compiled a manual for statisticians and planners to use the reconstructed cohort method to measure internal efficiency in the education system of Ethiopia. The major indicators that can be used to reflect the internal efficiency of an education system have been described in detail in this manual. The calculation of key rates to analyse the flow of learners through the system, namely promotion, repetition and dropout rates, are outlined. Through these rates (Wako refers to them as transition rates), the reconstructed cohort analysis is possible.

Wako (1995:26) also defined the term efficiency the same way economists apply it “as the optimal relationship between inputs and outputs”. An efficient activity therefore is when optimum output is obtained from a given input. Wako (1995: 26) used the concept of learner-year as a measurement for input. How to calculate the learner-years using the reconstructed cohort method is explained in detail in the manual. The cost of one learner-year includes all the resources (teachers, textbooks, classrooms, equipment, etc.) spent to keep one learner in school for one year. If a learner therefore repeats a grade, he is getting one year’s worth of education, but consuming two years’ worth of expenditure. In the reconstructed cohort analysis of efficiency, repeaters and dropouts represent wastage. Total cost can be estimated by multiplying the learner-years by cost per learner-year.

Wako (1995-35-47) divided the major indicators that can be derived from the reconstructed cohort analysis into two categories, namely indicators of retention and indicators of internal efficiency in education. Wako (1995:35-43) demonstrated how to calculate the following indicators of retention:

- a) Survival by grade: This indicator is used to show the extent to which the school system can retain learners, with or without repetition, and it measures the magnitude of dropouts. It is important to note that survival by grade measures the impact of repetition and dropout on internal efficiency.
- b) Dropouts by grades
- c) Graduation rate from final year
- d) Enrolment by years of study
- e) Dropouts by years of study
- f) Graduates by years of study
- g) Average duration of study for graduates
- h) Average duration of study for dropouts

The following indicators of internal efficiency have been included in the manual by Wako (1995: 43-47):

- a) Input-output ratio: This is a value that indicates the average learner-years expended to produce one graduate. Wako (1995: 45) derived an actual input-output ratio from the reconstructed cohort model as 1.43, indicating that graduates in the Ethiopian education system are being ‘produced’ at a cost 43% higher than ideal.
- b) Coefficient of efficiency: This indicator is the reciprocal of the input-output ratio, for example in Ethiopia it is $1/1.43$ or 69.7%, meaning that there is 30.3% wastage due to dropouts and repeaters (remember 100% indicates an optimal level of internal efficiency).
- c) Proportion of total wastage spent on repetition compared to dropouts.

Wako concluded by stating that the dropout and repetition rates in the education system in Ethiopia are the highest in grade 1. This finding replicates findings in other empirical studies

(Filmer & Pritchett, 1998a; UNESCO, 1998) about the effect of dropouts and repeaters on internal efficiency.

Wako (n.d.) applied the reconstructed cohort method on an eight-year cycle of primary grades to determine the level of internal efficiency in the entire primary education system in Ethiopia. He investigated the reasons behind the high dropout rates and repetition rates in the first half of primary grades in Ethiopia. The research enabled him to establish that the effects of repeaters and dropouts from grade one are the major causes of internal inefficiency in primary grades. The findings of Wako indicated that learners drop out of grade one for reasons such as no exposure to any form of pre-education, lack of teaching staff, insufficient textbooks or textbooks being above the standard of grade one. Other factors such as poor health, especially malaria and typhoid, are causes of absence from school. Other causes include hunger, distance from school, condition of classrooms, over-age, especially in rural areas.

Wako (n.d.:2) used the cohort analysis on the first four years of primary education in Ethiopia and in doing so was able to isolate the problems in the lower grades from those in upper grades. The separate analysis made it possible to see exactly where in the education system the problem occurred. Wako quoted a few problems with probable accompanying causes:

The coefficient of efficiency in primary grades (1-8) was 45%, which meant 55% of the education system resources in Ethiopia were wasted through dropouts and repetition and the situation was worse for girls than boys (Wako, n.d.:2). Higher rates of repetition and dropouts existed in grade 1. Again the problem was more serious for girls than boys.

Wako (n.d.: 3) attempted to show the extent to which grade 1 dropout and repetition added to the internal inefficiency of the first four grades and that both repetition and dropout affected the rates of internal efficiency in the first four grades of primary education in Ethiopia. He applied the reconstructed cohort method to obtain different kind of indicators for internal efficiency. Wako (n.d.:5), in keeping with a systems theoretical approach, defined the educational input as the number of learner-years spent in the cycle. He asserted that the quantity of educational inputs (teachers, classrooms, equipment, textbooks, etc.) rose as the number of students increased and the years passed by. Education output for Wako was the number of learners who successfully completed a given cycle.

Wako (n.d.:7) presented the results of the cohort analysis indicators by comparing a situation in grade 1 with repetition and without repetition. He explained these core indicators of internal efficiency with specific quantitative examples from the education system in Ethiopia:

- a) Survival by grade: This is an indicator of the number of learners remaining in the system and indirectly indicates an estimate of those learners that have left the system prematurely. The result of the cohort analysis showed that the level of survival in Ethiopia would increase from 71% to 96% if all the learners were promoted to the second grade. The survival rate would improve whenever further action like promoting learners in the following two grades (2 & 3) materialised. He even was of the opinion that if further action was taken with regard to dropouts, the Ethiopian education system could approach 100% internal efficiency at this level.
- b) Graduates from grade 4: Those who successfully completed the final grade of the first half of primary education would increase from 54% to 74% if there was no repetition in grade 1.

- c) Average duration of study for the cohort: This indicator showed the number of years learners spent in the school system.
- d) Coefficient of efficiency: Wako maintained that if Ethiopia could manage to keep at least 50% of early dropouts in the system, the coefficient of efficiency could improve significantly. For example the coefficient of efficiency in Ethiopia could increase from 69.7% to 80.3 % if there were no repeaters in grade one, but automatic promotion to the next grade.

Wako (n.d.: 8) concluded his argument with the following reasons for the repetition or dropout in Ethiopia:

- a) Lack of kindergarten schools: The majority of repeaters and dropout were in grade one, because learners came to school without exposure to any kind of formal training;
- b) contents of textbooks, especially in grades 1, 4 and 5 were regarded as too difficult;
- c) lack of textbooks;
- d) shortage of teaching staff;
- e) lack of qualified teachers;
- f) self-support: learners attended school half a day and then worked the other half to earn some money to survive;
- g) over aged learners; health problems;
- h) famine;
- i) family size;
- j) distance to school and
- k) parental attitude.

The result from the reconstructed cohort model as applied by Wako (n.d.:13) suggested that nearly 78% of the wastage in the first half of the primary cycle was caused by dropouts and repeaters in grade 1. His study also revealed that the wastage in the first half of primary cycle was more due to dropouts than repeaters.

Ernesto Cuadra and Birger Frederiksen (1992: 1) in measuring the effects of repetition and dropout on internal efficiency used the reconstructed cohort method. The supporting argument for their research is that reducing the high levels of repetition and dropout in most developing countries offers considerable scope for more efficient use of available resources (Cuadra & Frederiksen 1992: 1). Repetition according to them involves an element of waste since repeaters stay in school longer than the normal duration of the cycle, thus increasing cost per graduate and leading either to more crowded classrooms or to a reduction in the intake capacity of the corresponding grades. Their specific conclusion is that repetition tends to increase the probability of dropout. This argument corresponds to the findings by Grissom and Shepard (1989:34) “that holding children back increases rather than decreases their risk for dropping out of school”. Byrnes (1989: 13) is of a similar opinion, “that repeating a grade does not improve a student’s subsequent performance”. Cuadra and Frederiksen (1992: 1) asserted that policies to reduce repetition and dropout generally relied on quality-improvement measures as their main instrument for achieving such reductions. Quality-improvement measures, according to Cuadra and Frederiksen (1992), included in-service teacher training, and textbooks. These measures were likely to be more cost-effective than reduction in class size below 40 or the introduction of computers in the classroom.

Cuadra and Frederiksen (1992) used the reconstructed cohort method to assess the impact of repetition and dropout on efficiency. This method was used to estimate the progress of a cohort of learners who started school the same year through a cycle of education. Cuadra and Frederiksen (1992: 3) showed that through the reconstructed cohort method the efficiency of an education system could be determined quantitatively, for example a 100% efficient system would be one where there was no repetition or dropout. The input/output ratio was the indicator used to gauge how close to the 100% efficiency mark an educational system was.

Cuadra and Frederiksen (1992) discussed the input and output indicators when using the reconstructed cohort method to measure the internal efficiency of the education system. The output indicator is the number of learners graduating from the cycle. The number of learner-years as used by the cohort determines the input indicator. They defined learner-year in terms of the amount of resources consumed by the cohort as determined by the number of learners enrolled in each grade of the system (cycle) every year and the supplies and services (i.e. teachers, curriculum, learning materials, buildings, etc.) used to educate these learners.

Cuadra and Frederiksen (1992: 3) referred to two schools of thought pertaining to the effectiveness of repeating a grade. Firstly, there are those who believe that repetition is an effective measure of helping students who are not judged ready for promotion to the next grade. Secondly, there are those who argue that repetition has potentially harmful effects on the self-esteem and attitudes toward schooling and that holding learners back increases the likelihood of dropping out of school.

The same argument continued about dropouts with the question: “Is dropout a total waste?” For them there is no clear-cut answer to this question. Cuadra and Frederiksen (1992: 4), however, concluded that from a pure quantitative perspective repetition and dropout affect the internal efficiency of the system, because they influence both the amount of educational inputs used by the system and the total level of outputs generated by it. According to Cuadra and Frederiksen (1992) repetition affects the inputs used by the educational system because the longer a learner is enrolled beyond the minimum time required, the more teacher time, classroom space, textbooks and other teaching materials and educational services will be needed. Dropouts have a similar effect on educational outputs, because dropouts affect the number of learners who graduates.

Abagi and Odipo (1997) are of the opinion that the conceptualisation of the term school or education efficiency in a developing country, such as Kenya, should take a “process perspective” as opposed to an “outcome perspective”. They therefore used system theory as conceptual framework to produce a holistic operation model of efficiency in Figure 11.

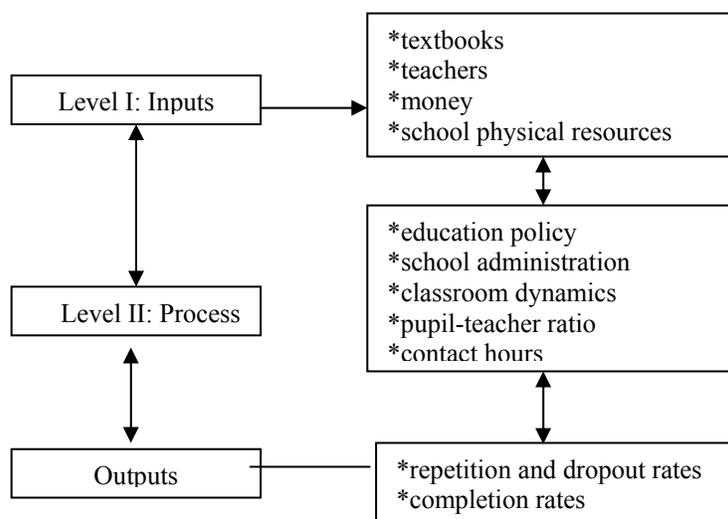


Figure 11: Holistic operation model of efficiency

Source: *Abagi & Odipo (1997: 9)*

Abagi and Odipo (1997: 10) used the measure of internal efficiency as “the amount of learning achieved during school age attendance, compared to resources provided”. Like others, they asserted that internal efficiency of an education system was revealed by the promotion, repetition and dropout rates. The results of their data analysis indicated that primary education in Kenya had internal efficiency problems. Abagi and Odipo (1997: 14) highlighted the factors behind the low internal efficiency in Kenya at primary school level. These factors were divided into three categories:

- a) Education policies and institutional processes: limited budget; political will; lop-sided priorities; poor management; monitoring and feedback
- b) School-based factors: rising poverty; over-loaded curriculum; teacher attitude; pupil motivation
- c) Household and community-based factors: household attitude; opportunity cost; gender issues and socialisation; religious factors

All the above-mentioned factors resulted in poor performance, repetition and dropout with eventual low completion rates at standard 8. Abagi and Odipo (1997: 24) indicated that there is a lot of wastage in primary education, because more than 50% of enrolled learners fail to complete the education cycle, yet education consumes about 55% of the government’s recurrent expenditure.

5.3.1.2. Educational attainment around the world

Filmer and Pritchett (1998a: 1) estimated educational enrolment and dropout around the world, including the effect of household wealth on educational attainment. They came up with the following findings

- The enrolment profiles of the poor differ across countries, but fall into distinctive regional patterns, e.g. typical of South America is that the poor reach nearly universal enrolment in the first grades, but then drop out in droves.
- There are huge differences between enrolment and educational attainment of the rich and the poor. Filmer and Pritchett (1998a: 1) refer to it as the “wealth gap”.
- Attainment profiles can be used as diagnostic tools to examine issues in the educational system.

The following structure will be used to discuss the exposition of Filmer and Pritchett (1998a): Data and Methodology; Enrolment and dropout patterns among the poor; “Wealth gaps” across countries; Attainment profiles as diagnostic tools.

5.3.1.2.1. Data and Methodology

Filmer and Pritchett (1998a) constructed an “asset index” from data of the Demographic and Health Surveys (DHS) to divide countries into three economic categories, namely “poor”, “middle” and “rich”. They also used the data to create an “attainment profile”. The attainment profile, according to Filmer and Pritchett (1998a: 10), is a graphical representation of the proportion of individuals aged 15 to 19 that completed each grade. For example the level at

grade 1 on the graph means the proportion that ever attended school and completed first grade. Filmer and Pritchett (1998a: 10) followed a different approach to estimate dropout rate namely “the slope of the enrolment profile on the graph is a simulation of dropout.” The dropout rate is a simulation because it is not obtained via the flow of a cohort through the education system, but a cross section of attainments of this cohort. It differs from the usual dropout rate in the sense that the denominator is all the children aged 15 to 19 as opposed to the proportion of those who reached the 5th grade. Filmer and Pritchett (1998a:10) explained that “in the attainment profile graphs the dropout rate is the vertical drop between grades as a proportion of the absolute height”.

Although Filmer and Pritchett (1998a: 11-17) showed the attainment profiles for each of 35 countries, only some of the results and patterns that emerged (refer to Appendix C) will be discussed here. These graphs and interpretations are of particular interest and key to this study. An attainment profile of the Western Cape will be created with the poor, middle and rich identified, based on the same principles used by Filmer and Pritchett (1998a).

To give a better understanding of their methodology and findings Filmer and Pritchett (1998a: 18) interpreted the attainment profile graph of Benin. The following deductions outlined their summary:

- Attainment profile of Benin: Only 26% of the poor aged 15 to 19 have completed grade 1 or higher, meaning 74 % have not completed even one year of schooling. Completion of grade 5 or higher was only 7.9 % and only 0.7% completed grade 9. They showed that

among the rich 80% completed grade 1 or higher, but dropout is such that only 54% completed grade 5.

- The “wealth gap” Filmer and Pritchett (1998a: 18) defined as the vertical distance between economic groups on the graph. For example the wealth gap in grade 5 is 0.46 (0.56 for the rich minus 0.10 for the poor).

From all these attainment profiles Filmer and Pritchett (1998a) asserted the following patterns among the poor.

5.3.1.2.2. Enrolment and dropout patterns among the poor

Filmer and Pritchett (1998a: 19) identified four enrolment and dropout patterns among the poor, which tend to follow regional patterns.

- Low enrolment and high dropout: In Western/Central Africa only between 4.6 (Mali) and 27 (Cote d’Ivoire) percent of poor children complete grade 5.
- Low enrolment and low dropout: In South Asia the fraction of poor children who did not complete grade 1 is very high (50%), but if they do start the retention is much higher.
- High enrolment and early high dropout: The Latin American pattern is one of high initial enrolment, but very steep dropout among the poor. All children start school, but then dropout is high. For example, in Brazil, of those that complete grade 1, only 16% of the poor economic group go on to complete primary school.
- High enrolment and late (East Africa) or very late dropout (East Asia and Central Asia/North Africa/Europe).

Filmer and Pritchett (1998a: 22, refer to Appendix C for graphical representation) asserted that from the above results the following findings for Latin America emerged: Grade 5 completion among the poor is 46% in Brazil, 57% in the Dominican Republic, 63% in Colombia, and peaks for Latin America at 75% in Peru. In contrast it is 89% in Zimbabwe, 84% in Kenya, 69% in Ghana, and 62% in Tanzania. The Eastern/Southern African countries retain higher proportions of the poor, as indicated by the flat portions in the graphs of the profiles of the poor in Kenya, Tanzania and Zimbabwe.

5.3.1.2.3. “Wealth gaps” across countries.

Filmer and Pritchett (1998a: 27) used the attainment profiles to determine the effect on attainment of wealth differentials within countries. In Western/Central Africa there are large attainment gaps at primary level, but these close at grade 9. In South Asia the wealth gap starts large and stays large. The wealth gaps in attainment in Eastern/Southern African countries are relatively small.

5.3.1.2.4. Attainment profiles as diagnostic tools

Filmer and Pritchett (1998a: 30-35) concluded that the attainment profiles can be used as a diagnostic tool to identify areas of concern in the system. Filmer and Pritchett (1998b) used the same methodology in another study to determine the effects of household wealth, gender, village, and state on educational enrolment and attainment in India. They likewise created attainment profiles of children aged 15 to 19 for the states of India. Based on these attainment

graphs they came to the conclusion that there are huge gaps between attainment of children for “rich” and “poor” economic groups in India.

In summary, the extant literature review shows how the Reconstructed Cohort Method is employed to determine the effects of repetition and dropout on internal efficiency. The basic principle of the method is that one can estimate the flow of a cohort of learners through a particular system in order to derive indicators of internal efficiency. The empirical studies show that dropout and repetition affect internal efficiency. Research findings indicate that repeating a grade does not improve a learner’s performance (Byrnes, 1989; Cuadra & Frederiksen, 1992; Grissom & Shepard, 1989; Holmes, 1989; UNESCO, 1998). Holding learners back increases rather than decreases their risk of dropping out of school. The findings of Grissom and Shepard (1989: 34) indicated that “retained students experience greater risk for dropping out that cannot be explained by their poor performance”. Some of the empirical studies on dropout and repeaters highlight the fact that there is a high dropout in grade 1, especially in developing countries (Filmer & Pritchett, 1998a UNESCO, 1998; Wako,n.d.).

The literature review on the Reconstructed Cohort Method and the Learner Attainment Profile shaped the conceptual and theoretical framework and served as a precursor for the empirical applications that follow.

5.3.2. The effect of socio-economic status, age and gender on the educational enrolment and attainment in the Western Cape

To construct an educational attainment profile for the Western Cape, this section follows a similar approach to the methodology of Filmer and Pritchett (1998a) using the 2001 census data.

5.3.2.1. Data and methodology

The variables in the South African census 2001 survey have been used to construct a poverty index and educational attainment profile for communities in the Western Cape.

5.3.2.1.1. Constructing a poverty index of the community

The basis of the poverty index of the community is an overall rating of each community's socio-economic status (SES). To measure the poverty of a community, a composite index has been constructed by using data of the 2001 census survey per suburb to describe the socio-economic status of the community. The overall rating of each community's socio-economic status is based on four indicators. Constructing each of the indices for these indicators is similar to that of the poverty index for census 1996 except that the income ranges are demarcated differently (refer to Appendix B).

The first step was to sort the suburbs by poverty index in descending order. The next step was to compute the cumulative proportion of the population ages 15 to 19 for each suburb. The cumulative population of any given suburb is obtained by adding its population proportion to the cumulative population of the immediately preceding suburb. From 40% of the population of individuals aged 15 to 19 in the sorted list the “poor” economy group is derived, “middle” from the next 40% and “rich” category from the bottom 20% (See Figure 12). The index is meant to identify a community in terms of its poverty and vulnerability. Communities high on the list are poor and vulnerable. Communities lower down the list are less vulnerable in terms of the criteria that were used to evaluate all communities (1 is poor and 0 is rich).

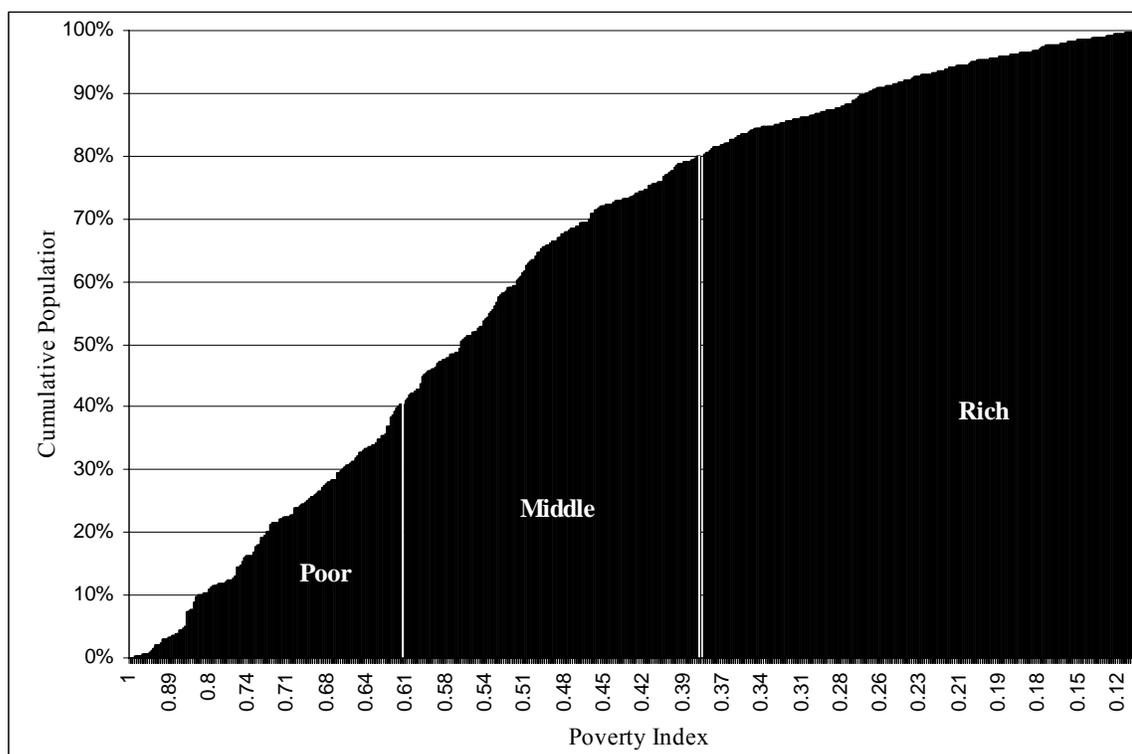


Figure 12: Cumulative percentage of population for ages 15 to 19 per suburb in the Western Cape according to poverty index

5.3.2.1.2. Attainment profiles for the age cohort 15 to 19 of the Western Cape

Table 11 shows the attainment profile for the age cohort 15 to 19, according to different economic groups in the Western Cape.

Table 11: Attainment profile for the ages 15 to 19 per grade, by economic group in the Western Cape

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9
poor	0.98	0.98	0.97	0.96	0.94	0.90	0.83	0.70	0.52
middle	0.99	0.99	0.99	0.98	0.97	0.95	0.92	0.84	0.68
rich	0.99	0.99	0.99	0.99	0.99	0.98	0.97	0.93	0.82

Table 11 displays the totals that were calculated based on the following variables in the census 2001) survey (Statistics South Africa, 2001):

- a) highest educational level (grade) obtained per person
- b) children ages 15 to 19 within a suburb. The 15 to 19 year olds are used in the analysis for three reasons:
 - Filmer and Pritchett (1998a) used the same cohort in their analysis, thus doing so would make comparisons with other countries possible by drawing from their study.
 - The cohort 15 to 19 are all supposed to be at school and would give a clear indication of the dropout phenomenon, unlike older cohorts who may have already

left their parental home at the time of the census and whose socio-economic status in childhood can thus not be established from the census data.

- It is fair to say that this is the age cohort that takes the economic status of their parents unlike older cohorts who may have already left the parental home. It therefore makes sense to use this cohort in the analysis of different economic groups, namely poor, middle and rich.

For example the proportion of attainment (enrolment ratio) for grade 1 was obtained by taking the number of individuals aged 15 to 19 within an area minus the total of individuals who never attended school. The proportion of attainment for grade 2 was obtained by subtracting the total individuals with grade 1 as the highest education level from the estimated proportion attainment for grade 1. The dropout rate is a simulation because it is not obtained via the flow of a cohort through the education system over time, but from a cross section of attainments of this cohort. It differs from the usual dropout rate in the sense that the denominator is all the children ages 15 to 19 as opposed to the proportion of those who reached the specific grade.

Figure 13 is a graphical representation of Table 11 and shows the enrolment and dropout per grade for three different economic groups in the Western Cape.

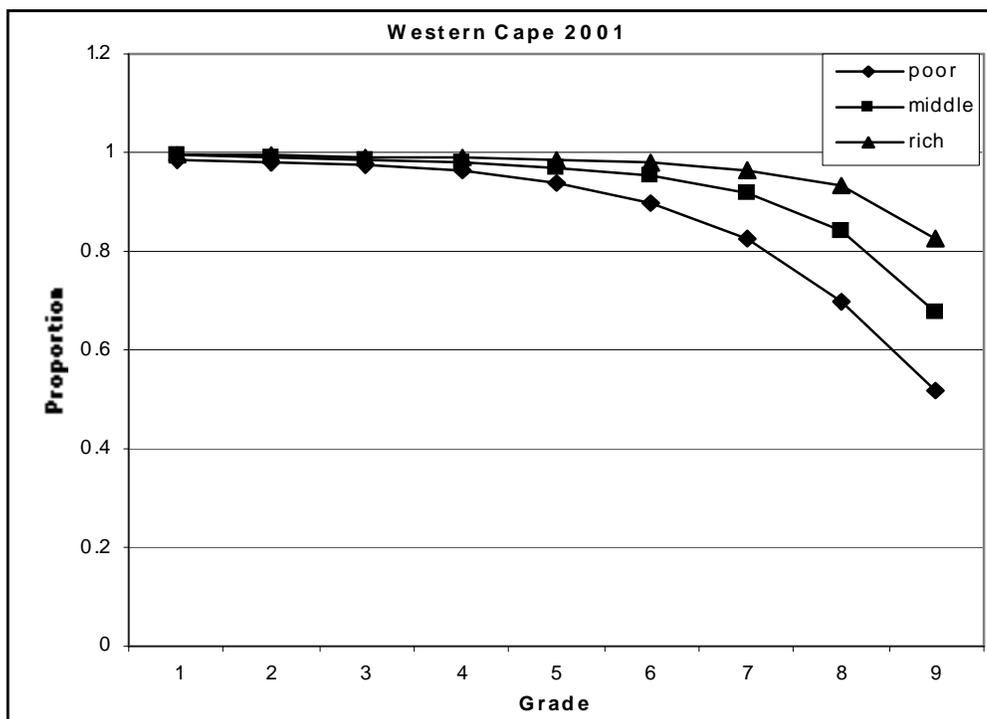


Figure 13: Attainment profile for ages 15 to 19, by economic group

In Figure 13, the attainment graph for the Western Cape, the dropout rate is the vertical drop between grades as a proportion of the absolute height. The census data indicates the highest grade obtained by each person. These individuals with the highest education level for a particular grade become the dropouts for the next grade. For example, individuals with grade 1 as the highest education level become the dropouts between grades 1 and 2. In other words the difference between the proportion of grade 5 enrolment and proportion of grade 6 enrolment, is an estimate of the proportion of all the children that dropped out between 5th and 6th grade.

The attainment profile of the Western Cape (Figure 13 & Table 11) shows that for all economic groups the attainment of learners in primary school education (grade 1 to 7) is very high in the early grades, as the flat portion in the graph indicates. There is, however, a high

dropout rate in all economic categories in the latter grades with the highest attrition (0.18) in the “poor” group between the grades 8 and 9.

The attainment profiles of 15 – 19 year olds, although it is a cross-section of attainments, show what was happening over the past 10-15 years. These profiles are not a snapshot of what is happening today. The attainment profile of the 15-19 year olds in 2001 (Census 2001 dataset was used) really reflects the schooling system as it existed starting in 2001 – 19 + 6, meaning 1988. It is not a snapshot of what the system’s flow rates would be at the present.

5.3.2.1.3. Wealth gaps

The Wealth Gap is represented graphically (Figure 13) as the vertical distance between two economic groups. For example the wealth gap in grade 7 enrolment is 0.14 (0.97 for rich versus 0.83 for the poor enrolment in Table 11). The wealth gaps in the primary school (grades 1 to 7) are very small in the early grades, but then increases in the latter grades of the primary phase.

5.3.2.1.4. Attainment profile for the age cohort 20 to 30 by gender of the Western Cape

To investigate attainment at higher grades, we turn to the cohort 20 to 30 years of age. Unfortunately, the socio-economic status of most members of this cohort whilst they were in the education system is not known from the census, as most of them no longer reside in their

parental home. The attainment profile of the cohort 20 to 30 by gender is depicted in Table 12 with a graphical representation in Figure 14.

Table 12: Attainment profile for the ages 20 to 30 per grade by gender: Western Cape

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Male	0.97	0.97	0.96	0.95	0.93	0.90	0.86	0.79	0.70	0.61	0.49	0.42
Female	0.97	0.97	0.97	0.96	0.95	0.93	0.90	0.84	0.77	0.67	0.56	0.47
All	0.97	0.97	0.96	0.95	0.94	0.91	0.88	0.82	0.74	0.64	0.53	0.44

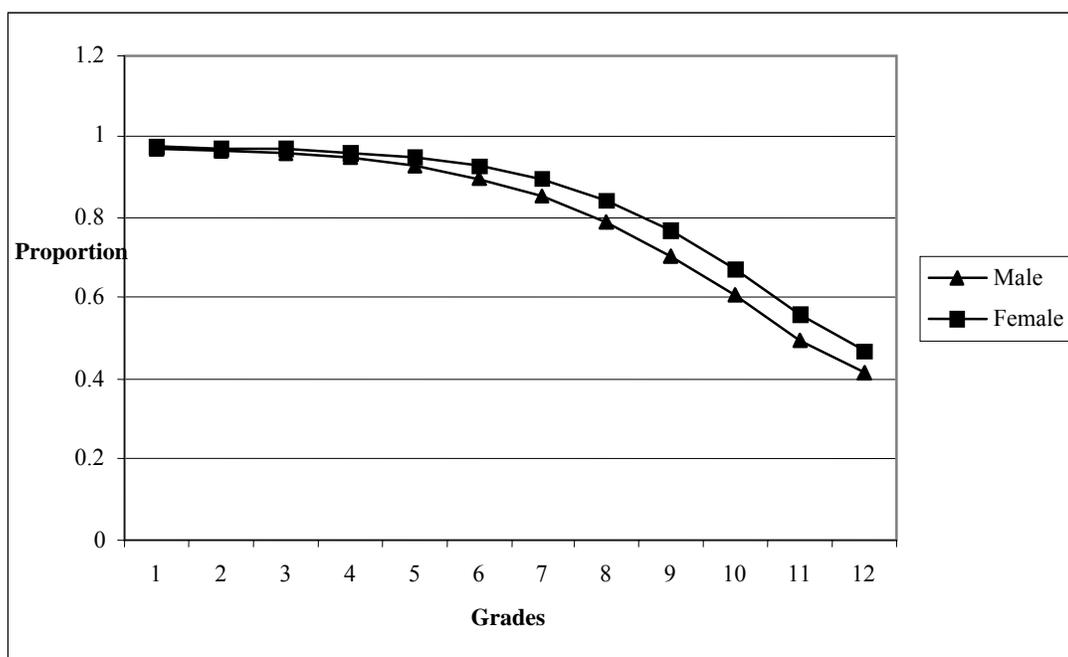


Figure 14: Attainment profile for ages 20 to 30: Western Cape

At this point it might be important to compare the findings of the attainment profile of the cohort 20 to 30 (Table 12 and Figure 14) with other data resources. The attainment profile for the ages 20 to 30, that was constructed similar to the cohort 15 to 19, offered interesting results. The estimates in the profile (Table 12 and Figure 14) have the same meaning as the survival rates (Table 23 and Figure 27) derived from the reconstructed cohort method. Although different data sources and methods were used the outcomes reflect the same results (refer to the patterns in Figures 17, 18, 21, 22 and 27):

- The results in Figure 14 compare well with the patterns of the actual enrolment data (Figure 17).
- The flat structure of the graph (Figure 14) indicates that there is a low dropout in lower grades with high dropout in higher grades.
- The steepness of the graph (Figure 14) shows that there is high dropout between grades 10 to 12, with the highest dropout between grades 10 and 11, a phenomenon that is the same in survival rates (see Table 23).
- Differences by gender in the attainment profile show that there is a higher dropout of males than females, especially in the higher grades. The attainment of female learners in the higher grades is better than that of male learners (see Table 12 and Figure 14).

5.3.2.1.5. Comparison of the attainment profile of the Western Cape with other provinces

A comparison of attainment profiles for the age cohort 20 to 30 across Provinces is summarised in Table 13 and shows the differences in attainment profiles across provinces.

Table 13: Attainment profile for the ages 20 to 30 per grade by Province

Province	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Eastern Cape	0.90	0.89	0.88	0.86	0.83	0.80	0.76	0.70	0.62	0.53	0.42	0.31
Free State	0.94	0.94	0.93	0.92	0.90	0.87	0.83	0.76	0.69	0.60	0.48	0.36
Gauteng	0.96	0.96	0.95	0.95	0.94	0.92	0.90	0.86	0.81	0.75	0.65	0.53
KwaZulu-Natal	0.90	0.89	0.88	0.87	0.85	0.82	0.79	0.74	0.68	0.61	0.50	0.39
Limpopo	0.87	0.87	0.86	0.85	0.84	0.81	0.79	0.73	0.66	0.57	0.46	0.30
Mpumalanga	0.89	0.89	0.88	0.87	0.85	0.82	0.79	0.73	0.67	0.59	0.48	0.36
Northern Cape	0.91	0.91	0.90	0.88	0.85	0.81	0.75	0.67	0.59	0.49	0.39	0.32
North West	0.91	0.91	0.90	0.88	0.85	0.82	0.78	0.73	0.67	0.59	0.49	0.38
Western Cape	0.97	0.97	0.97	0.96	0.94	0.92	0.88	0.82	0.74	0.64	0.53	0.44

From Table 13 it becomes evident that:

- In all the provinces there is a high attainment of learners in the lower grades (Figure 15)
- In all the provinces there is high dropout in the higher grades
- The Western Cape is, after Gauteng, the best in retaining its learners in the system. The Eastern Cape is one of the Provinces with the highest dropout of learners.

This trend becomes clearer when presented graphically. Figure 15 shows the decline across Provinces, especially in the higher grades.

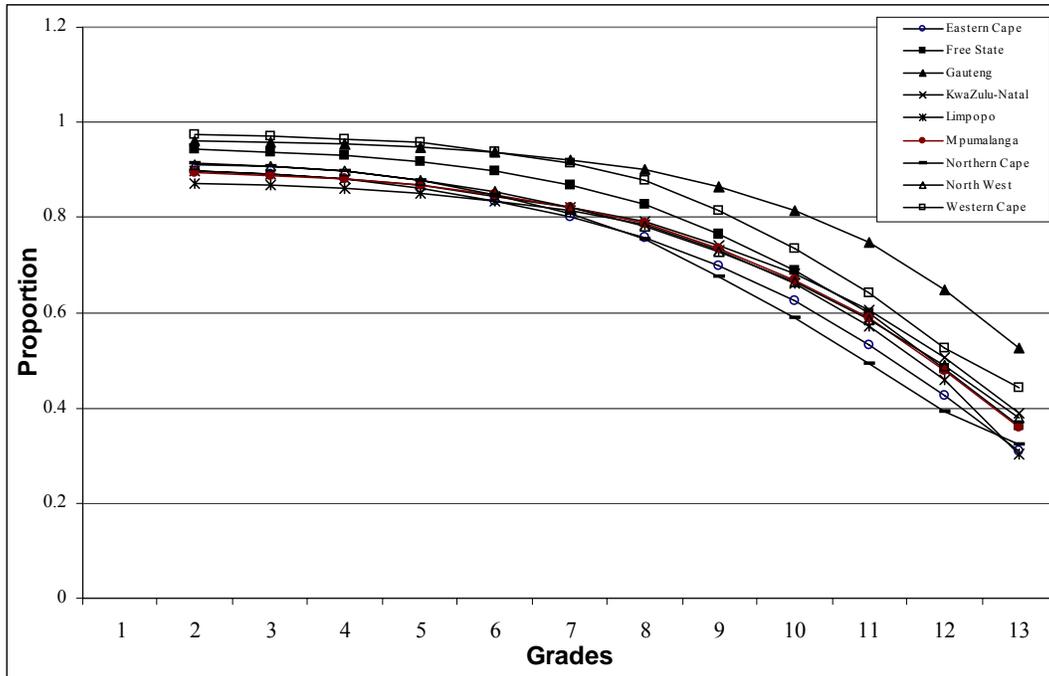


Figure 15: Attainment profile for the ages 20 to 30 per Province

5.3.2.1.6. Attainment profile for different age cohorts

A comparison of attainment profiles across different age cohorts in the Western Cape is summarised in Table 14.

Table 14: Attainment profile for different cohorts in the Western Cape

Cohorts	Grade1	Grade2	Grade3	Grade4	Grade5	Grade6	Grade7	Grade8	Grade9	Grade10	Grade11	Grade12
20-24	0.98	0.97	0.97	0.96	0.95	0.93	0.90	0.84	0.77	0.67	0.55	0.45
25-29	0.97	0.97	0.96	0.95	0.93	0.90	0.86	0.79	0.71	0.62	0.50	0.42
30-34	0.97	0.96	0.95	0.94	0.91	0.88	0.83	0.75	0.66	0.56	0.44	0.37
35-39	0.95	0.95	0.94	0.92	0.88	0.84	0.78	0.69	0.58	0.49	0.36	0.30
40-44	0.94	0.93	0.92	0.89	0.86	0.81	0.75	0.65	0.54	0.46	0.33	0.29
45-49	0.92	0.91	0.90	0.87	0.83	0.78	0.71	0.62	0.50	0.43	0.30	0.27
50-54	0.91	0.90	0.89	0.86	0.82	0.77	0.70	0.61	0.49	0.43	0.30	0.27
55-59	0.90	0.89	0.88	0.85	0.82	0.77	0.71	0.63	0.50	0.45	0.32	0.29
60-64	0.88	0.87	0.86	0.83	0.79	0.75	0.69	0.62	0.49	0.44	0.31	0.28
65-69	0.88	0.87	0.86	0.83	0.80	0.76	0.70	0.63	0.49	0.44	0.31	0.29
70-74	0.88	0.87	0.86	0.84	0.80	0.77	0.72	0.65	0.52	0.48	0.33	0.32
75-79	0.89	0.88	0.87	0.84	0.81	0.78	0.74	0.68	0.55	0.50	0.37	0.35
80-84	0.88	0.87	0.86	0.84	0.81	0.78	0.74	0.69	0.55	0.51	0.38	0.36

The attainment profiles of all the cohorts in Table 14 make it possible to infer how a group of learners proceeded through the system and compare it with other age cohorts. Individualised data, through the census survey, now make it possible to simulate the schooling career of each learner, for example how many graduate, how many dropout, where is the system, over time, the least efficient, etc.

Figure 16 is a graphical representation of the different age cohorts in the Western Cape

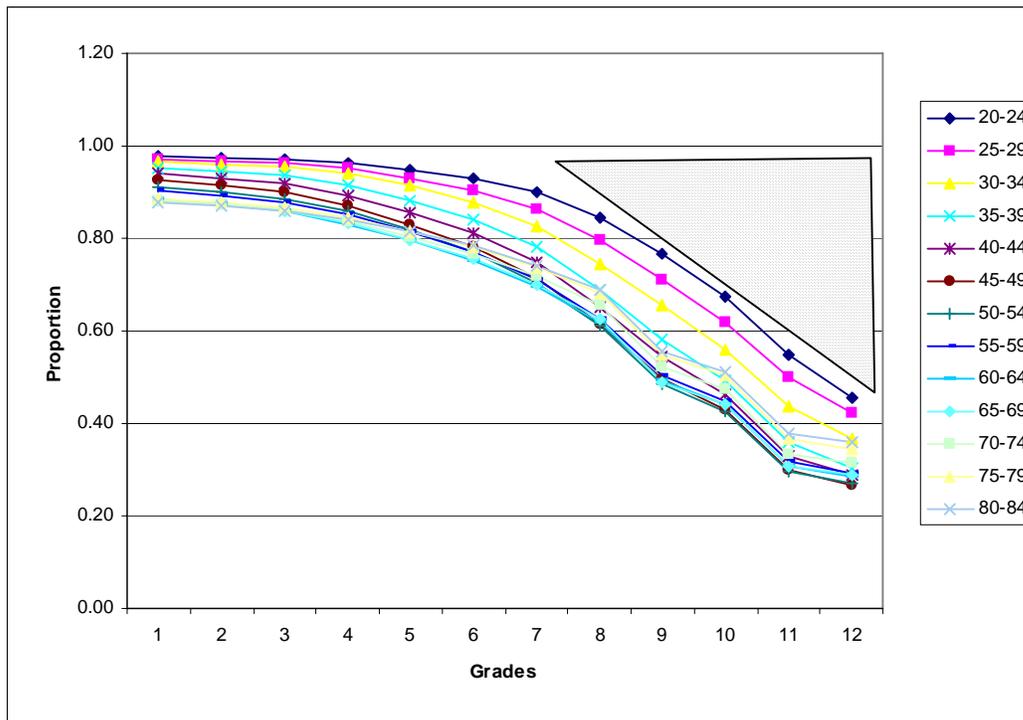


Figure 16: Attainment profile in Western Cape for different age cohorts

The flat structure of the graph (Figure 16) for all cohorts shows that the system simply retains its learners for the entire primary school phase (grade 1 to grade 7).

The dropout is the highest between grades 10 and 11, a phenomenon that is also evident in other datasets (refer to Figure 25 – dropout rate). The size of the triangle in Figure 16 is the size of the dropout problem in the system. The differences in attainment profiles of cohorts could indicate if there was a significant improvement in the efficiency of the education system. In other words did the educational attainment improve over a certain period and in which age cohort? In the Western Cape the percentage that reach grade 12 is the highest for the age cohort 20- 24.

5.3.2.2. Attainment profile as a diagnostic tool in the education system of the Western Cape

The attainment profile of the Western Cape can be used as a diagnostic tool to determine where key concerns in the education system are. From Figures 14, 15 and 16 it becomes evident that there are high dropout rates in the higher grades as indicated by the vertical steepness of the graph.

The dropout rate is a simulation because it is not obtained via the flow of a cohort through the education system over time, but a cross section of attainments of this cohort.

It is important to note that the attainment profile could be used as a diagnostic tool in the education system. The results of such analysis could serve as a catalyst for further research. However, the reliability of the education attainment profile is totally reliant on the accuracy and the quality of the census data.

The next section will describe how the Reconstructed Cohort Method will be used to transform data for sensemaking and better understanding of the internal efficiency of education system. .

5.3.3. The Reconstructed Cohort Method in the education system of the Western Cape

5.3.3.1. Introduction

This section reports on the empirical application of the Reconstructed Cohort Method, also called the Grade Transition Model (GTM) to the education system of the Western Cape. The basic principle of this method is that one can construct (estimate) the flow of an entrance cohort of learners through a particular system and predict the eventual outcomes at the end of the cycle.

The essential factors necessary to do this kind of reconstruction are the rates that influence the flow (increase or decrease) from one period to another, called the transition period. The duration of a transition period could be from one year to another in the case of an education system, or it could be five years in the case of population studies when census data are used to determine for example population growth. Four basic transition rates are necessary to determine the flow of a cohort through a cycle, namely the rates of those:

- progressing through the system (promoting)
- not progressing through the system (repeating)
- leaving the system (dropping out)¹⁶
- entering into the system (learner transfers).

¹⁶ In this study, dropouts refer to the learners dropping out of the WCED system. However, they could have continued their education in other education systems, such as private institutions, colleges or home schools, moved to other Provinces or countries or even died.

In this study the reconstructed cohort method will be dealt with in two major parts. The first part describes the raw data needed to apply this method and how the raw data are utilised to calculate the transition rates. The second part is an extensive outline of the application of these transition rates to ascertain the indicators of efficiency.

5.3.3.2. The reconstructed cohort method: Data requirements

In order to construct this model, the data need to conform to certain requirements.

To reconstruct the flow of learners through an education cycle, transition rates such as repeaters, promotions and dropouts per grade are needed. The calculation of these rates requires minimum raw data of enrolment per grade of at least two consecutive years (transition period) and repeaters per grade of the latter year.

Data of transfers from other systems, such as from other countries or provinces, private institutions, those who were not enrolled in any school the previous year and home schools would enhance the accuracy and reliability of the estimation of these rates. The same applies for transfers to other systems. In the Western Cape the data are obtained annually through a survey that is completed by each school. The quality of the data in terms of completeness, relevance, accuracy, timeliness and accessibility are key to the application of the reconstructed cohort method in measuring the internal efficiency of the education system. As mentioned before in Chapter 3 transactional processes such as data collection, data capturing, data verification and validation, and data integration are therefore important in the assurance of quality data.

5.3.3.2.1. Computation of flow rates using data on repeaters and enrolment

Analysis of enrolment by year and gender per grade from a sensemaking perspective

The data in Table 15 is a summary of the enrolment by year, grade and gender in the Western Cape and facilitates the calculation of transition rates (repeaters, promotion and dropouts) between any two years. It elicits the relationships and patterns in the enrolment over time.

Table 15: Enrolment by year, grade and gender in the Western Cape (1995 to 2002)

YEAR	GENDER	GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11	GRADE 12
1995	Female	48 024	41 283	39 674	39 115	38 155	36 291	35 182	34 808	32 015	26 864	22 213	19 700
1995	Male	51 134	43 680	41 247	40 084	38 634	36 002	33 613	32 665	29 064	23 834	18 911	14 959
1995	TOTAL	99 158	84 963	80 921	79 199	76 789	72 293	68 795	67 473	61 079	50 698	41 124	34 659
1996	Female	48188	41759	39891	39115	38813	37095	35295	36955	32920	29122	23820	20789
1996	Male	51810	43086	41246	40881	38860	37047	34179	35161	29776	25318	19754	15975
1996	TOTAL	99 998	84 845	81 137	79 996	77 673	74 142	69 474	72 116	62 696	54 440	43 574	36 764
1997	Female	46931	41153	39556	39132	37486	37171	35523	37504	33973	30291	24099	22226
1997	Male	50923	43739	41052	42059	39408	37197	35444	35757	30919	26779	20487	16714
1997	TOTAL	97 854	84 892	80 608	81 191	76 894	74 368	70 967	73 261	64 892	57 070	44 586	38 940
1998	Female	48033	44108	42286	41722	40007	38273	37706	38861	35031	32917	25390	23295
1998	Male	51347	46060	43339	43466	40874	38443	36222	37823	32358	28771	20601	17622
1998	TOTAL	99 380	90 168	85 625	85 188	80 881	76 716	73 928	76 684	67 389	61 688	45 991	40 917
1999	Female	42841	45083	43635	43196	41276	39207	37611	40309	36183	34017	27282	23202
1999	Male	44595	47842	44978	44818	41798	39288	37050	38734	33491	29462	21965	17004
1999	TOTAL	87 436	92925	88613	88014	83074	78495	74661	79043	69674	63479	49247	40206
2000	Female	31083	40237	44964	44770	42378	42378	38169	40467	36922	34369	26874	23488
2000	Male	33761	41628	47379	47179	43388	43388	37644	39559	33712	29471	22060	17508
2000	TOTAL	64844	81865	92343	91949	85766	85766	75813	80026	70634	63840	48934	40996
2001	Female	40048	30237	40457	45606	43954	41713	39698	41334	37504	36419	27769	22536
2001	Male	41742	32723	41375	48696	45300	41592	38080	40856	34462	30615	22437	17374
2001	TOTAL	81 790	62 960	81 832	94 302	89 254	83 305	77 778	82 190	71 966	67 034	50 206	39 910
2002	Female	42599	37959	30897	40736	45452	43143	40998	38889	40970	37530	28650	22564
2002	Male	44370	39067	33237	42286	47736	43643	39867	36712	39480	32222	22968	17904
2002	TOTAL	86 969	77 026	64 134	83 022	93 188	86 786	80 865	75 601	80 450	69 752	51 618	40 468

Source: Annual School Survey (WCED)

The graph, Figure 17 is based on the enrolment data in Table 15. The similar pattern of all the lines in the graph over the years is a reflection of data accuracy and the stability of the education system.

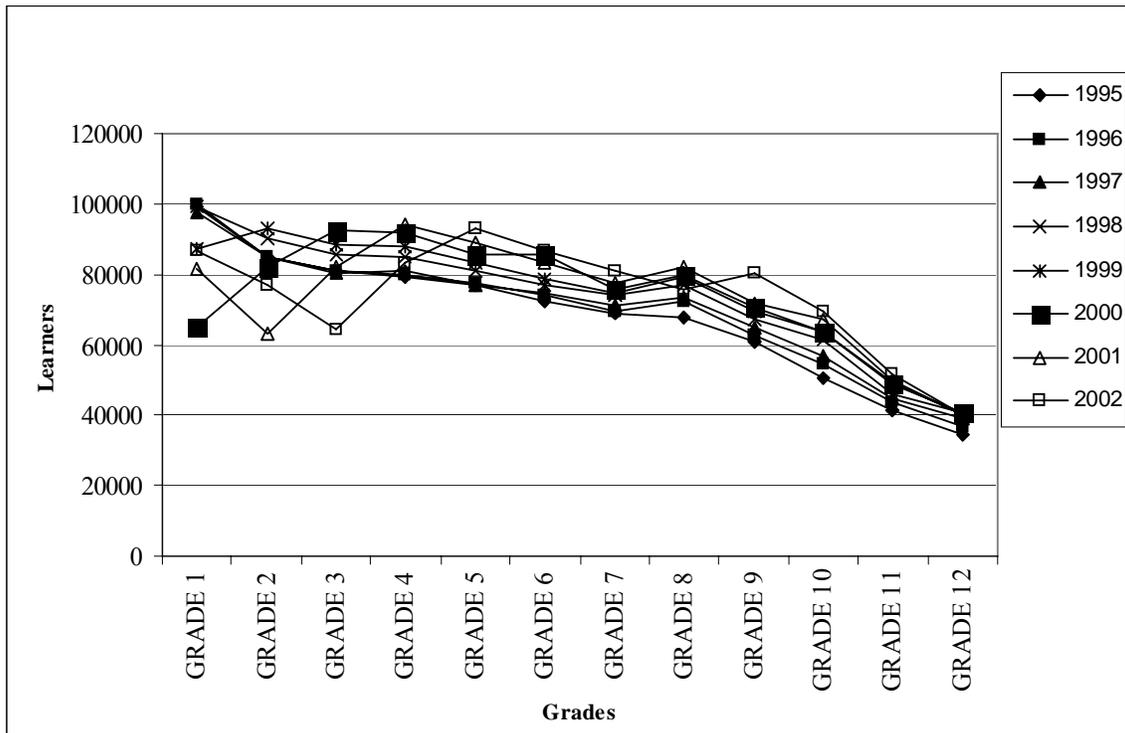


Figure 17: Enrolment by year, grade and gender per grade in the Western Cape (1995 to 2002)

An analysis of the data in Table 15 and the graph (Figure 17) answers a few critical questions and can therefore be used as a diagnostic tool, such as at which grades are the dropout the highest. Another interesting inference is that Figure 17 (enrolment by year and gender per grade) has a similar pattern as the attainment profile for the cohort 20 to 30 in the Western Cape (Figure 14), although different data sources were used.

The following analysis is based mainly on data in Table 15 and Figure 17. In order to extract knowledge and make better sense of the data the research focuses on a few areas with the aid

of other graphs and figures. It elicits the relationships and patterns in the enrolment over time and facilitates further research and investigation in smaller groups and areas.

- 1) Based on the data in Table 15 and the graph (Figure 17) it is clear that there is a high dropout rate of learners between grades 8 and 12. To make better sense of this dropout phenomenon amongst secondary learners would require further research. It could be due to the poor preparation of learners during the early childhood development, but it is entirely outside the scope of this research.
- 2) Further analysis reveals that the highest dropout of learners seems to occur between grades 10 and 11 (that is where the gradient in the graph of Figure 17 is the steepest). This could either be the result of a high repeater rate in grade 10 or a high dropout after grade 10. Four issues could shed light on this occurrence:
 - Over and under age analysis.
 - Age specific comparison (Comparison of learners enrolled in the education system with those appropriate age groups in the population according to Census data).
 - The repeater rate over a period (Figure 23 confirms this). Examining the reasons why specifically grade 10 has the highest repeater rate would make for an interesting and insightful analysis, but it falls beyond the scope of this research.
 - The average dropout rate (Figure 25) confirms the high dropout rate in grade 10.
- 3) Likewise, the apparent dropout for Grade 1 is high and should be investigated. It could be that the actual repetition in Grade 1 is higher than the reported repetition and that learners included in original enrolment are not regarded as repeaters when they fail, because they are too young – the age specific comparison (Figure 19 & 20) confirms this. If that is the case, then drop-off in enrolment between Grades 1 and 2 is not really a dropout, but unofficial (unreported) repetition in Grade 1 of underage learners. The influence of such

under-estimation of repetition will be further investigated under the Reconstructed Cohort Method to illustrate the sensitivity of the learner through-put pertaining to repetition.

- 4) The enrolment figures for Grade 1 in 2000, Grade 2 in 2001 and Grade 3 in 2002 seem to be lower than the rest of the Grades. In 2000 the Western Cape Education Department implemented a policy that only learners reaching the age of 7 in that year may enter school (Grade1) and therefore these figures reflect a similar population and enrolment comparison as indicated in Figure 19.

- 5) Figure 18, which was created from the data in Table 15, is a presentation of the actual flow of the *same* cohort of learners (grade 5 in 1995) through to the end of the cycle (grade 12 in 2002). The reason why the cohort starts with Grade 5 in 1995 is because that is the first year when school survey data became available for all the schools in the Western Cape. Another reason why it was necessary to start the cohort at grade 5 and not at grade 1 is because it takes the flow of the cohort to the end of the cycle.

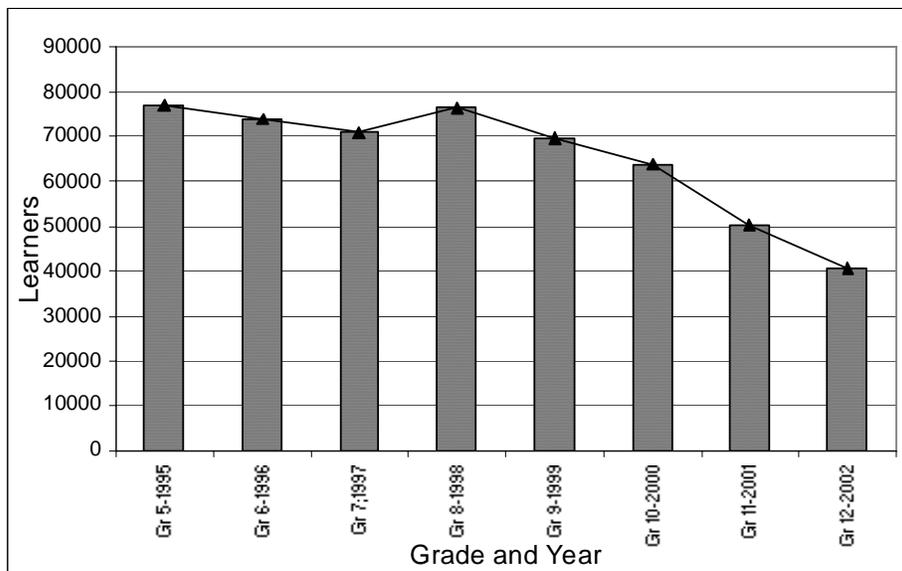


Figure 18: Continuous flow of learners from grade 5 (1995) to grade 12 (2002): Western Cape

Figure 18 emphasises the continuous flow of learners and shows the high dropout rate between grade 8 and grade 12. The blip is again clearly visible at grade 8. It seems that there are more learners in grade 8 in 1998 than learners in grade 7 in 1997. The possible influx of learners from other provinces or countries has been mentioned before. On the other hand, for certain grades there are a drop-in (blip in the graph in Figure 18), i.e. in grade 8 there are more learners in the system than the previous year. A possible reason for this could be that learners are coming into the system from other provinces, countries, private and home schooling. Another reason could be an exceptionally high repeater rate in grade 8.

6) Figures 19 and 20 are closely related to Table 15 and discussed below. Figure 19 is a comparison between the enrolment by year in grade 1 according to Table 15 and the population for the appropriate age (7) in grade 1. Figure 20 is the percentage of learners enrolled in grade 1 in comparison with the appropriate age in the population.

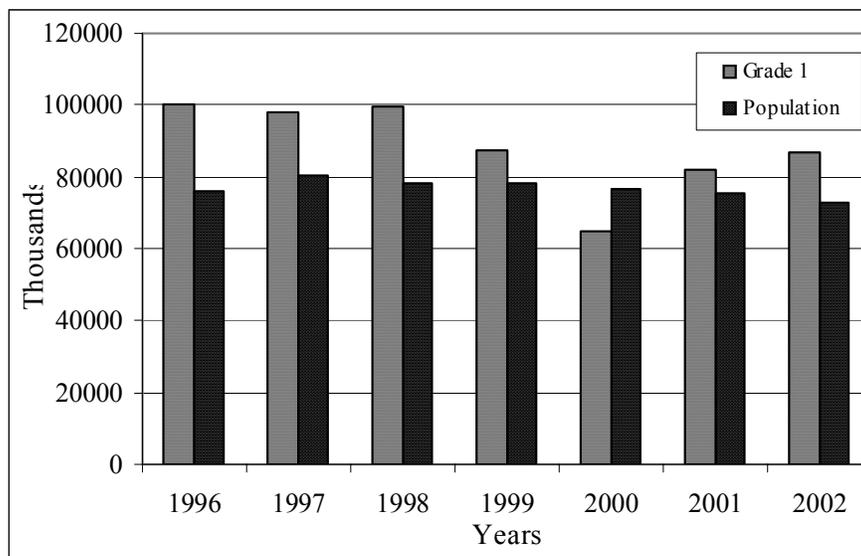


Figure 19: Age specific comparison (Comparison of Grade 1 enrolment with corresponding age of grade 1 learners in the in the Population): Grade 1 by year (1996 to 2002): Western Cape

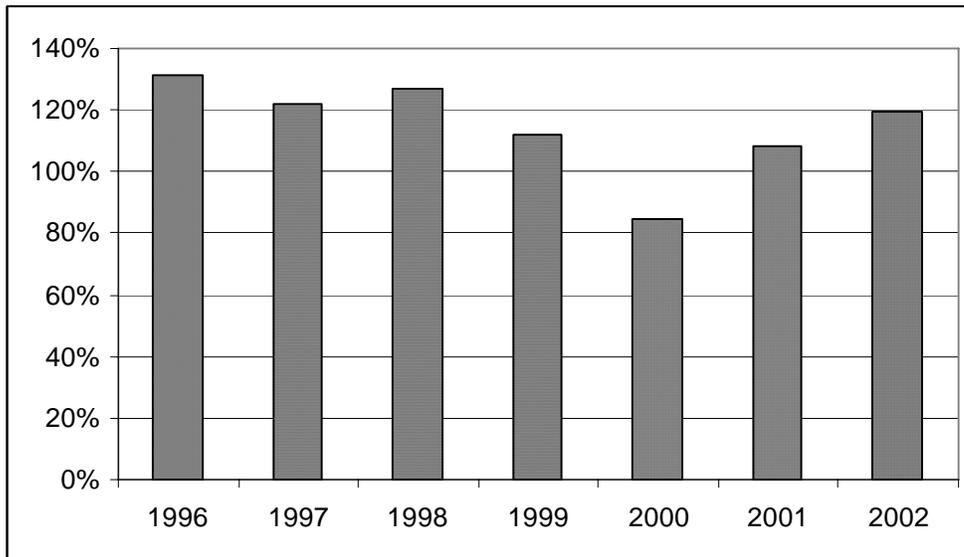


Figure 20: Age specific ratio (The percentage of learners in the Population with the corresponding age of those in grade 1): Grade 1 by year (1996 to 2002): Western Cape

7) Figures 19 and 20 confirm the under age problem in Grade 1 (as discussed earlier) and the subsequent drop out of the system. According to Figures 19 and 20 there are clearly more learners enrolled in the education system than the appropriate school age in the population indicated. The only exception is in 2000 when the Western Cape Education Department implemented a policy that only learners reaching the age of 7 in that year may enter school (grade1) and therefore reflects the correct population and enrolment comparison as indicated in Figure 19.

8) The Figures 21 and 22 confirm the dropout from the system in the higher grades. Figure 21 is a comparison between the population and enrolment in grades with similar age. Figure 22 is an age-specific enrolment ratio that relates the enrolment of a given grade in a given year to the population of the same age in the same year. The result concurs with the

findings in other data sources such as the attainment profile in Table 12 and the survival rates in Table 23.

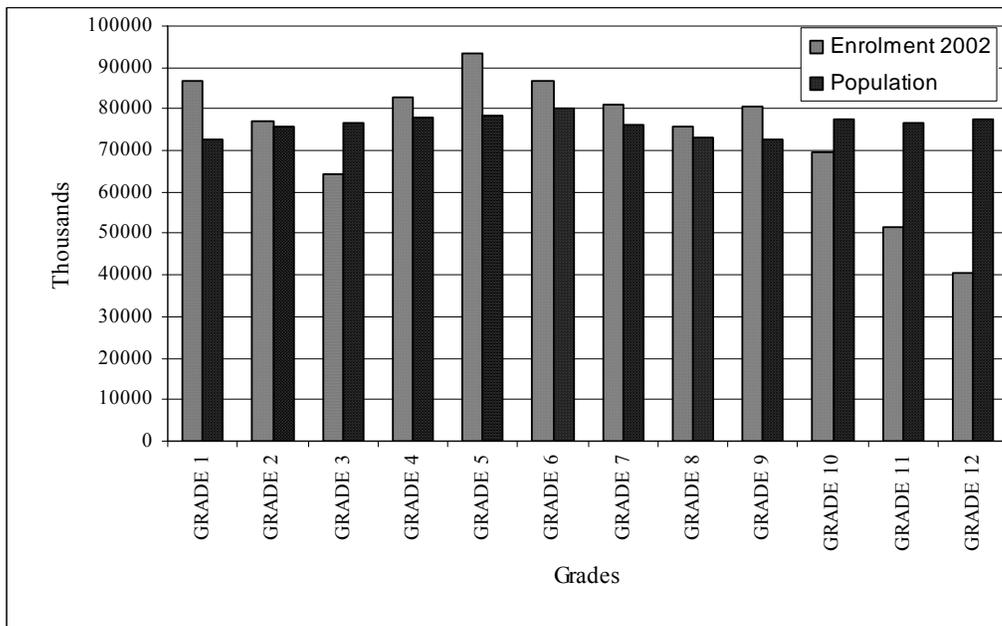


Figure 21: Comparison of Enrolment per Grade and Population in grade-appropriate age for 2002: Western Cape

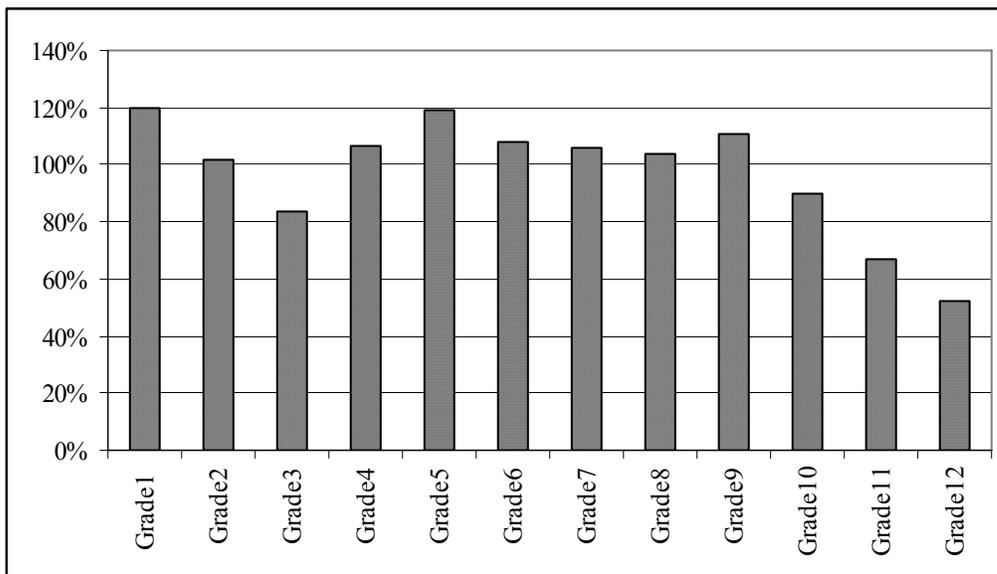


Figure 22: Age specific ratio 2002 (The comparison of appropriate age in Population with the corresponding age in the respective grades): Western Cape

Repeaters by year and gender per grade

Repeaters or those learners not progressing through the system are necessary to construct the flow of learners through the system. Table 16 displays the repeaters in the Western Cape by year, grade and gender. The data was obtained through the annual school survey. The data in Table 16 together with the enrolment data (Table 15) will be used in the calculation of the repeater rate.

Table 16: Repeaters by year, grade and gender in the Western Cape (1995 to 2002)

YEAR	GENDER	GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11	GRADE 12
1995	Female	5 776	3 502	2 794	2 259	2 315	2 160	2 209	3 717	4 444	2 980	2 153	445
1995	Male	7 942	5 326	4 322	3 574	3 467	2 767	2 613	4 141	4 686	3 141	2 078	353
1995	TOTAL	13 718	8 828	7 116	5 833	5 782	4 927	4 822	7 858	9 130	6 121	4 231	798
1996	Female	5399	2755	2134	1914	1737	1412	1009	3323	3839	2904	1811	602
1996	Male	7493	4027	2768	3061	2395	1704	1145	3505	3833	2856	1570	624
1996	TOTAL	12 892	6 782	4 902	4 975	4 132	3 116	2 154	6 828	7 672	5 760	3 381	1 226
1997	Female	5060	2802	1922	2064	1430	1047	756	3369	3111	2668	1336	2331
1997	Male	6890	3778	2616	3160	2122	1427	931	3877	3615	2833	1314	1151
1997	TOTAL	11 950	6 580	4 538	5 224	3 552	2 474	1 687	7 246	6 726	5 501	2 650	3 482
1998	Female	4791	2612	1893	2237	1547	1056	759	3533	3475	3771	1798	2748
1998	Male	6651	3602	2612	3716	2434	1589	1008	3993	3812	3913	1764	1187
1998	TOTAL	11 442	6 214	4 505	5 953	3 981	2 645	1 767	7 526	7 287	7 684	3 562	3 935
1999	Female	2571	2276	1687	2083	1581	991	698	3994	3887	3974	2252	1569
1999	Male	3417	3117	2304	3370	2419	1536	1010	4766	4178	4102	2222	713
1999	TOTAL	5 988	5 393	3 991	5 453	4 000	2 527	1 708	8 760	8 065	8 076	4 474	2 282
2000	Female	1967	1676	1675	2339	1672	1233	795	3994	3608	3933	2366	1254
2000	Male	2916	2380	2295	3747	2630	1702	1230	4757	4137	4062	2115	617
2000	TOTAL	4 883	4 056	3 970	6 086	4 302	2 935	2 025	8 751	7 745	7 995	4 481	1 871
2001	Female	1395	1273	1347	1965	1229	883	380	3831	3921	4633	2651	1111
2001	Male	2036	2088	2101	3345	2107	1321	600	4944	4574	4704	2690	689
2001	TOTAL	3 431	3 361	3 448	5 310	3 336	2 204	980	8 775	8 495	9 337	5 341	1 800
2002	Female	1654	830	1099	1209	1049	774	444	882	3029	4896	2844	990
2002	Male	2428	1265	1697	2269	1901	1177	751	1405	3453	4991	2642	791
2002	TOTAL	4 082	2 095	2 796	3 478	2 950	1 951	1 195	2 287	6 482	9 887	5 486	1 781

Source: Annual School Survey (WCED)

Transfers by year and gender per grade

Table 17 displays data on transfers by grade and gender in the Western Cape. The data on learners that have been transferred from other systems into the education system of the Western Cape are only available for the year 2002 (Table 17).

Table 17: Transfers of learners from other systems into the education system of the Western Cape in the year 2002

Year	Gender	Grade1	Grade 2	Grade3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
2002	Female	229	598	702	717	876	729	756	972	1011	1277	545	61
2002	Male	226	697	707	743	863	729	710	754	805	928	465	52
2002	Total	455	1295	1409	1460	1739	1458	1466	1726	1816	2205	1010	113

The data in Table 17 was obtained through the annual school survey. An added benefit when this kind of data is available is that it could be used to give a more accurate estimation of the promotions of learners and eventually of the reconstructed cohort method.

Promotions in the Western Cape by year and gender per grade

Table 18 displays the promotions by year and gender between 1995 to 2001 in the Western Cape and has been calculated based on the enrolment and repeaters data as presented in Tables 15 and 16 respectively. For example the promotees of grade 1 in 2001 are calculated by subtracting the number of repeaters in grade 2 in 2002 from the enrolment of that grade in 2002. To calculate the total promotees of grade 1 in 2001 the 2095 repeaters of grade 2 in 2002 are subtracted from the 77026 of grade 2 in 2002 to give the result of 74931 as indicated

in the Table 18. By subtracting the 1295 transfers (Table 17) for grade 2 in 2002 from this result the promotees for grade 1 in 2001 of 73636 is a better estimation of the promotions. Unfortunately only data of the transfer of learners into the system for 2002 are available. The last three rows in the Table 18 indicate the promotees for 2001 with provision also for the transfers from outside into schools for each grade¹⁷.

¹⁷ Because no data is available on transfers to other provinces or to private schools, such people will be recorded as dropouts by this method. This would exaggerate the dropout rate. However, Census data show that migration from the Western Cape is far lower than migration into the Province.

Table 18: Promotees by year and gender (1995 to 2001): Western Cape

YEAR	GENDER	GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11
1995	Female	39 004	37 757	37 201	37 076	35 683	34 286	33 632	29 081	26 218	22 009	20 187
1995	Male	39 059	38 478	37 820	36 465	35 343	33 034	31 656	25 943	22 462	18 184	15 351
1995	TOTAL	78 063	76 235	75 021	73 541	71 026	67 320	65 288	55 024	48 680	40 193	35 538
1996	Female	38351	37634	37068	36056	36124	34767	34135	30862	27623	22763	19895
1996	Male	39961	38436	38899	37286	35770	34513	31880	27304	23946	19173	15563
1996	TOTAL	78 312	76 070	75 967	73 342	71 894	69 280	66 015	58 166	51 569	41 936	35 458
1997	Female	41496	40393	39485	38460	37217	36947	35328	31556	29146	23592	20547
1997	Male	42458	40727	39750	38440	36854	35214	33830	28546	24858	18837	16435
1997	TOTAL	83 954	81 120	79 235	76 900	74 071	72 161	69 158	60 102	54 004	42 429	36 982
1998	Female	42807	41948	41113	39695	38216	36913	36315	32296	30043	25030	21633
1998	Male	44725	42674	41448	39379	37752	36040	33968	29313	25360	19743	16291
1998	TOTAL	87 532	84 622	82 561	79 074	75 968	72 953	70 283	61 609	55 403	44 773	37 924
1999	Female	38561	43289	42431	40706	41145	37374	36473	33314	30436	24508	22234
1999	Male	39248	45084	43432	40758	41686	36414	34802	29575	25409	19945	16891
1999	TOTAL	77 809	88 373	85 863	81 464	82 831	73 788	71 275	62 889	55 845	44 453	39 125
2000	Female	28964	39110	43641	42725	40830	39318	37503	33583	31786	25118	21425
2000	Male	30635	39274	45351	43193	40271	37480	35912	29888	25911	19747	16685
2000	TOTAL	59 599	78 384	88 992	85 918	81 101	76 798	73 415	63 471	57 697	44 865	38 110
2001	Female	37129	29798	39527	44403	42369	40554	38007	37941	32634	25806	21574
2001	Male	37802	31540	40017	45835	42466	39116	35307	36027	27231	20326	17113
2001	TOTAL	74 931	61 338	79 544	90 238	84 835	79 670	73 314	73 968	59 865	46 132	38 687
PROMOTEES CALCULATED WITH APPLICATION OF TRANSFERS												
2001	Female	36531	29096	38810	43527	41640	39798	37035	36930	31357	25261	21513
2001	Male	37105	30833	39274	44972	41737	38406	34553	35222	26303	19861	17061
2001	TOTAL	73 636	59 929	78 084	88 499	83 377	78 204	71 588	72 152	57 660	45 122	38 574

Dropouts in the Western Cape by year and gender per grade

Table 19 displays the promotions by year and gender between 1995 to 2001 in the Western Cape and has been calculated based on the enrolment, repeaters and promotees as in Tables 15, 16 and 18. For example the drop-outs of grade 1 in 2001 are calculated by subtracting the number of repeaters in grade 1 in 2002 and promotees in grade 1 in 2001 from the enrolment of that grade in 2001. The total dropout for grade 1 in 2001 are the 4081 repeaters of grade 1 in 2002 plus the 74931 promotees of grade 1 in 2001 subtracted from the total enrolment of grade 1 in 2001 of 81790 giving 2777.

Table 19: Dropouts by year and gender (1995 to 2001): Western Cape

YEAR	GENDER	GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11
1995	Female	3 621	771	339	125	735	593	541	2 404	1 958	1 951	215
1995	Male	4 582	1 175	659	558	896	1 264	812	3 217	2 769	2 794	1 990
1995	TOTAL	8 203	1 946	998	683	1 631	1 857	1 353	5 621	4 727	4 745	2 205
1996	Female	4 777	1 323	901	995	1 259	1 281	404	2 724	2 186	3 691	2 589
1996	Male	4 959	872	-269	435	968	1 107	1 368	3 980	2 215	3 312	2 877
1996	TOTAL	9 736	2 195	632	1 430	2 227	2 388	1 772	6 704	4 401	7 003	5 466
1997	Female	644	-1 852	-1 822	-1 565	-1 278	-832	-564	2 415	1 352	2 928	1 754
1997	Male	1 814	-590	-1 310	-97	120	394	606	3 218	2 249	4 029	2 288
1997	TOTAL	2 458	-2 442	-3 132	-1 662	-1 158	-438	42	5 633	3 601	6 957	4 042
1998	Female	2 655	-116	-514	-56	210	369	693	2 571	1 101	3 913	1 505
1998	Male	3 205	269	-413	717	703	867	1 244	3 744	2 820	4 926	2 088
1998	TOTAL	5 860	153	-927	661	913	1 236	1 937	6 315	3 921	8 839	3 593
1999	Female	2 313	118	-471	151	-1 541	600	343	3 001	2 139	5 576	2 682
1999	Male	2 431	378	-749	313	-2 518	1 172	1 018	4 402	3 945	5 455	2 959
1999	TOTAL	4 744	496	-1 220	464	-4 059	1 772	1 361	7 403	6 084	11 031	5 641
2000	Female	724	-146	-24	80	319	2 177	286	3 053	1 215	4 618	2 798
2000	Male	1 090	266	-73	641	1 010	4 587	1 132	4 727	3 227	5 020	2 685
2000	TOTAL	1 814	120	-97	721	1 329	6 764	1 418	7 780	4 442	9 638	5 483
2001	Female	1265	-391	-169	-6	536	385	1247	2511	1841	5717	3351
2001	Male	1512	-82	-339	592	933	1299	2022	3424	3778	5298	2682
2001	TOTAL	2 777	-473	-508	586	1 469	1 684	3 269	5 935	5 619	11 015	6 033
DROPOUTS CALCULATED WITH APPLICATION OF TRANSFERS												
2001	Female	1863	311	548	870	1265	1141	2219	3522	3118	6262	3412
2001	Male	2209	625	404	1455	1662	2009	2776	4229	4706	5763	2734
2001	TOTAL	4 072	936	952	2 325	2 927	3 150	4 995	7 751	7 824	12 025	6 146

5.3.3.2.2. The calculation of the transition rates (repetition, promotion and drop out rates)

The next step of the reconstructed method is the calculation of the transition rates (repetition, promotion and drop out rates) using the raw data as presented in Tables 16, 18 and 19. The different transition rates are obtained by applying the same type of computation.

Repetition Rate in the Western Cape

To calculate repetition rate the repeater data (Table 16) is expressed as a percentage of the enrolment (Table 15). For example to calculate the repetition rate for grade 1 in 2001 use the 4082 repeaters of grade 1 in 2002 and divide it by the enrolment of 81790 for grade 1 in 2001 and express it as a percentage, viz. 5.0% (see Table 20). Figure 23 shows the average repetition rate calculated over a period of 7 years (1995 to 2002). There are more repeaters in grades 8 to 12 with the peak in grade 10. The male repetition is higher in all grades than the female. It is also noticeable that the repetition rate for grade 1 is high.

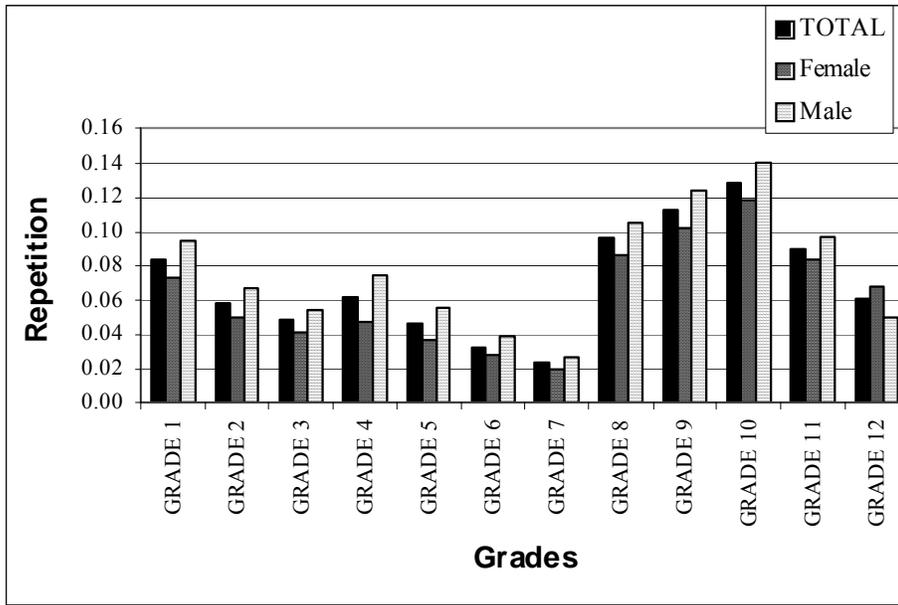


Figure 23: Average repetition rate (1995 to 2001) by gender and grade

Promotion rate

To calculate the promotion rate the promotees' data (Table 18) is expressed as a percentage of the enrolment (Table 15). To calculate the promotion rate for grade 1 in 2001 use the 74931 promotees of grade 1 in 2001 and divide it by the total enrolment of 81790 and times it by 100 as a percentage, viz. 91.6% (Table 20).

Figure 24 shows the average promotions rate calculated over a period of 7 years (1995 to 2002).

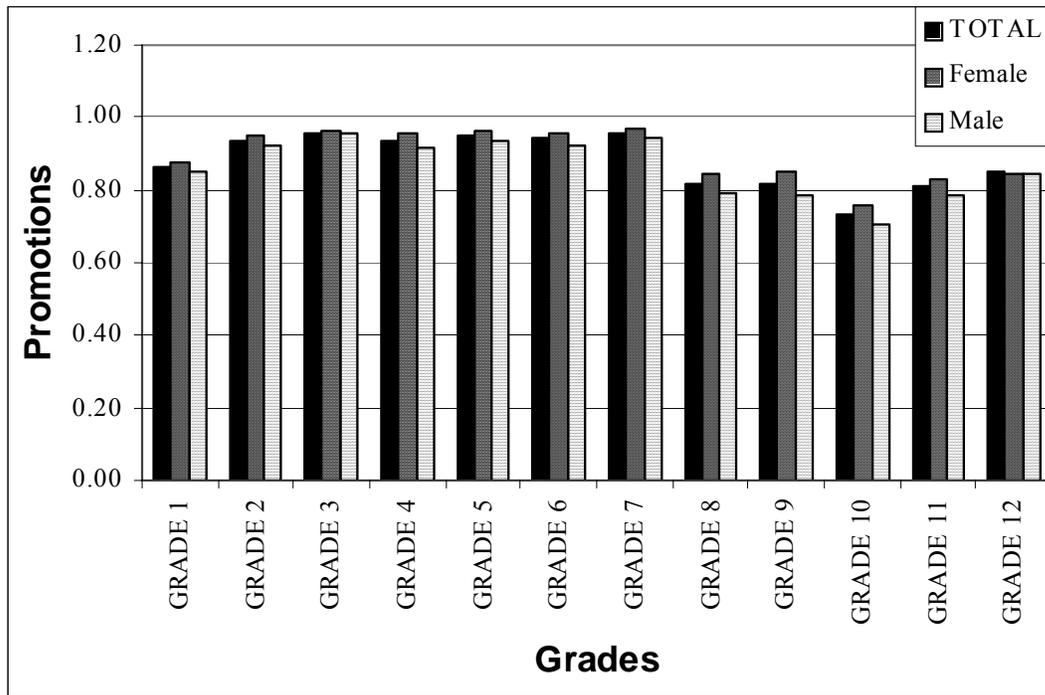


Figure 24: Average promotion rate in the Western Cape by gender and grade (1995 to 2001)

Dropout Rate

The dropout rate expresses dropouts (Table 19) as a percentage of the enrolment (Table 15).

To calculate the dropout rate for grade 1 in 2001 use the 2777 drop out in grade 1 in 2001 and divide it by the total 81790 for grade 1 in 2001 and times it by 100 to express it as a percentage, which results in 3.4% (Table 20).

Again dropout is the highest in the grades 8 to 12 with grade 10 the highest. The dropout rate for male learners is higher than for female learners for all the grades (Figure 25).

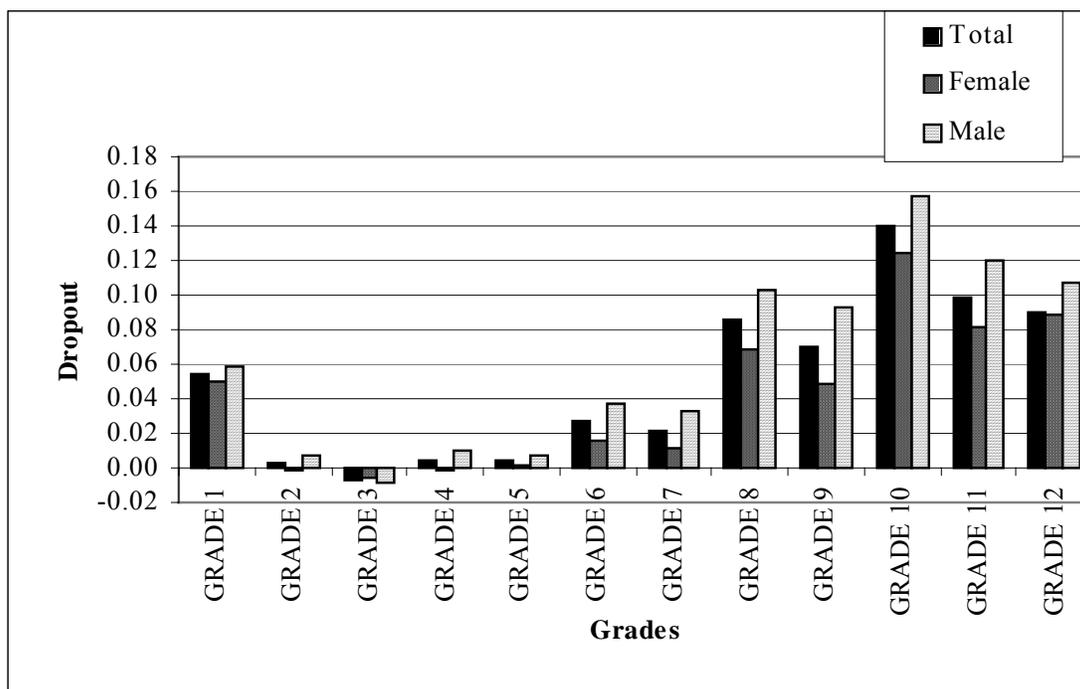


Figure 25: Average dropout rate in the Western Cape by grade and gender (1995 to 2001)

Table 20 is a summary of the calculated flow rates (promotion, dropout and repetition) by year and grade in the Western Cape and will be applied in the Reconstructed Cohort Method. Table 21 displays the calculated flow rates for female learners and Table 22 shows a summary of flow rates for male learners by year and grade in the Western Cape.

Table 20: Transition rates for male and female learners in the Western Cape

Year		GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11	GRADE 12
1995	Promotion rate	0.79	0.90	0.93	0.93	0.92	0.93	0.95	0.82	0.80	0.79	0.86	0.85
1996	Promotion rate	0.78	0.90	0.94	0.92	0.93	0.93	0.95	0.81	0.82	0.77	0.81	0.85
1997	Promotion rate	0.86	0.96	0.98	0.95	0.96	0.97	0.97	0.82	0.83	0.74	0.83	0.85
1998	Promotion rate	0.88	0.94	0.96	0.93	0.94	0.95	0.95	0.80	0.82	0.73	0.82	0.85
1999	Promotion rate	0.89	0.95	0.97	0.93	1.00	0.94	0.95	0.80	0.80	0.70	0.79	0.85
2000	Promotion rate	0.92	0.96	0.96	0.93	0.95	0.90	0.97	0.79	0.82	0.70	0.78	0.85
2001	Promotion rate	0.92	0.97	0.97	0.96	0.95	0.96	0.94	0.90	0.83	0.69	0.77	0.85
	Average	0.86	0.94	0.96	0.93	0.95	0.94	0.96	0.82	0.82	0.73	0.81	0.85
1995	Repetition rate	0.13	0.08	0.06	0.06	0.05	0.04	0.03	0.10	0.13	0.11	0.08	0.04
1996	Repetition rate	0.12	0.08	0.06	0.07	0.05	0.03	0.02	0.10	0.11	0.10	0.06	0.09
1997	Repetition rate	0.12	0.07	0.06	0.07	0.05	0.04	0.02	0.10	0.11	0.13	0.08	0.10
1998	Repetition rate	0.06	0.06	0.05	0.06	0.05	0.03	0.02	0.11	0.12	0.13	0.10	0.06
1999	Repetition rate	0.06	0.04	0.04	0.07	0.05	0.04	0.03	0.11	0.11	0.13	0.09	0.05
2000	Repetition rate	0.05	0.04	0.04	0.06	0.04	0.03	0.01	0.11	0.12	0.15	0.11	0.04
2001	Repetition rate	0.05	0.03	0.03	0.04	0.03	0.02	0.02	0.03	0.09	0.15	0.11	0.04
	Average	0.08	0.06	0.05	0.06	0.05	0.03	0.02	0.10	0.11	0.13	0.09	0.06
1995	Dropout rate	0.08	0.02	0.01	0.01	0.02	0.03	0.02	0.08	0.08	0.09	0.05	0.11
1996	Dropout rate	0.10	0.03	0.01	0.02	0.03	0.03	0.03	0.09	0.07	0.13	0.13	0.06
1997	Dropout rate	0.03	-0.03	-0.04	-0.02	-0.02	-0.01	0.00	0.08	0.06	0.12	0.09	0.05
1998	Dropout rate	0.06	0.00	-0.01	0.01	0.01	0.02	0.03	0.08	0.06	0.14	0.08	0.09
1999	Dropout rate	0.05	0.01	-0.01	0.01	-0.05	0.02	0.02	0.09	0.09	0.17	0.11	0.10
2000	Dropout rate	0.03	0.00	0.00	0.01	0.02	0.08	0.02	0.10	0.06	0.15	0.11	0.11
2001	Dropout rate	0.03	-0.01	-0.01	0.01	0.02	0.02	0.04	0.07	0.08	0.16	0.12	0.11
	Average	0.05	0.00	-0.01	0.00	0.00	0.03	0.02	0.09	0.07	0.14	0.10	0.09
Transfers taken into account in calculating the transition rates													
2001	Prom Transfers	0.90	0.95	0.95	0.94	0.93	0.94	0.92	0.88	0.80	0.67	0.77	0.85
2001	Rep Transfers	0.05	0.03	0.03	0.04	0.03	0.02	0.02	0.03	0.09	0.15	0.11	0.04
2001	Drop Transfers	0.05	0.01	0.01	0.02	0.03	0.04	0.06	0.09	0.11	0.18	0.12	0.11

Table 21: Transition rates for female learners in the Western Cape

YEAR		GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11	GRADE 12
1995	Promotion rate	0.81	0.91	0.94	0.95	0.94	0.94	0.96	0.84	0.82	0.82	0.91	0.80
1996	Promotion rate	0.80	0.90	0.93	0.92	0.93	0.94	0.97	0.84	0.84	0.78	0.84	0.85
1997	Promotion rate	0.88	0.98	1.00	0.98	0.99	0.99	0.99	0.84	0.86	0.78	0.85	0.85
1998	Promotion rate	0.89	0.95	0.97	0.95	0.96	0.96	0.96	0.83	0.86	0.76	0.85	0.85
1999	Promotion rate	0.90	0.96	0.97	0.94	1.00	0.95	0.97	0.83	0.84	0.72	0.81	0.85
2000	Promotion rate	0.93	0.97	0.97	0.95	0.96	0.93	0.98	0.83	0.86	0.73	0.80	0.85
2001	Promotion rate	0.93	0.99	0.98	0.97	0.96	0.97	0.96	0.92	0.87	0.71	0.78	0.85
	Average	0.88	0.95	0.97	0.95	0.96	0.96	0.97	0.85	0.85	0.76	0.83	0.84
1995	Repetition rate	0.11	0.07	0.05	0.05	0.05	0.04	0.03	0.10	0.12	0.11	0.08	0.03
1996	Repetition rate	0.11	0.07	0.05	0.05	0.04	0.03	0.02	0.09	0.09	0.09	0.06	0.11
1997	Repetition rate	0.10	0.06	0.05	0.06	0.04	0.03	0.02	0.09	0.10	0.12	0.07	0.12
1998	Repetition rate	0.05	0.05	0.04	0.05	0.04	0.03	0.02	0.10	0.11	0.12	0.09	0.07
1999	Repetition rate	0.05	0.04	0.04	0.05	0.04	0.03	0.02	0.10	0.10	0.12	0.09	0.05
2000	Repetition rate	0.04	0.03	0.03	0.04	0.03	0.02	0.01	0.09	0.11	0.13	0.10	0.05
2001	Repetition rate	0.04	0.03	0.03	0.03	0.02	0.02	0.01	0.02	0.08	0.13	0.10	0.04
	Average	0.07	0.05	0.04	0.05	0.04	0.03	0.02	0.09	0.10	0.12	0.08	0.07
1995	Dropout rate	0.08	0.02	0.01	0.00	0.02	0.02	0.02	0.07	0.06	0.07	0.01	0.17
1996	Dropout rate	0.10	0.03	0.02	0.03	0.03	0.03	0.01	0.07	0.07	0.13	0.11	0.04
1997	Dropout rate	0.01	-0.05	-0.05	-0.04	-0.03	-0.02	-0.02	0.06	0.04	0.10	0.07	0.03
1998	Dropout rate	0.06	0.00	-0.01	0.00	0.01	0.01	0.02	0.07	0.03	0.12	0.06	0.08
1999	Dropout rate	0.05	0.00	-0.01	0.00	-0.04	0.02	0.01	0.07	0.06	0.16	0.10	0.10
2000	Dropout rate	0.02	0.00	0.00	0.00	0.01	0.05	0.01	0.08	0.03	0.13	0.10	0.10
2001	Dropout rate	0.03	-0.01	0.00	0.00	0.01	0.01	0.03	0.06	0.05	0.16	0.12	0.11
	Average	0.05	0.00	-0.01	0.00	0.00	0.02	0.01	0.07	0.05	0.12	0.08	0.09
Transfers taken into account in calculating the transition rates													
2001	Prom Transfers	0.91	0.96	0.96	0.95	0.95	0.95	0.93	0.89	0.84	0.69	0.77	0.85
2001	Rep Transfers	0.04	0.03	0.03	0.03	0.02	0.02	0.01	0.02	0.08	0.13	0.10	0.04
2001	Drop Transfers	0.05	0.01	0.01	0.02	0.03	0.03	0.06	0.09	0.08	0.17	0.12	0.11

Table 22: Transition rates for male learners in the Western Cape

YEAR		GRADE 1	GRADE 2	GRADE 3	GRADE 4	GRADE 5	GRADE 6	GRADE 7	GRADE 8	GRADE 9	GRADE 10	GRADE 11	GRADE 12
1995	Promotion rate	0.76	0.88	0.92	0.91	0.91	0.92	0.94	0.79	0.77	0.76	0.81	0.80
1996	Promotion rate	0.77	0.89	0.94	0.91	0.92	0.93	0.93	0.78	0.80	0.76	0.79	0.85
1997	Promotion rate	0.83	0.93	0.97	0.91	0.94	0.95	0.95	0.80	0.80	0.70	0.80	0.85
1998	Promotion rate	0.87	0.93	0.96	0.91	0.92	0.94	0.94	0.78	0.78	0.69	0.79	0.85
1999	Promotion rate	0.88	0.94	0.97	0.91	1.00	0.93	0.94	0.76	0.76	0.68	0.77	0.85
2000	Promotion rate	0.91	0.94	0.96	0.92	0.93	0.86	0.95	0.76	0.77	0.67	0.76	0.85
2001	Promotion rate	0.91	0.96	0.97	0.94	0.94	0.94	0.93	0.88	0.79	0.66	0.76	0.85
	Average	0.85	0.93	0.95	0.92	0.94	0.92	0.94	0.79	0.78	0.70	0.78	0.84
1995	Repetition rate	0.15	0.09	0.07	0.08	0.06	0.05	0.03	0.11	0.13	0.12	0.08	0.04
1996	Repetition rate	0.13	0.09	0.06	0.08	0.05	0.04	0.03	0.11	0.12	0.11	0.07	0.07
1997	Repetition rate	0.13	0.08	0.06	0.09	0.06	0.04	0.03	0.11	0.12	0.15	0.09	0.07
1998	Repetition rate	0.07	0.07	0.05	0.08	0.06	0.04	0.03	0.13	0.13	0.14	0.11	0.04
1999	Repetition rate	0.07	0.05	0.05	0.08	0.06	0.04	0.03	0.12	0.12	0.14	0.10	0.04
2000	Repetition rate	0.06	0.05	0.04	0.07	0.05	0.03	0.02	0.12	0.14	0.16	0.12	0.04
2001	Repetition rate	0.06	0.04	0.04	0.05	0.04	0.03	0.02	0.03	0.10	0.16	0.12	0.05
	Average	0.09	0.07	0.05	0.07	0.06	0.04	0.03	0.11	0.12	0.14	0.10	0.05
1995	Dropout rate	0.09	0.03	0.02	0.01	0.02	0.04	0.02	0.10	0.10	0.12	0.11	0.16
1996	Dropout rate	0.10	0.02	-0.01	0.01	0.02	0.03	0.04	0.11	0.07	0.13	0.15	0.08
1997	Dropout rate	0.04	-0.01	-0.03	0.00	0.00	0.01	0.02	0.09	0.07	0.15	0.11	0.08
1998	Dropout rate	0.06	0.01	-0.01	0.02	0.02	0.02	0.03	0.10	0.09	0.17	0.10	0.11
1999	Dropout rate	0.05	0.01	-0.02	0.01	-0.06	0.03	0.03	0.11	0.12	0.19	0.13	0.11
2000	Dropout rate	0.03	0.01	0.00	0.01	0.02	0.11	0.03	0.12	0.10	0.17	0.12	0.11
2001	Dropout rate	0.04	0.00	-0.01	0.01	0.02	0.03	0.05	0.08	0.11	0.17	0.12	0.10
	Average	0.06	0.01	-0.01	0.01	0.01	0.04	0.03	0.10	0.09	0.16	0.12	0.11
Transfers taken into account in calculating the transition rates													
2001	Prom Transfers	0.89	0.94	0.95	0.92	0.92	0.92	0.91	0.86	0.76	0.65	0.76	0.85
2001	Rep Transfers	0.06	0.04	0.04	0.05	0.04	0.03	0.02	0.03	0.10	0.16	0.12	0.05
2001	Drop Transfers	0.05	0.02	0.01	0.03	0.04	0.05	0.07	0.10	0.14	0.19	0.12	0.10

5.3.3.3. The reconstructed cohort method: analysing the flow of learners through the education system using transition rates

The final step of the reconstructed cohort method, also called the Grade Transition Model (GTM) is the application of these transition rates. The reconstruction of a hypothetical flow of learners through the system facilitates the computation of the indicators of internal efficiency. The methodology of the reconstructed cohort flow model is based on the concept that for learners enrolled in a given grade and given year, three possible and mutually exclusive events might have occurred:

- some learners may have been promoted to the next higher grade,
- others may have repeated the same grade they were attending the previous year
- the remaining learners may have dropped out of the system.

(Refer to UNESCO, 1998:46; Wako, 1995:27)

The three key transition rates, namely promotion, repetition and dropout that are used to analyse the flow of learners through the system will be applied in four different ways. Each one of these four scenarios of the reconstructed cohort method will be discussed in the following paragraphs to explain how the different indicators of internal efficiency could be obtained.

5.3.3.3.1. Reconstructing a cohort of learners using the transition rates of a particular year

Figure 26 is a prediction model to reconstruct the flow of a cohort of learners through the education cycle using the transition rates of any particular year (as an example the rates of the year 2001 as applied in Figure 26) as presented in Table 20 with the assumption that:

- At any given grade, the rates of repetition, promotion, and drop-out will be fixed for the duration of the life cycle of the cohort.
- There will be no transfers into the system in any of the subsequent years during the lifetime of the cohort.

Figure 26 is based on these principles, which is a generic model used in other studies (UNESCO, 1998; Thonstad, 1981; Wako, 1995).

This kind of estimation helps in education planning, particularly with regard to the building of schools, staff establishments, etc. To reconstruct the history of learners entering grade 1 in the Western Cape using such a model, it is easier to express this starting cohort as an index of 1000 learners, and all operations are consequently expressed in 'per thousand' terms (cf. UNESCO, 1998:46).

To obtain the promotions, repeaters or dropouts for any particular grade the learners of that grade in any given year are multiplied by the rate according to Table 20.

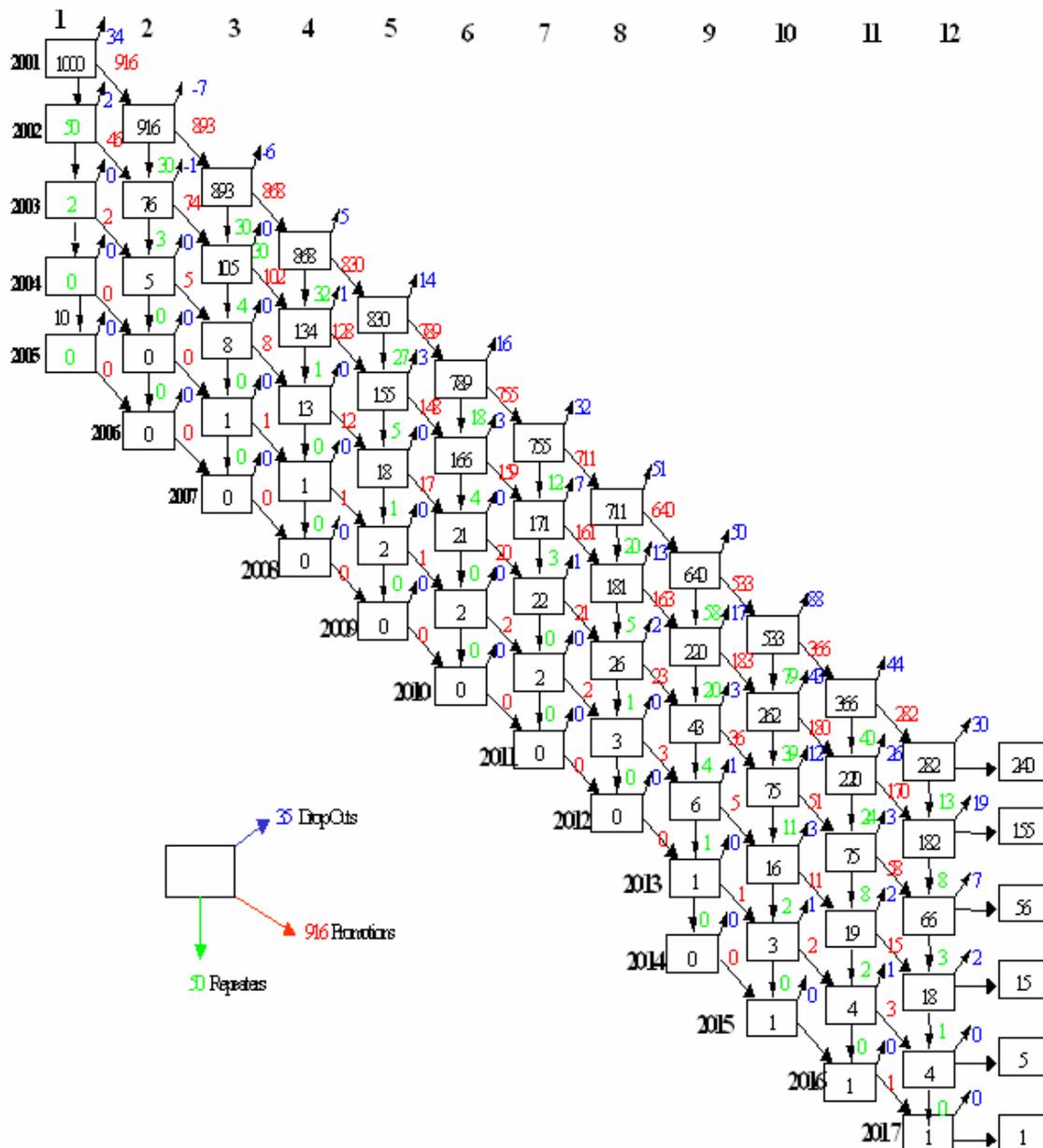


Figure 26: The reconstructed cohort method according to method 1

5.3.3.3.2. Indicators of internal efficiency

Graduates as an indicator of internal efficiency

When applying these flow rates as it occurred between the years 2001 and 2002 the estimated number of learners that will graduate in 2012 without repeating any grade are 240; 155 are estimated to repeat one grade; 56 are estimated to repeat two grades, etc. (cf. Figure 26).

In this model (Figure 26), the flow rates of the 2001 cohort predict that only 471 learners (240 + 155 + 56 + 5) out of the starting 1000 pupils, or 47.1%, will eventually pass Grade 12 in 2012.

Coefficient of efficiency as an indicator of internal efficiency

The coefficient of efficiency per grade in Figure 29 could be calculated in one of the following two ways:

- First: by multiplying the promotees per grade (Figure 26) by the ideal number of learner-years and dividing it by the actual number of learner-years. The ideal number of learner-years is the exact number of learner-years it takes promotees to reach a particular grade. The actual number of learner-years is the total number of years it takes to reach a particular grade. The coefficient of efficiency for a given grade is therefore:

promotees times ideal years to reach particular grade divided by actual number of learner-years it took the promotes to reach the particular grade.

Therefore the coefficient of efficiency for grade 12 is:

$$(240 + 155 + 56 + 15 + 5 + 1) \times 12 \div (1053 + 997 + 1006 + 1015 + 1005 + 978 + 950 + 921 + 911 + 889 + 687 + 553) = (471 \times 12) \div 10965. \text{ The coefficient so obtained is}$$

51.5%. This indicates that internal efficiency is just above half of the optimal level, due to learners repeating grades and dropping out of the system. A coefficient of 100% indicates optimal internal efficiency. Figure 29 presents the coefficient of efficiency per grade in the education system of the Western Cape when the transition rates for 2001 are applied in a fixed manner for the duration of the cycle in the reconstructed cohort method.

- The second method to calculate a coefficient of efficiency is by first calculating the input-output ratio. Firstly, calculate the learner-years per promotee for a particular year. For example learner-years per promotee for grade 12 = $10965 \div 471$. The answer is 23.3. (Note: Learner-years per promotee for given grade = actual number of learner-years it takes promotees to reach a particular grade divide by the number of promotees).

Ideally it takes 12 years to pass grade 12. This cohort took 11.3 years more to pass grade 12. This measure could be done for each grade to determine the grade where the most learner-years per promotee are wasted. The input-output ratio for grade 12 = $23.3/12 = 1.94$. This means that the Western Cape Education system is characterised by an input-output ratio of 1.94. If there were no repeaters or dropouts, this ratio would equal unity. In this system, the grade 12 promotees are being educated at a cost 94 percent higher than the ideal, in terms of resources required over the school cycle to achieve a Grade 12 pass. The ideal is an input-output ratio of unity, everything over one is extra cost (refer to Figure 30).

The coefficient of efficiency is a reciprocal of the input-output ratio. The coefficient of efficiency is therefore also $51.5\% = 1/1.94 * 100$. Both methods of calculating the

coefficient of efficiency yielded the same answer. The closer to 100%, the higher the internal efficiency. This system is least efficient in the higher grades, according to Figure 29.

Learner-years wastage of those not completing the cycle

Wastage in learner-years due to dropout and repetition per grade is calculated as follows:

Firstly, calculate the ideal number of learner-years per grade by multiplying the promotees times the number of years it takes to reach that grade, i.e. grade 4 is $972 \times 4 = 3888$.

Secondly, deduct the actual number of learner-years it took to reach a particular grade from the ideal number of learner-years to calculate the wastage learner-years, i.e. $4072 - 3888 = 185$.

Multiply the dropouts of a given grade by the years it takes to reach that grade to get the wastage learner-years per grade.

The actual wastage for grade 4 dropouts is calculated by adding the learner-years for grade 1 dropouts (36×1), grade 2 dropouts (-7×2), grade 3 dropouts (-6×3) and the grade 4 dropouts (6×4). This results in $(36 + -14 + -18 + 24) = 28$. So the wastage by dropouts up to grade 4 is $28/185 = 15.1\%$. The rest of the wastage is due to repetition. Similarly, when the same formula is applied to grade 12, the actual learner-years are 10965. Ideal learner-years amount to $5650 = 471 \times 12$. Wastage is thus equal to $5316 = 10965 - 5650$. Learner-years wasted due to dropouts = 4738. Percentage due to dropouts is $89.1\% = 4738/5316$. Only 10.9% is therefore due to repetition.

Analysis of the indicators of internal efficiency as derived from the Reconstructed Cohort Method

Based on the above computations Table 23 is a summary of indicators of efficiency for the reconstructed cohort model 1.

Table 23: Summary of all the indicators of internal efficiency according to reconstructed cohort method for model 1

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Promotion rate	91.6%	97.4%	97.2%	95.7%	95.0%	95.6%	94.3%	90.0%	83.2%	68.8%	77.1%	85.0%
Repetition rate	5.0%	3.3%	3.4%	3.7%	3.3%	2.3%	1.5%	2.8%	9.0%	14.7%	10.9%	4.5%
Dropout rate	3.4%	-0.8%	-0.6%	0.6%	1.6%	2.0%	4.2%	7.2%	7.8%	16.4%	12.0%	10.5%
Survival rate	100%	96.4%	97.2%	97.8%	97.2%	95.5%	93.5%	89.5%	82.9%	75.8%	61.2%	52.9%
Pupil-years	1,053	997	1,006	1,015	1,005	978	950	921	911	889	687	553
Coefficient of efficiency	-	-	-	95.5%	94.1%	92.7%	89.5%	83.7%	77.2%	62.9%	55.9%	51.5%
Promotions per 1000	-	-	-	972	955	935	895	829	758	612	529	471
Total learner-years per grade	-	-	-	4072	5076	6055	7005	7926	8837	9726	10412	10965
Ideal Learner-years per promotee				4	5	6	7	8	9	10	11	12
Actual Learner-years per promotee				4.19	5.31	6.47	7.82	9.56	11.66	15.90	19.68	23.29
Input-output ratio				1.05	1.06	1.08	1.12	1.20	1.30	1.59	1.79	1.94
Ideal Learner-years per promotee				3887	4776	5612	6268	6632	6820	6117	5821	5650
Learner-years by dropouts per grade				28	113	233	513	1041	1680	3140	4042	4738
Wasted learner-years				185	301	442	736	1294	2016	3608	4591	5316
Percentage learner-years due to dropouts				15.15%	37.58%	52.69%	69.67%	80.44%	83.32%	87.03%	88.04%	89.14%

- 1) The system appear to be least efficient in grade 10, because the promotion rate is the lowest, the repetition and dropout rates the highest (see also Figure 28). The decrease in coefficient of efficiency rate, between grade 9 and grade 10 is the highest for the entire cycle. This definitely makes micro research and investigation necessary. A possible point of departure could be to address the dropout problem, and more specifically the grade 10 dropout problem, on a micro level, for example at school level, through quantitative as well as qualitative research at the level of individuals in affected schools or communities..
- 2) The calculated input-output ratio at grade 12 is 1.94 (refer to Table 23). The ideal ratio is unity, everything over one is extra cost. If there are no repeaters or dropouts, this ratio would equal unity. In this system, the grade 12 promotees are being educated at a cost 94 % higher than the ideal.
- 3) The actual learner years per promotee indicate that if the education system maintains the current transition rates it would then take over 23 years to attain one successful matriculant. That is 11 years more than the ideal. Dropout and repetition therefore could have major cost implications in any education system.
- 4) It was indicated in an earlier section that the repetition rates could be under-estimated in the education system of the Western Cape (and possibly in SA in general). It was noted in paragraph 5.3.3.2.1 point number 3, that the high dropout in grade 1 could be due to unreported repeaters. This can be a strong determinant of low apparent throughput rates in all the cohort analyses. Therefore, a sensitivity analysis was done to illustrate the influence of the under-estimation of repeaters on the eventual results of the learner flow through. When the repetition rates are doubled for the first two grades, for instance, results changed and affected the overall internal efficiency. In Figure 26 the flow rates of the 2001 cohort predict that only 471 learners (240 + 155 + 56 + 5) out of the starting 1000 pupils, or

47.1%, will eventually pass Grade 12 in 2012. But when the repetition rate is changed to twice what it is reported to be in the first two grades, the 47% changed to 50% and the survival rate of 52% as indicated in Table 23 changed to 56.2% indicating the influence of under-estimation of repeaters on the overall efficiency of the education system. As these are relatively small changes, it appears that the results are not very sensitive to such under-estimation.

- 5) Table 24 is a summary of the survival rates with and without repetition based on the reconstructed Cohort Model 1 as depicted in Figure 26. Survival rate is an indicator of internal efficiency as derived from the Reconstructed Cohort Method 1

Table 24: Survival rates with and without repetition by grade

	Grade1	Grade 3	Grade4	Grade5	Grade6	Grade7	Grade8	Grade9	Grade10	Grade11	Grade12
Total survivals	1000	972	978	972	955	935	895	829	758	612	529
Survival with repetition	0	79	110	141	166	181	184	189	225	245	247
Survival without repetition	1000	893	868	830	789	755	711	640	533	366	282
Dropouts	0	28	22	28	45	65	105	171	242	388	471

- 6) Figure 27 depicts the survival rate by grade and is based on the data in Table 24. It clearly indicates the big drop out of learners between grade 10 and 11.

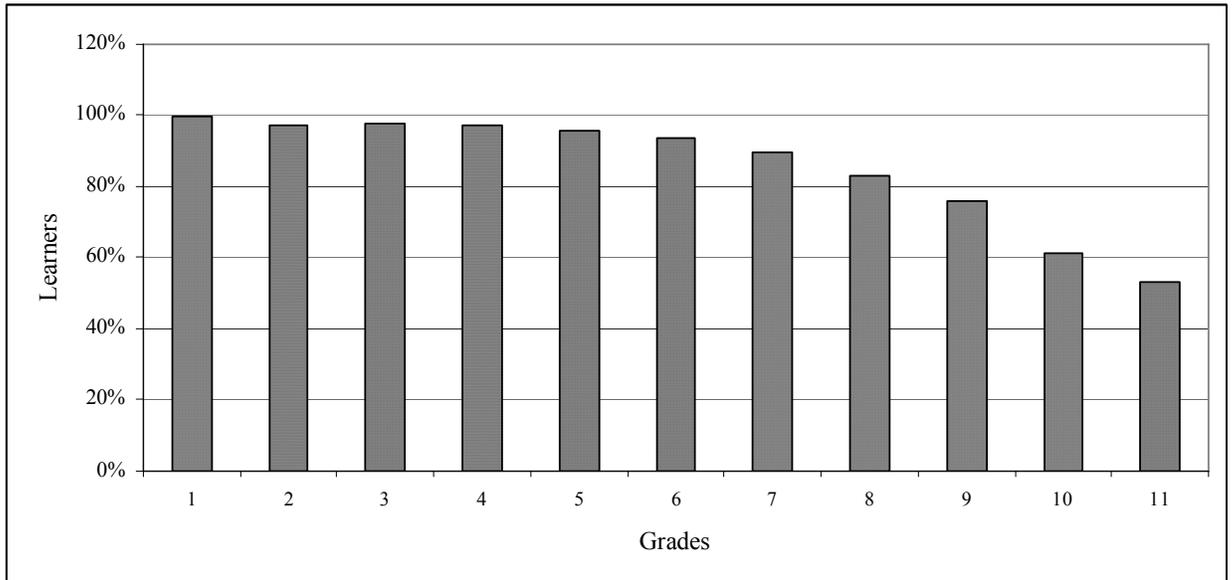


Figure 27: Survival rate per grade, based on the first row of Table 24

- 7) Figure 28 is based on the data in Table 24 and clearly indicates the big drop out of learners in all categories between grade 10 and 11 (according to gradient of the graph).

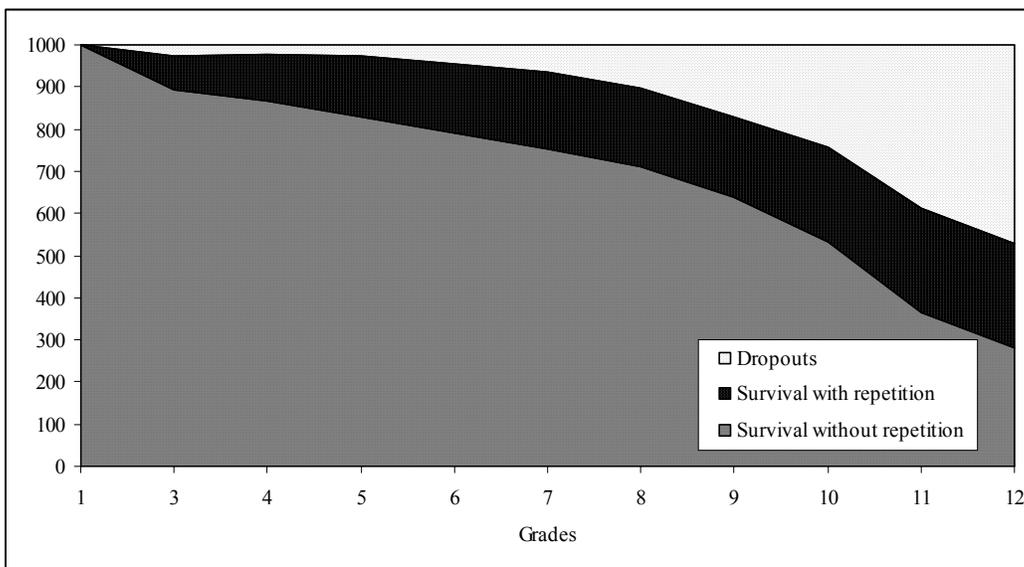


Figure 28: Survival rates and dropouts based on data in Table 24

- 8) Figure 29 presents the coefficient of efficiency per grade in the education system of the Western Cape when the transition rates for 2001 are applied in a fixed manner for the duration of the cycle in the reconstructed cohort method. The education system is the least efficient between the grades 10 and 12.

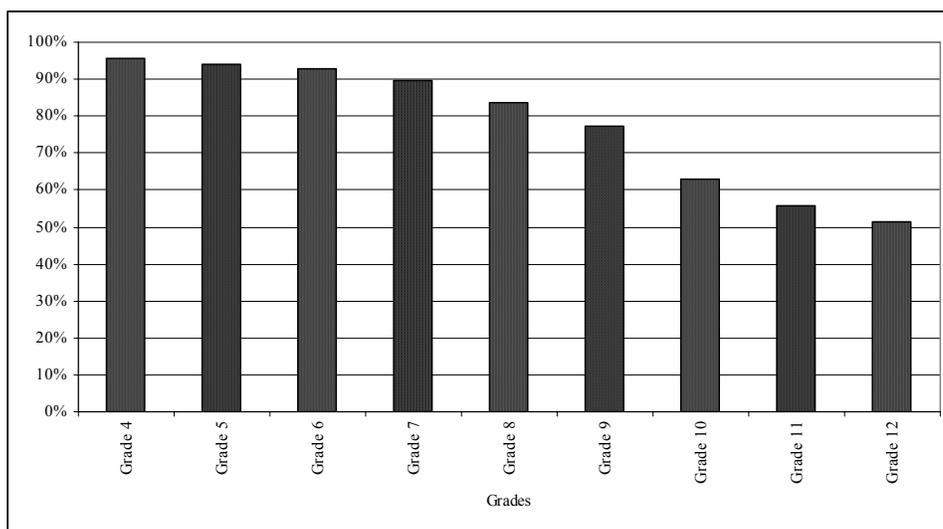


Figure 29: Coefficient of efficiency

- 9) Figure 30 shows the input-output ratio per grade according to Table 23. The ideal is an input-output ratio of unity, everything over one is extra cost.

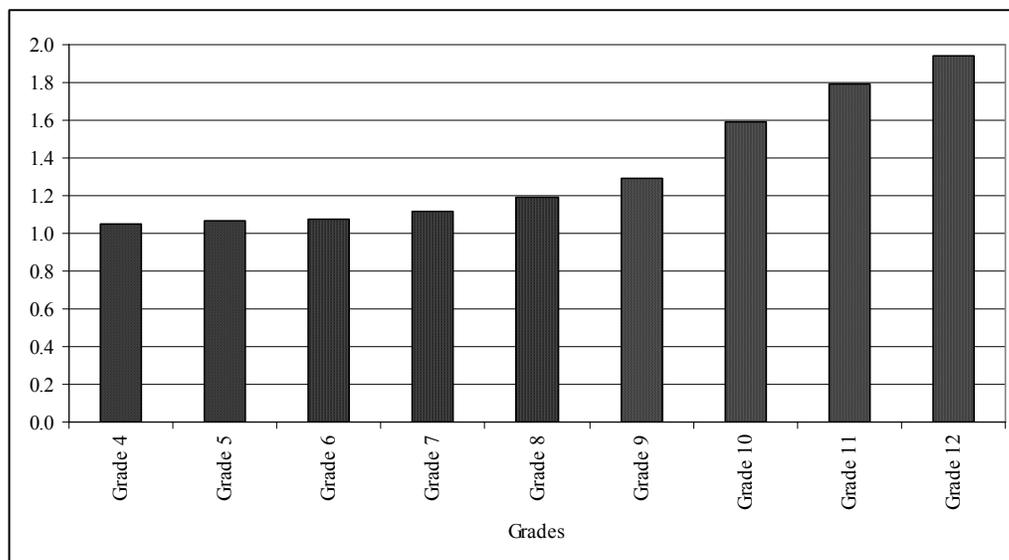


Figure 30: Input-Output Ratio

5.3.3.3.3. Reconstructing a cohort of learners using the transition rates of a particular year and applying the transfers of learners from other systems

The second model is an augmentation of the first model. This model additionally takes into account the transfer of learners into the system throughout the cycle. Table 20 summarises the data on transfers in the Western Cape. Computing flow diagrams for a cohort of learners the education cycle rests on the assumption that:

- the transfer rate of learners into the system is available and that at any given grades, the same transfer rate will apply
- at any given grade, the same rates of repetition, promotion, and drop-out apply
- An implicit additional assumption is that all dropouts of the system are actual dropouts rather than learners transferring out of the system (e.g. by migrating or moving into private schools.)

Based on the above computations the following is a summary of indicators of efficiency for reconstructed cohort method 2 based on Figure 31. In comparison with the previous model it is clear that when the transfer rates are applied the measured efficiency of the system further decreases. Some of those measured in the higher grades are drop-ins from other provinces, systems, etc.

Table 25 displays the indicators of internal efficiency by grade for the Reconstructed Cohort Model 2 based on Figure 31

Table 25: Summary of all the indicators of internal efficiency according to Reconstructed

Cohort Method for model 2

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Promotion rate	90.0%	95.2%	95.4%	93.8%	93.4%	93.9%	92.0%	87.8%	80.1%	67.3%	76.8%	85.0%
Repetition rate	5.0%	3.3%	3.4%	3.7%	3.3%	2.3%	1.5%	2.8%	9.0%	14.7%	10.9%	4.5%
Dropout rate	5.0%	1.5%	1.2%	2.5%	3.3%	3.8%	6.4%	9.4%	10.9%	17.9%	12.2%	10.5%
Survival rate	100%	94.8%	93.3%	92.2%	89.8%	86.8%	83.4%	78.0%	70.4%	62.0%	49.0%	42.2%
Pupil-years	1,053	980	966	957	929	889	847	802	774	727	549	441
Coefficient of efficiency	-	-	-	90.8%	88.8%	86.7%	82.4%	75.9%	68.1%	54.9%	49.0%	45.5%
Promotions per 1000	-	-	-	898.2	867.7	834.1	779.7	704.1	620.0	489.5	422.2	375.6
Total learner-years per grade	-	-	-	3955.9	4884.8	5773.3	6620.4	7422.5	8196.3	8923.5	9473.0	9914.3
Ideal Learner-years per promotee	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0
Actual Learner-years per promotee				4.4	5.6	6.9	8.5	10.5	13.2	18.2	22.4	26.4
Input-output ratio				1.1	1.1	1.2	1.2	1.3	1.5	1.8	2.0	2.2
Ideal Learner-years per promotee				3592.8	4338.6	5004.8	5458.1	5632.7	5579.7	4895.0	4644.4	4507.7
Learner-years by dropouts per grade				211.0	361.0	565.0	943.0	1543.0	1627.0	2927.0	3664.0	4216.0
Wasted learner-years				363.1	546.1	768.5	1162.4	1789.8	2616.6	4028.5	4828.6	5406.6
Percentage learner-years due to dropouts				58%	66%	74%	81%	86%	62%	73%	76%	78%

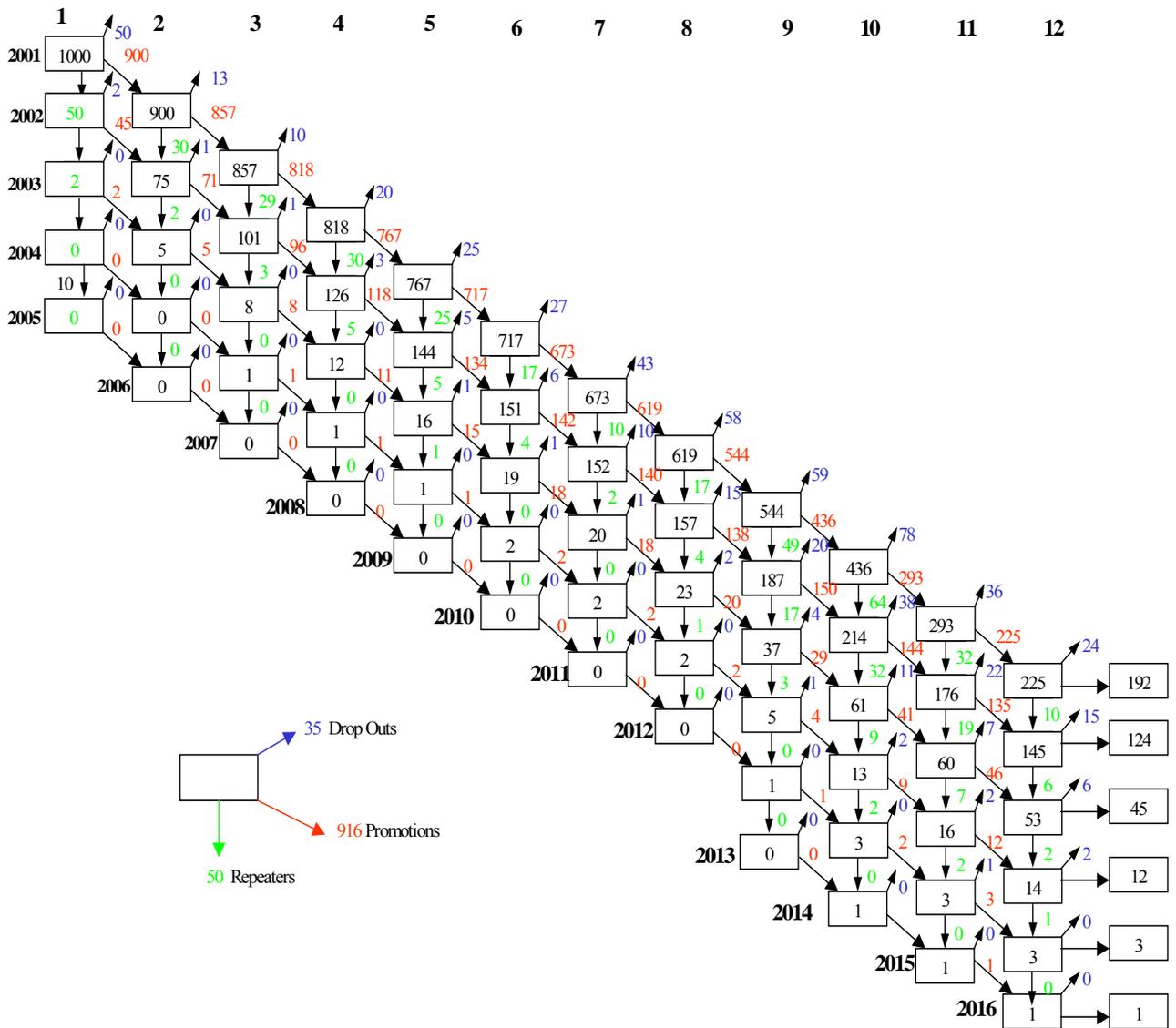


Figure 31: The Reconstructed Cohort Method according to method 2

5.3.3.3.4. Reconstructing a cohort of learners using the average transition rates for the years 1995 to 2001

The third version of the reconstructed cohort model is based on the same principles as models 1 and 2 above. The only difference is that an average flow rate of repetition, promotion and dropouts (Table 20) is used, rather than actual figures for a particular year. The application of average flow rates gives a better indication of the flow through of learners and is less subject to annual fluctuations, because it takes into account the flow rates over a period of time. (On the other hand, it may not fully capture improvements over time.) In this case, the average was calculated for the years 1995 to 2001, which gives an indication of the flow rates expected in the Western Cape. The assumptions remain the same, namely that:

- at any given grade, the same rates of repetition, promotion, and drop-out apply
- there will be no transfers in any of the subsequent years during the life time of the cohort.

Table 26 displays the indicators of internal efficiency by grade for the Reconstructed Cohort Model 3 based on Figure 32

Table 26: Summary of all the indicators of internal efficiency according to Reconstructed Cohort Method for model 3

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Promotion rate	86.2%	93.9%	95.9%	93.4%	94.9%	94.0%	95.6%	81.9%	81.8%	73.2%	81.1%	85.0%
Repetition rate	8.4%	5.8%	4.8%	6.1%	4.6%	3.3%	2.3%	9.5%	11.2%	12.9%	9.0%	6.0%
Dropout rate	5.4%	0.3%	-0.7%	0.5%	0.4%	2.7%	2.2%	8.6%	7.0%	13.9%	9.9%	9.0%
Survival rate	100%	94.1%	93.8%	94.5%	94.0%	93.6%	91.0%	89.0%	80.6%	74.2%	62.3%	55.5%
Pupil-years	1,091	999	985	1,007	986	968	931	983	907	851	685	589
Coefficient of efficiency	-	-	-	92.1%	92.3%	90.4%	89.4%	81.1%	75.4%	64.2%	58.8%	54.9%
Promotions per 1000	-	-	-	940.06	935.94	909.70	889.63	805.54	742.06	623.24	555.26	502.22
Total learner-years per grade	-	-	-	4081	5067	6035	6966	7949	8856	9708	10393	10982
Ideal Learner-years per promotee	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00
Actual Learner-years per promotee				4.34	5.41	6.63	7.83	9.87	11.94	15.58	18.72	21.87
Input-output ratio				1.09	1.08	1.11	1.12	1.23	1.33	1.56	1.70	1.82
Ideal Learner-years per promotee				3760	4679	5458	6227	6444	6678	6232	6107	6026
Learner-years by dropouts per grade				64.00	84.00	240.00	380.00	1052.00	1619.00	2809.00	3557.00	4181.00
Wasted learner-years				321.42	387.71	577.17	738.81	1505.19	2178.42	3476.04	4285.24	4955.39
Percentage learner-years due to dropouts				20%	22%	42%	51%	70%	74%	81%	83%	84%

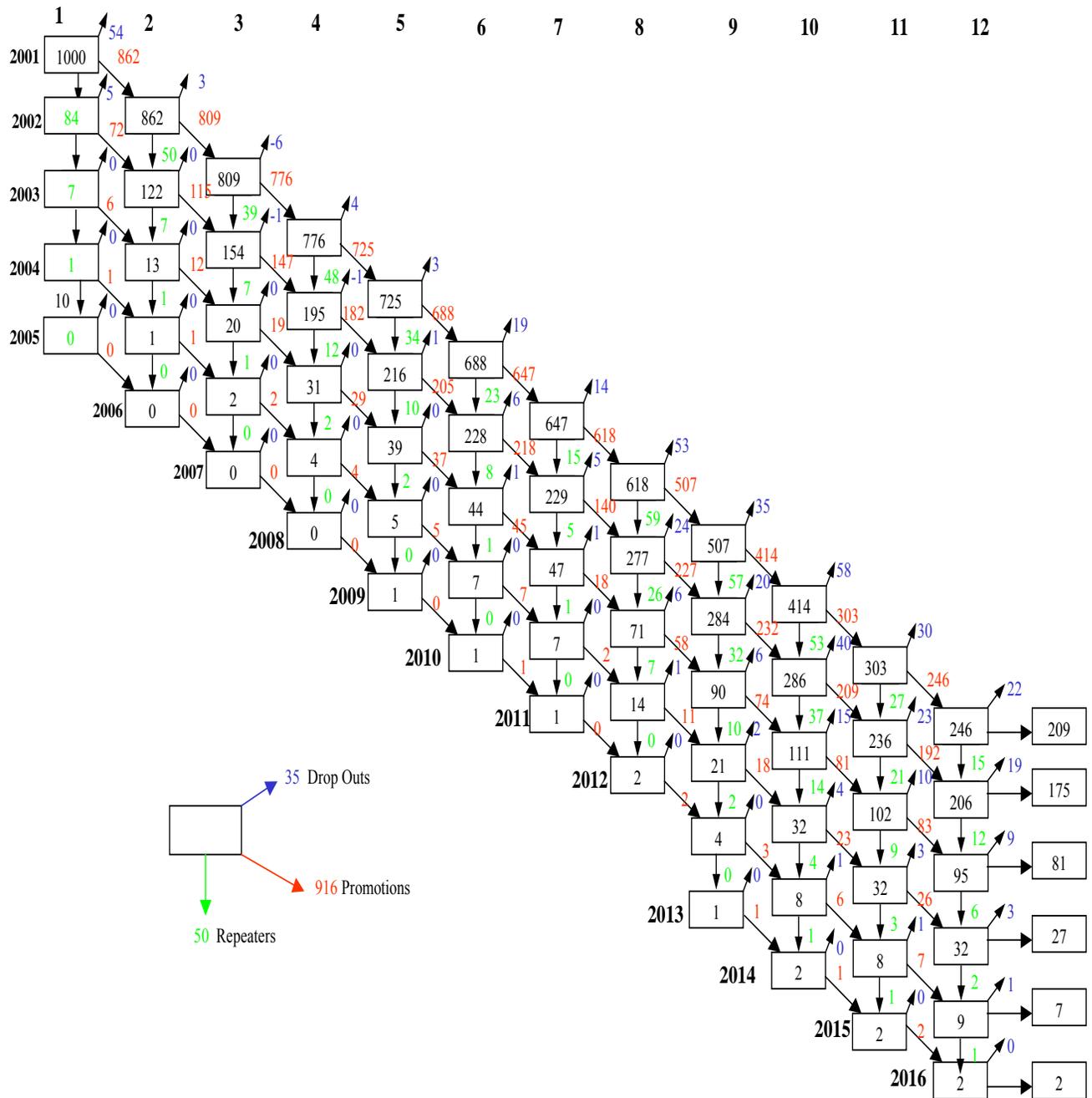


Figure 32: The Reconstructed Cohort Method according to method 3

5.3.3.3.5. Reconstructing a cohort of learners using the actual transition rates for the years 1995 to 2001

The fourth model presented in this dissertation is a flow diagram of a cohort of learners through the secondary cycle using the actual flow rate for a particular grade as it occurred in that specific year. The actual transition rates for the years 1995 to 2001 are available (Table 20) and were used as follows:

The flow rates of 1995 were used for grade 8; the flow rates of 1996 were used for grade 9 and eventually the flow rates of 2001 for grade 12, the end of the cycle. This is a good measure to determine how efficient the system was over a certain period of time. In the South African context it is also a good measure to see how efficient the Western Cape was in the post apartheid South Africa. To further use the model as an analytical (diagnostic) tool the following assumptions apply:

- The actual flow rates per grade are available for the entire cycle.
- A learner can fail only once.
- At any given grade, the same rates of repetition, promotion, and drop-out apply. (For learners repeating a grade, the actual transition rates of the second consecutive year could be used to further increase the realistic and accurate calculation of efficiency indicators).

Based on the above computations the following is a summary of indicators of efficiency for reconstructed cohort method 4. Table 27 displays the indicators of internal efficiency by grade for the Reconstructed Cohort Model 4 based on Figure 33

Table 27: Summary of all the indicators of internal efficiency according to Reconstructed Cohort Method for model 4

	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Promotion rate	82.0%	82.2%	70.0%	77.9%	85.0%
Repetition rate	10.3%	12.0%	12.6%	10.9%	4.5%
Dropout rate	7.7%	5.8%	17.4%	11.2%	10.5%
Survival rate	100.0%	91.5%	85.6%	69.3%	60.9%
Pupil-years	1,103	1,013	941	744	626
Coefficient of efficiency	83%	81%	68%	64%	61%
Promotions per 1000	1000	856	693	609	543
Total learner-years per grade	1103	2116	3058	3802	4427
Ideal learner-years per promotee	1	2	3	4	5
Actual learner-years per promotee	1.10	2.47	4.42	6.24	8.15
Input-output ratio	1.10	1.24	1.47	1.56	1.63
Ideal learner-years per promotee	1000.00	1712.58	2077.55	2436.78	2716.36
Learner-years by dropouts per grade	74.00	146.00	536.00	752.00	1022.00
Wasted learner-years	103.00	403.89	980.14	1364.85	1710.87
Percentage learner-years due to dropouts	71.84%	36.15%	54.69%	55.10%	59.74%

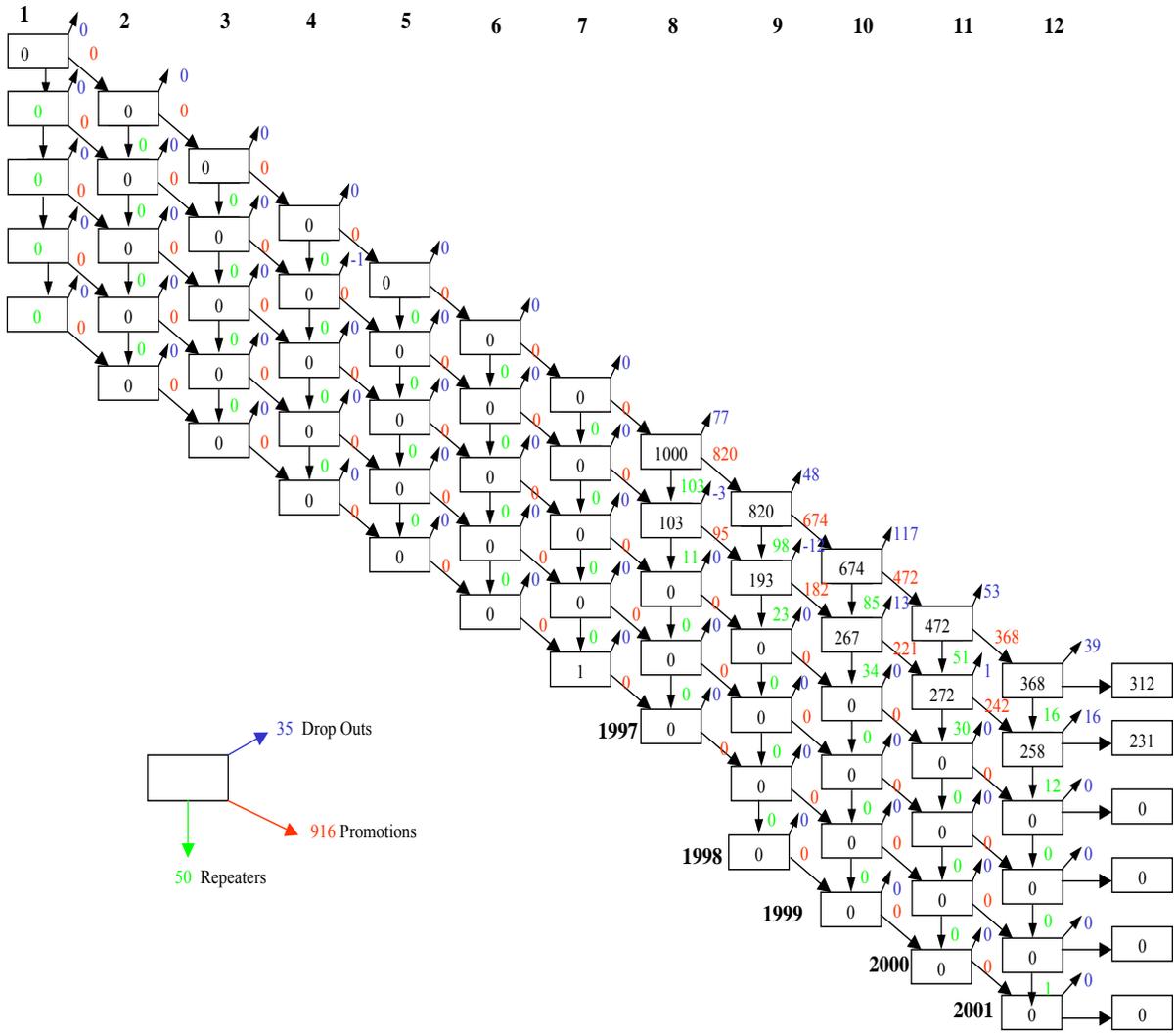


Figure 33: The Reconstructed Cohort Method according to method 4

5.3.4. Concluding remarks

The Reconstructed Cohort Model is an appropriate quantitative technique in any education system to study the flow of learners into, through, and out of the cycle. Student flow models are extremely useful tools for education planners and policy makers (Fredriksen, 1991: 37) suggests that

“student flow analyses are a must for all sector and appraisal reports that recommend or support policies and investments to improve access to, internal efficiency of and output from a cycle of education. It is true that the indicators normally used to assess access, internal efficiency and output also are rather crude, and they need to be supplemented with more qualitative information in order to allow an adequate diagnosis to be established”.

The quantitative analysis is a good point of departure to make sense of the throughput of learners and the functioning of a cycle of education. The education authorities could use the estimations of the model to do the necessary interventions and remedy the situations.

CHAPTER SIX

SUMMARY, FINDINGS AND IMPLICATIONS FOR FURTHER RESEARCH

The study covered three major aspects namely,

- a) The development of an ISD model based on the extant literature.
- b) An application of the ISD model using the Practice-in-Action methodology to illustrate how bricolage, improvisation and sensemaking could facilitate the development of EMIS.
- c) An illustration of sensemaking through the application of quantitative methods on data obtained through the ISD model.

Below is a brief summary of each of these major aspects as listed above.

6.1 The development of an ISD-model

The literature review on information systems development revealed how approaches changed from the traditional (systems development life cycle) based on rational positivist assumptions, to systems thinking and emergent design approaches that accommodate epistemological shifts. Through the different approaches in the subject literature, a set of epistemological assumptions was defined pertaining to knowledge and how to acquire it in order to build a conceptual framework for ISD. A model based on the assumptions of the general systems theory (input, transformation processing, feedback and control), and a set of guiding principles based on emergent approaches such as improvisation, sensemaking and bricolage was proposed.

This explorative research indicated that although information systems development takes place within a normative framework, it allows for systems thinking and showed that beyond the

rational and quantifiable world there is a social reality that needs to be captured. The proposed model endeavours to indicate that there is more to the rational approach in the development of information systems than a linear step by step approach and the stored data in the database. Beyond the numbers and quantifiable world a complex reality exists that, amongst others things, includes the atmosphere, culture and structure of the organization. In addition to this the behaviour, emotions, knowledge and experiences of all the people, who in one way or another interact with the information system, need to be considered. The continuous or cyclical feed-back links provide for continuous sense-making in complex existing information systems.

6.2 Using the ISD model in a practical and real situation based on a Practice-in-Action methodology

The ISD model informed construction of the information systems development during the ten years 1996 to 2005 in the Western Cape in which the author was narrowly involved as a developer and a manager of the ISD team. Developing a system that would make information available for the management of the education for planning, monitoring and decision-making was a driving motive for this study. To provide the most complete, relevant, accurate and timely information that is accessible to the entire organization poses tremendous challenges in an education systems context. Given the complexities of the changing technology and the publicly contested education policy environments, more demands are put on the information systems developers to shorten the development cycle and make quality information available in a shorter time. Enhancing the traditional formal systems approach by using local design

improvisations and sensemaking, the application development time and cycle were shortened considerably.

This study described the application of the ISD model on the EMIS of the Western Cape Education Department through the Practice-in-Action approach. It explored how a web-based information system could be developed through the work practices and experiences of key role players and participants.

6.3 An illustration of sensemaking through the application of quantitative methods to data obtained through the ISD model

Data analysis forms an integral part of the ISD process and to illustrate the possibilities thereof, quantitative methods were applied on data sets of EMIS and other information systems to complete the cycle. Data analyses cannot be conducted prior to or without the establishment of ISD. ISD and data analysis are complementary processes, the one without the other would not be practicable.

Given this and the context of the particular situation, the study utilised a bricolage design (using existing data sources) in order to make sense (analysing the data sets) of the education system. This supported the contribution of these quantitative methods towards making sense of certain dimensions of the education system and enhanced the value of these data sets to inform decision-making and planning. In order to back up some of these claims the research presented a number of aspects of the education situation that were obtained through the application of the quantitative methods that otherwise would have remain hidden. It must be

pointed out that it was not the purpose of the study to develop a model to ‘represent’ the education system and its complex social interrelatedness. Bricolage and sensemaking were used to enable the application of existing approaches (as reviewed in the relevant chapters).

From the **education production function** analyses, the evidence indicated that average qualifications of teachers (REQV), dropout rate per school (dropout rate), percentage of learners who repeated grade 11 (repeaters), poverty of the community (poverty) and the number of teachers paid by the school (SGB teachers) significantly influenced the grade 12 results and were important in predicting student outcomes. Limitations of quantitative approaches taken into consideration, these estimates revealed a good deal about the variables that affected grade 12 results in the Western Cape Education Department, although some of these variables could probably be proxies for other factors, such as school management. A central theme that emerged from this research was that the application of the education production function in order to elicit knowledge from raw data sets was equally as important as the actual results. The education production function therefore has important educational and policy formulation implications, since these kinds of quantitative analyses are important initial steps that alert decision-makers and planners on focus areas and could provide a basis for further qualitative research and investigation. The education production function results aid the identification of smaller groups, such as individual schools and learners for a more detailed study and sensemaking processes.

Although these estimates tell us a good deal about the variables that influence grade 12 outcomes, it does not reveal any detail about behaviour at the classroom and the school level that influenced the quality of the education. For example, the reasons why SGB teachers

significantly affect student outcomes in the Western Cape or why certain schools with low socio-economic status perform better than others cannot be explicated. Knowledge about the importance of these two factors (poverty and SGB teachers) with regard to SGB teachers in the determination of student outcomes remains tentative, and point to the limitations of quantitative analysis of large datasets of this nature.

Similarly, issues such as teachers who have the same numerical REQV but appear to have a better quality of training, or the assumption that former CED schools better utilise well-qualified teachers, as indicated by the education production function analysis, are areas for future research. This is indicative of how the present data sets need to be supplemented by qualitative research.

The final part of the education production function analysis compared actual scores with predicted scores. The education production function results were used to predict how well each school could perform. Using production function estimates in this way identified over-achieving and under-achieving schools, information that would have remain hidden if a simple average were used. In particular, it indicated that some poor schools are over-achievers, despite their socio-economic background and resources. Once such schools have been identified, there is a need to qualitatively analyse the functioning of such schools to identify what makes them excel despite adversity. This could contribute to the sensemaking process and facilitating the establishing of best practices in education.

The education production function therefore is not only important for the results that it produces, but the purpose is to utilize it in such a way that it reveals knowledge hidden in raw

data for decision making. By utilizing such a statistical technique, more is revealed than what the actual numbers tell us. The research was not particularly interested in the individual estimated parameters, but rather in illustrating how to produce data for analysis and then using the elicited knowledge as a guiding framework for further research and analysis.

This study contributes to the Department of Education's quest for information and the analysis thereof by using the education production functions in the education system. According to the Department of Education (2003: 104) such quantitative methods can be used as

“research input-output trends in South African schools (as part of the research into production functions) and in other, similar schooling systems in order to arrive at normative scores that can be used to gauge the performance success of schools with varying levels of resourcing, and varying levels of socio-economic disadvantage.”

This study could contribute to this objective of researching input-output trends in South Africa in order to gauge the performance of schools.

The study could not find any data sources about school management or learner ability and with this limited data sources the research used whatever was available (bricolage approach) to make sense of what influenced learner performance in order to inform education decision-making. Future research should increase our understanding of the school management and learner ability that could influence student outcomes.

Future research options regarding educational production functions include using the same model specification and applying it on data sets of other years (2003, 2004, 2005 and 2006).

This would provide fuller information as to the quality of the data and the strength of the

underlying relationships. It is envisaged that this research model will provide significant and consistent information and knowledge for strategic decision-making and policy making.

A further quantitative technique consisted of constructing an **educational attainment profile** for the Western Cape, based on the Census 2001 data according to the methodology of Filmer and Pritchett (1998a). This educational attainment model can be used on future census data sets to indicate the extent of the dropout of learners from the education system. Such a model could make sense of the data and be applied as a diagnostic tool to alert education management about areas of concern at an early stage. Again not only the results of the model are important, but the application of the model facilitated the revelation of knowledge hidden in large data sets.

The research used the **Reconstructed Cohort Model or Grade Transition Model (GTM)** as a quantitative method. It was applied on data sets that were derived from EMIS as discussed in this study. The Grade Transition Model was used to project the flow of learners through the education system. The research showed how average transitions rates could be used to estimate the dropout rates and to determine the efficiency of the education system. The application of the Reconstructed Cohort Model resulted in only estimations of the throughput of learners based on the transitions rates. Further research and investigation are required to establish if it holds true against reality (sensemaking perspective).

The application of the Reconstructed Cohort Model made possible certain indicators of efficiency, which was an important sensemaking exercise. Amongst other things, it revealed that there was a high dropout in the secondary school phase. It further indicated that the education system was least efficient in grade 10, because the promotion rate is the lowest, the

repetition and dropout rates the highest. The decrease in coefficient of efficiency rate, between grade 9 and grade 10 is the highest for the entire cycle. This would be indicative of the need for additional research and investigation to understand what leads to such dropouts and repetition. The utilisation of such a model is as important as the actual results because it remains an important diagnostic tool to inform authorities about the efficiency of the education system and the areas of concern. The application of the Grade Transition Model facilitated the discovery of knowledge that can be applied for decision-making, policy formulation and future qualitative research. Indicators that otherwise would have been invisible (hidden) have now been elicited. The application of such a model has important educational implications, since it is a quantitative analysis that leads to a better understanding of the patterns and relationships within the data sets.

The GTM is a simple mathematical model that could be used as a powerful diagnostic tool to recommend or support policies to improve education system. However, a main weakness of the model is that it is too reliant on a single indicator (repetition rate), a variable that is not often used for any other management decision. If the quality of the data set were not good, then all the other indicators derived from the model would not be accurate and could produce incorrect (even wrong) projections. From a sense making perspective it is important that in the data gathering and data processing phases, active role players in these processes understand the important function of the data. The reconstructed cohort model quantitatively determines the flow of learners into, through, and out of the education cycle. It calculates the dropout rate, repetition rate and projects the number of graduates at the end of the cycle. However, the model is not able to comment on the quality of the teaching or the learner dynamics (movement) inside the education cycle. To make sense of the education situation, quantitative

estimates must be regarded as a starting point for further qualitative research and micro investigation of smaller groups or focus areas (even at school and learner level).

6.4 The EMIS Process Model

The study culminated in the development of a four-phase EMIS process model that is cyclical, interactive and integrated. Figure 34 graphically depicts the EMIS process model that was developed and provides a synopsis of the phases, processes and activities of EMIS as described in the study. The four phases of EMIS development, when using the survey method, namely Systems Analysis and Design, Application Development, Information Systems Data Processing, and Data Analysis and Reporting are illustrated. The diagram indicates the different activities involved in developing an education management information system. This study has illustrated that EMIS is more than just a technological solution, but it involves activities, processes, procedures and the key role players to manage such a system. The activity diagram attempts to establish a model that indicates the flow of processes in EMIS as conveyed by Table 2.

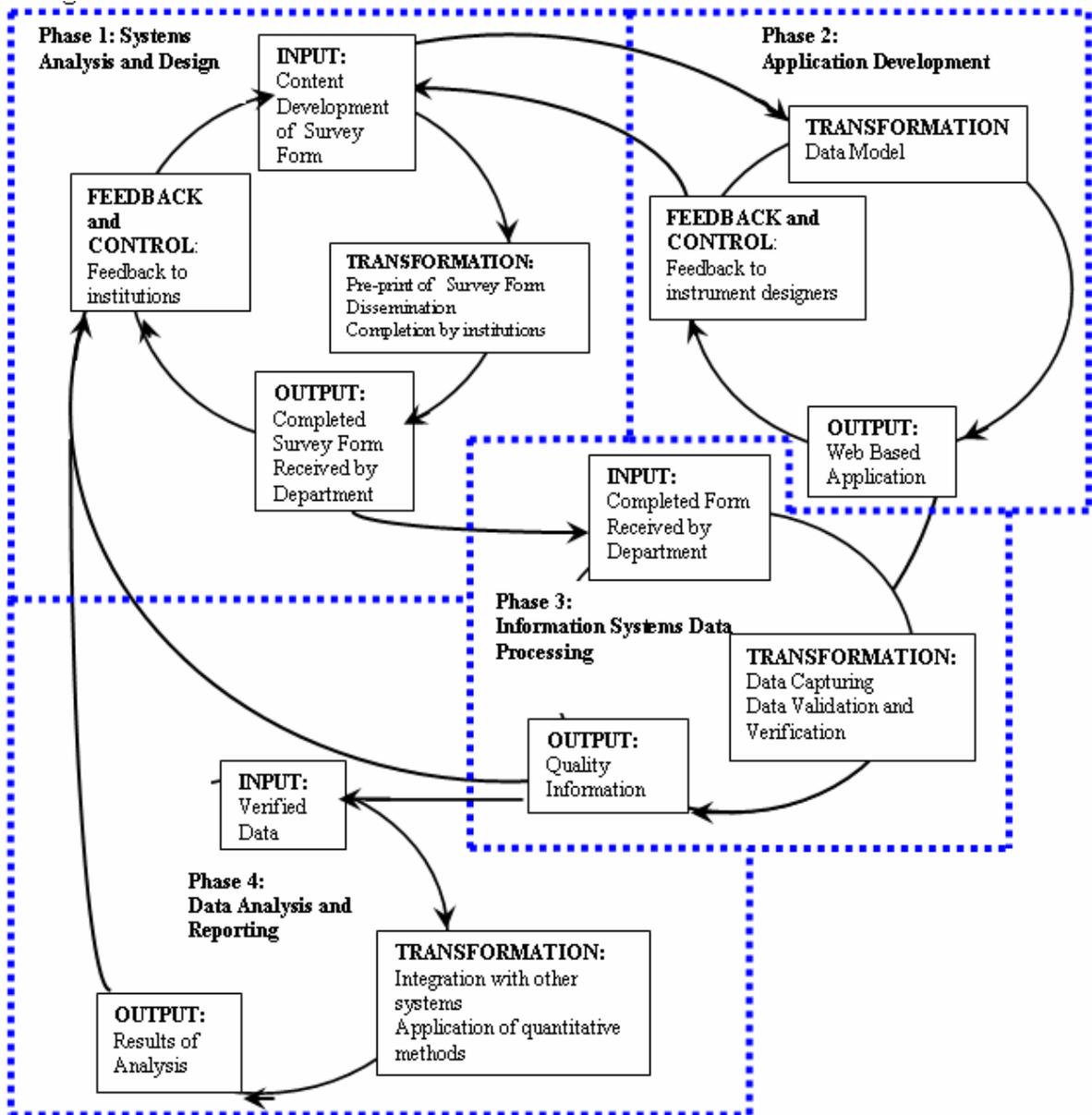


Figure 34: EMIS Process model as a Research Outcomes

Figure 34 indicates the ongoing criterion of sensemaking through the arrows in the diagram. The diagram indicates how, through the systems concept, the specific input elements are transformed through the specified processes into particular outputs. However, the output in one phase becomes the input in the next phase. EMIS in this research is a system where data collection (questionnaire development, web-based computer application development, data capturing and data verification) and data analysis processes work in a cycle to produce the information (datasets) according to the needs of the various role players. The diagram may also be viewed as an elaboration of the input-transformation-output systems concept.

The first three phases of the model focused on developmental processes of information systems while the last phase concentrated on the data analysis practices. The developmental processes based on sensemaking and bricolage approaches produced the EMIS data that serves as essential building blocks for making sense of an education system. However, the EMIS data sets developed through this model are not the only data sources available. The need for EMIS to be integrated with other data sources such as PERSAL, Examinations and Census for a more comprehensive and inclusive information system constituted an important bricolage design exercise.

A Practice-in-Action approach was followed in this study in order to understand how information systems development is located within the context of the education environment. This approach that espouses sensemaking, improvisation and bricolage design principles developed in-house as an alternative to the traditional approaches to ISD. The aim of this research was similar to that described by Heiskanen and Newman (1997:121) namely, “to report on and explore the possibilities of how a practicing information systems professional is

able to describe and analyze her or his experiences in such a way that it makes sense to the professional and scholar communities, thus bridging the gap between two camps”. The study therefore contends that the experiences and the activities (Practice-in-Action) outlined in this study could be utilised to contribute to the formulation of ISD theories.

Linked to the aim of such an approach, namely to investigate the possibilities of translating work experience and practice to inform practitioners about the development of IS, the proposed process model could serve as a roadmap for EMIS development. It would not only be useful for the initial design and development of an information system but could inform the refinement and continuous review of historically embedded information systems as it exists in management systems of education systems.

6.5 Concluding remarks about quantitative analysis: An Exploration from a Sensemaking Perspective

The development of an EMIS together with the application of quantitative methods is not only a technical exercise. Information systems development in an organisation, in general and data analysis in particular, must not be perceived as a process-based and numbers-driven exercise only. It ought to provide quality information that is as complete, relevant, accurate, timely and accessible as possible for improving the quality of decision-making in the organisation. The policy-makers, planners and educators need the processed (produced) datasets to “make sense” of education.

The principles of sensemaking and bricolage that underpin the pillars of ISD in this research were employed to counter the technical emphasis placed on statistical analysis. This study

supports Gorard's (2004:1) view on the fallacy "among social scientists that statistical analysis is somehow a technical, essentially objective, process of decision-making". It endeavoured to foreground the dual nature of statistical analysis emphasising both the important role that quantitative analysis plays on the one hand, and the need for further qualitative research on the other hand. Both these processes are essentially a part of sensemaking with bricolage a central element through the use of whatever data and models can be found to assist in this process.

Data analysis starts from a quantitative foundation to elicit knowledge that is not so apparent. Unfortunately, data does not announce its meaning or significance. It is through data analysis that sensemaking aspects become visible in data. The data analysis practices reveal through quantitative techniques, the not so obvious that could be further researched through smaller samples and qualitative studies. Gorard et al. (2002:5) espouse the notion that "it is actually impossible to conduct purely 'qualitative' or purely 'quantitative' work. Both approaches rely heavily on each other". The research has shown how the application of quantitative methods and statistical analysis could be used to generate insights into organisational sensemaking in this case an education system. Using numbers, formulas and statistical analysis seems to be the conventional approach for managers to solving a problem. . However, such an approach is not a panacea because beyond the numbers and the quantifiable world is a reality that includes culture, emotions, behaviour and experience. It needs to be supplemented with other information to enhance understanding of the functioning of the whole system and the behaviour of actors within it. This research presents important ways for understanding the role of sensemaking in information systems development and data analysis practices. Sensemaking through data analyses give insights, often even to those who do not know that they do not know. It thus makes visible what otherwise may remain invisible.

7 REFERENCES

- Abagi, O. & Odipo, G. 1997. Efficiency of primary education in Kenya: Situational analysis and implications for educational reform. Discussion paper No. DP 004/97, Institute of Policy Analysis and Research, Nairobi, Kenya, September 1997. Available from <http://www.ipar.or.ke/DP4.pdf>.
- Ahituv, N. & Neumann, S. 1984. A flexible approach to information systems development. *MIS Quaterly*, 69-78.
- Ahituv, N. & Neumann, S. 1986. *Principles of information systems for management*. Dubuque, Iowa: Wm. C. Brown Publishers.
- Ahuja, V. & Filmer, D. 1996. Educational attainment in developing countries: New estimates and projections disaggregated by gender. *Journal of Educational Planning and Administration*, Vol. X (3): 229-254.
- Alatalo, T., Oinas-Kukkonen, H., Kurkela, V. & Siponen, M. 2002. Information systems development in emergent organisations. Available from <http://www.owla.oulu.fi/publications/isd2002/isd-owla-final.pdf>.
- Allison, P.D. 1999. *Multiple regression*. Thousand Oaks, CA: Pine Forge Press.
- Archer, S. 1995. Quality and efficiency indicators in education: A survey research report prepared for EDUPOL. Johannesburg: EDUPOL
- Arends, F. 2006. Methodological issues in measuring learner flow-through in the education system, in marking matric. *Colloquium Proceedings*, chapter 8. Cape Town, South Africa: HSRC Press. Available from <http://www.hsrcpress.ac.za>.

- Avgerou, C. 2001. The significance of context in information systems and organizational change. *Information Systems Journal*, 11: 43-63.
- Avison, D., Baskerville, R. & Myers, M. 2001. Controlling action research projects. *Information Technology & People*, Vol. 14(1): 28-45.
- Avison, D., Lau, F., Myers, M. & Nielsen, P. 1999. Action research. *Communications of the ACM Transactions on Information Systems*, Vol. 42(1): 94-97.
- Bansler, J.P. & Bodker, K. 1993. A reappraisal of structured analysis: Design in an organizational context. *ACM Transactions on Information Systems*, 11(2): 165-193.
- Bansler, J.P. & Havn, E. 2004. Improvisation in information systems development. In Kaplan, B., Truex, D., Wastell, D., Wood-Harper, A. & DeGross, J. (Eds.). *Information systems research. Relevant theory and informed practice*. Manchester, UK.: Kluwer Academic Publishers. Proceedings of IFIP TC8/WG8. July 15-17, 2004. Available from http://is.lse.ac.uk/Support/ifip_wg82/Manchester/Bansler.pdf.
- Barrett, F.J. 1998. Creativity and improvisation in jazz and organisations: Implications for organisational learning. *Organisation Science*, 9(5): 605-622.
- Barro, R. & Jong-Wha, L. 1993. International comparisons of educational attainment. *Journal of Monetary Economics*, 32: 363-394. Available from <http://nber.org/papers/w4349.pdf>.
- Barry, C. (n.d.). Web-based information systems – Time for the revisionists. Available from <http://www.is.nuigalway.ie/collector/papers%5Cbarry.pdf#search=%22web%20based%20information%20revisionists%22>.

- Barry, C. 2004. A Conceptual Framework for Information Systems Development from a Decision-Making Perspective. In Vasilecas, O. et al. (Eds.), *Proceedings of 13th International Conference on Information Systems Development*, Vilnius, Lithuania, September 9-11, pp. 170-181.
- Barry, C. & Brown, J. 2003. Finding a home for web-based information systems – Perusing the landscape. In *Proceedings of 11th European Conference on Information Systems (ECIS)*, Naples, Italy, June 16-21. Available from <http://csrc.lse.ac.uk/asp/aspecis/20030008.pdf>
- Baskerville, R. 1999. Investigating information systems with action research. *Communications of the Association for Information Systems*, Vol. 2 (19). Available from http://www.cis.gsu.edu/~rbaskerv/CAIS_2_19/CAIS_2_19.html.
- Baskerville, R. & Wood-Harper, A. 1996. A critical perspective on action research as a method for information systems research. *Journal of Information Technology*, 11: 235–246.
- Bell, S. & Wood-Harper, T. 2003. How to set up information systems: A non-specialist's guide to the multiview approach. London: Earthscan Publications Ltd.
- Benyon, D. 1990. *Information and data modeling.*: Oxford: Blackwell Scientific Publications.
- Beynon-Davies, P. & Williams, M. 2003. The diffusion of information systems development methods. *Journal of Strategic Information Systems*, 12: 29–46.
- Bødker, K. & Carstensen, P. 2004. Development and use of web-based information systems. *Scandinavian Journal of Information Systems*, 16: 3-10.

- Bohte, J., Polinard, J.L., Meier, K.J. & Wrinkle, R.D. 1999. Latino student improvements on the TAAS exam. A report of the Texas Educational Excellence Project. Texas A & M University, University of Texas-Pan American. Available from <http://teep.tamu.edu/reports/report006.pdf>.
- Bravo, D., Contreras, D. & Sanhueza, C. 1999. Educational achievement, Inequalities and private/public gap: Chile 1982-1997. Department of Economics Universidad de Chile. July 1999. Available from http://www.econ.uchile.cl/academic/papers/politicas_sociales/educacion_cs.pdf.
- Brody, M. 2006. *What research says to the practitioner. Action research defining paths towards empowerment and change*. Montana State University. Available from http://naaee.org/conferences/biloxi/m_brody_3_10159a.pdf.
- Büscher, M., Gill, S., Mogensen, P. & Shapiro, D. 2001. Landscapes of practice: Bricolage as method for situated design. *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, Vol. 10(1): 1-28.
- Byrnes, D.A, 1989. Attitudes of Students Parents and Educators toward repeating a Grade. In Shepard, L.A. & Smith, M.L. (Eds.). *Flunking grades: Research and policies on retention*. UK: Falmer Press.
- CACM (Communications of the ACM). 1998. *Special issue on Web-Based Information Systems*, Vol. 41: 7.
- Carrizo, L., Sauvageot, C. & Bella, N. 2003. *Information tools for the preparation and monitoring of education plans*. Paris, France: UNESCO. Available from <http://unesdoc.unesco.org/images/0013/001323/132306e.pdf>.

- Carstensen, P.H. & Vogelsang, L. 2001. Design of web-based information systems – New challenges for systems development? Global co-operation in the new millenium, *The 9th European Conference on Information Systems*, Bled, Slovenia, June 2001. Available from <http://csrc.lse.ac.uk/asp/aspecis/20010017.pdf#search=%22Design%20of%20web-based%20information%20systems%20-%20New%20challenges%20pdf%22>.
- Case, A. & Deaton, A. 1999. School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics*, CXIV(3):1047-1084.
- Case, A. & Yogo, M. 1999. Does school quality matter? Returns to education and the characteristics of schools in South Africa. *NBER Working Paper W 7399*. Cambridge, Mass.: National Bureau of Economic Research.
- Casey, D. & Brugha, C.M. 2005. Understanding the situation of information systems development failure: A role of pragmatism. *Proceedings of the Third International Conference on Action in Language, Organisations and Information Systems (ALOIS*2005)*. Agerfalk, P.J., Bannon L. & Fitzgerald B. (Eds.), Limerick. Ireland, 15-16 March 2005.
- Cassidy, T. 2005. Education Management Information System (EMIS). Development in Latin America and the Caribbean: Lessons and challenges.
- Checkland, P. 1999. *Soft systems methodology: A 30-year retrospective*. England: John Wiley & Sons Ltd.
- Chetty, D. 1992. School efficiency and effectiveness: Pointers for educational transformation in South Africa. Paper presented at the *Economics of Education Conference*, University of Cape Town.

- Chiasson, M. & Dexter, A. 2001. System development conflict during the use of an information systems prototyping method of action research: Implications for practice and research. *Information Technology & People*, Vol. 14 (1): 91-108.
- Choo, C.W. 1996. The knowing organisation: How organisations use information to construct meaning, create knowledge and make decisions. *International Journal of Information Management*, Vol. 15(5): 329-340.
- Ciborra, C.U. 1996. *Teams, markets and systems*. Cambridge: Cambridge University Press.
- Ciborra, C.U. 1998. Crisis and foundations: An inquiry into the nature and limits of models and methods in the information systems discipline. *Journal of Strategic Information Systems*, 7: 5-16.
- Ciborra, C.U. 1999(a). Notes on improvisation and time in organisations. *Accounting Management and Information Technologies*, 9: 77-94.
- Ciborra, C.U. 1999(b). Theory of information systems based on improvisation. In Currie, W. & Galliers, B. (Eds.). *Rethinking management information systems*. Oxford: Oxford University Press: pp. 136-155.
- Clarke, R.J. 2005. Workpractice theory as a foundation for IS research: A systematic semiotic perspective. *Proceedings of the Third International Conference on Action in Language, Organisations and Information Systems (ALOIS*2005)*. Agerfalk, P.J., Bannon L. & Fitzgerald B. (Eds.), Limerick, Ireland, 15-16 March 2005.
- Clotfelter, C., & Ladd, H. 1996. Recognizing and rewarding success in public schools. In Ladd, H.F. (Eds.). *Holding schools accountable. Performance-based reform in education*. Washington, D.C.: The Brookings Institution, pp. 23-59.

- Coleman, J.S. et al. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.
- Cowley, P. & Easton, S. 2004. Report card on British Columbia's secondary schools: Studies in Education Policy, 2004 Edition. Vancouver BC: The Fraser Institute. Available from <http://www.fraserinstitute.ca/admin/books/chapterfiles/Complete%20Publication-BCSSR04.pdf>.
- Crouch, L. 1999. Education indicators in South Africa: Who needs them? What for? Some theoretical and practical considerations. Mimeo: Pretoria.
- Crouch, L. & Mabogoane, T. 1998a. When the residuals matter more than the coefficients: An educational perspective. *Studies in Economics and Econometrics*, 22(2): 1-14.
- Crouch, L. & Mabogoane, T. 1998b. *No Magic Bullets, Just Tracer Bullets*. Mimeo: Pretoria: Department of Education.
- Crouch, L. & Mthethwa, L. 2002. Decentralisation of information or information for decentralisation? The case of South Africa. Paper presented at the *Conference on "Education and Decentralisation: African experiences and comparative analyses"*, June 2002, Johannesburg, South Africa.
- Cuadra, E. 1991. Data collection strategies and methods for monitoring student flows. PHREE *Background paper series*, Document No. PHREE/91/43. Washington, DC: Education and Employment Division and Human Resources Department The World Bank.
- Cuadra, E. & Fredriksen, B. 1992. Scope for efficiency gains resulting from reduction in repetition and dropout. PHREE *Background Paper* No. PHREE/92/55. Washington, DC: Education and Employment Division Population and Human Resources Department, The World Bank, June 1992. Available from [http://www.reduc.cl/raes.nsf/0/4962ba9f1ba70dd5042568340071bbb4/\\$FILE/rae07.336.pdf](http://www.reduc.cl/raes.nsf/0/4962ba9f1ba70dd5042568340071bbb4/$FILE/rae07.336.pdf)

- Davenport, T.H. & Prusak, L. 1997. *Information ecology: Mastering the information and knowledge environment*. Oxford: Oxford University Press.
- Deaton, A. & Case, A. 1996. School quality and educational outcomes in South Africa, draft. Research programme in development studies. Princeton: Princeton University, USA.
- Department of Education 1998. *South African Schools Act No. 84 of 1996 – National Norms and Standards for School Funding*. Pretoria: DoE.
- Department of Education 2002. *Government Notice No. 1120. Government Gazette Vol. 446 No. 23783*. Pretoria: DoE.
- Department of Education 2003. *Report to the Minister: A review of financing, resourcing and costs of education in public schools*. Pretoria: DoE.
- Department of Education 2004. *Education Information Policy*. Pretoria: Department of Education.
- Department of Public Administration. (n.d.). *Education Law and Policy Handbook 1999. Personnel Administration Measures (PAM)*. Pretoria.
- Filmer, D. 1999. Education attainment and enrollment profiles: A resource book based on an analysis of demographic and health survey data. Available from <http://www.worldbank.org/research/projects/edattain/edattain.htm>.
- Filmer, D. & Pritchett, L. 1998a. The effect of household wealth on educational attainment around the world: Demographic and health survey evidence. Policy research working paper no. 1980, September 1998. The World Bank. Washington, DC. Available from <http://www.worldbank.org/html/dec/Publications/Workpapers/WPS1900series/wps1980wps1980.pdf>.

- Filmer, D. & Pritchett, L. 1998b. Educational enrollment and attainment in India: Household wealth, gender, village, and state effects. September 1998. The World Bank. Washington, DC. Available from <http://www.worldbank.org/research/projects/edattain/edindia.pdf>
- Fiske, E.B. & Ladd, H.F. 2002. *Financing schools in post apartheid South Africa: Initial steps towards fiscal equity*. Mimeo. Stanford: Stanford Institute.
- Fiske, E.B. & Ladd, H.F. 2003. Balancing public and private resources for basic education: School fees in post-apartheid South Africa. Draft paper prepared for *Education and social change in South Africa: The challenge of policy*. Pretoria, South Africa: Human Sciences Research Council.
- Fitzgerald, B. 1998. An empirical investigation into the adoption of systems development methodologies. *Information & Management*, 34(98): 317-328.
- Fredriksen, B. 1991. An introduction to the analysis student enrollment and flows. PHREE *Background paper series*, Document No. PHREE/91/39. Washington: Education and Employment Division and Human Resources Department, The World Bank.
- Fuller, B. & Prema, C. 1994. How to raise the effectiveness of secondary schools? Universal and locally tailored investment strategies. Education and Social Policy Department, ESP *Discussion paper series No. 28*. Washington: The World Bank.
- Goldkuhl, G. 2005. Workpractice theory: What it is and why we need it. *Proceedings of the Third International Conference on Action in Language, Organisations and Information Systems (ALOIS*2005)*. Agerfalk, P.J., Bannon L. & Fitzgerald B. (Eds.), Limerick. Ireland, 15-16 March 2005.

- Gorard, S. 2004. Judgement-based statistical analysis. Paper presented at the *British Educational Research Association Annual Conference*, University of Manchester September 2004.
- Gorard, S., Prady, K. & Roberts, S. 2002. An introduction to the simple role of numbers in social science research. *Occasional Paper Series, Paper 53*. ESRC Teaching and Learning Research Programme, Research Capacity Building Network, Cardiff University: School of Social Sciences.
- Grissom, J.B. & Shepard, L.A. 1989. Repeating and dropping out of school. In Shepard, L.A. & Smith, M.L. (Eds.). *Flunking grades: Research and policies on retention*, UK: Falmer Press.
- Gustafsson, M. 2005. The relationships between schooling inputs and outputs in South Africa: Methodologies and policy recommendations based on the 2000 SACMEQ dataset. Paper delivered at *SACMEQ International Invitational Research Conference*, Paris, September 2005.
- Hanushek, E.A. 2002a. Publicly provided education. *Working paper No. 8799*, National Bureau of Economic Research, Cambridge, MA. Available from <http://www.nber.org/papers/w8799>
- Hanushek, E.A. 2002b. The failure of input-based schooling policies. Working paper 9040, National Bureau of Economic Research, Cambridge, MA Cambridge, MA. Available from http://edpro.stanford.edu/hanushek/admin/pages/files/uploads/input_based.EJ.pdf#search=%22hanushek%20the%20failure%20of%20input-based%20schooling%20policies%22.

- Hanushek, E.A. & Raymond, M.E. 2002a. *Lessons about the design of state accountability systems*. Hoover Institution, Stanford University: National Bureau of Economic Research.
- Hanushek, E.A. & Raymond, M.E. 2002b. *The effect of school accountability systems on the level and distribution of student achievement*. Hoover Institution, Stanford University: National Bureau of Economic Research.
- Heiskanen, A. & Newman, M. 1997. Bridging the gap between information systems research and practice: The reflective practitioner as a researcher. *Proceedings of the Eighteenth International Conference on Information Systems*, Kumar, K. & DeGross, J. (Eds.), Atlanta, Georgia, December 15-17, pp. 121-131.
- Henriksen, D.L. 2003. Project web as Practice: On the relevance of radical localism for information systems development research. *Ph.D. Dissertation*, June 2003. Roskilde University, Computer Science. Available from <http://www.ruc.dk/upload/application/pdf/b138e5ea/DS96.pdf>.
- Hirschheim, R. & Klein, H.K. 1989. Four paradigms of information systems development. *Communications of The AC*, Vol. 32(10): 1199-1216.
- Holmes, C.T. 1989, Grade level retention effects: A meta-analysis of research studies. In Shepard, L.A. & Smith, M.L. (Eds.). *Flunking grades: Research and policies on retention*, UK: Falmer Press.
- Hosking, S. 1994. An education production function analysis of Ciskei schools. *Journal for Studies in Economics and Econometrics*, 18(3): 77-90.
- Human Development Report 2001. Calculating the Human Development Indices. Available from <http://hdr.undp.org/reports/global/2001/en/pdf/calculating.pdf>.

- Iivari, H. & Hirschheim, R. 1996. Analysing information systems development: A comparison and analysis of eight IS development approaches. *Information Systems*, 21(7): 551-575.
- Ives, B. & Olson, M. 1984. User involvement and MIS success: A review of research. *Management Science*, Vol. 30(5), 586-603.
- Jaccard, J., Turrisi, R. & Wan, C.K. 1990. *Interaction effects in multiple regression*. Thousand Oaks, CA: Sage Publications Series.
- Jackson, M.C. 2004. *Systems thinking: Creative holism for managers*. England, UK: John Wiley and Sons, Ltd.
- Jacucci, G. 2005. Workpractice theory: Where does it stand? *Proceedings of the Third International Conference on Action in Language, Organisations and Information Systems (ALOIS*2005)*. Agerfalk, P.J., Bannon L. & Fitzgerald B. (Eds.), Limerick, Ireland, 15-16 March 2005.
- Kalermo, J. & Rissanen, J. 2002. Agile software development in theory and practice. Software Business Program, *Master's thesis*. University of Jyväskylä [Online] Available from http://www.cs.jyu.fi/sb/Publications/KalermoRissanen_MastersThesis_060802.pdf.
- Khazanchi, D. & Munkvold, B.E. 2000. Is information systems a science? An inquiry into the nature of the information systems discipline. *The DATA BASE for Advances in Information Systems*, 31(3): 24-42.
- Koufaris, M., Isakowitz, T. & Bieber, M. 1999. Web information systems: Introduction to the Minitrack. *Proceedings of the 32nd Hawaii International Conference on System Sciences*. 1999 IEEE.
- Krueger, A. 1999. Experimental estimates of education production function. *Quarterly Journal of Economics*, May: 497-532.

- Ladd, H.F. 1996. Catalysts for learning: Recognition and reward programs in the public schools. *Brookings Review*, Summer (14)3:14-17. Available from: <http://www.brook.edu/dybdocroot/press/review/laddsu96.html>.
- Lang, M. 2006 Issues and challenges in the development of hypermedia information systems. School of Business & Economics, National University of Ireland, Galway. Available from http://www.is.nuigalway.ie/mlang/research/BIT_2001.pdf.
- Lanzara, G.F. 1999. Between transient constructs and persistent structures: Designing systems in action. *Journal of Strategic Information Systems*, 8: 331-349.
- Laudon, K.C. & Laudon, J.P. 1995. Information systems: A problem solving approach. Fort Worth, TX: The Dryden Press, Harcourt Brace College Publishers.
- Laudon, K.C. & Laudon, J.P. 1998. *Information systems and the internet: A problem solving approach*. Fort Worth, TX: The Dryden Press, Harcourt Brace College Publishers.
- Leedom, D. 2001. Final report: Sensemaking symposium. Technical report (Paper presented at Command and Control Research Council, US. Evidence Based Research Inc. Available from http://www.au.af.mil/au/awc/awcgate/ccrp/sensemaking_final_report.pdf.
- Levin B 1993. Students and educational productivity. *Education Policy Analysis Archives*, Volume 1(5), May 1993. ISSN 1068-2341.AZ: Arizona State University, AZ. Available from <http://epaa.asu.edu/epaa/v1n5.html>.
- Lewis-Beck, M.S. 1995. *Data analysis: An introduction*. Beverly Hills, CA: Sage.

- Lind, M. 2005. Contextual understanding of information systems – Characteristics of process oriented information systems. In: Cambell, B., Underwood, J.& Bunker, D. (Eds.). *Proceedings of the 16th Australasian Conference on Information Systems (ACIS)*, University of Technology, Sydney. Available from <http://www.vits.org/publikationer/dokument/550.pdf>.
- Louridas, P. 1999. Design as bricolage: Antropology meets design thinking. *Design Studies*, 20: 517-535.
- Lundeberg, M., Goldkuhl, G. & Nilsson, A. 1981. *Information systems development: A systematic approach*. New Jersey: Prentice-Hall.
- Lyytinen, K. & Robey, D. 1999. Learning failure in information systems development. *Information Systems Journal*, 9: 85-101.
- Mahood, M.A. 1987. System development methods – A comparative investigation. *MIS Quaterly*, September: 292-311.
- Marchand, D.A. 2001. Is your company effective at using information? *European Business Forum*, Issue 8. Available from http://www.ebfonline.com/pdf/issue_8/p54-57%20ID%20marchand.pdf.
- Marchand, D.A, Kettinger, W.J. & Rollins, J.D. 2002. *Information orientation: The link to business performance* .Oxford: Oxford University Press.
- McEwan, P. & Carnay, M. (1999). The effectiveness and efficiency of private schools in Chile's voucher system. Unpublished manuscript, Stanford University. Available from http://www.refrat.hu-berlin.de/aktion/texte/hintergrund/CM_chile.pdf#search=%22The%20Effectiveness%20and%20Efficiency%20of%20Private%20Schools%22.

- McKay, J. & Marshall, P. 2001. The dual imperatives of action research. *Information Technology & People*, Vol. 14(1), 46-59.
- Meier, K.J., Wrinkle, R.D. & Polinard, J. 1999. Black student improvements on the TAAS Exam. A report of the Texas Educational Excellence Project. Texas A & M University, University of Texas-Pan American. Available from <http://teep.tamu.edu/reports/report007.pdf>.
- Mendro, B., Olivarez, R., Peters, M., Milanowski, T., Kellor, E. & Odden, A. 1999. A case study of the Dallas public school school-based performance award program. Consortium for Policy Research in Education. Wisconsin Center for Education Research, University of Wisconsin-Madison. January 1999. Available from <http://www.wcer.wisc.edu/cpre>.
- Meyer, R. 1997. Value-added indicators of school performance: A primer. *Economics of Education Review*, 16(3): 283-301.
- Miner, A.S. 1998. Organisational improvisation and organisational memory. *Academy of Management Review*, 23(4): 698-723.
- Miner, A.S., Bassof, P. & Moorman, C. 2001. Organizational improvisation and learning: A field study. *Administrative Science Quarterly*, 46: 304-337.
- Moorman, C. & Miner, A.S. 1998(a). The convergence of planning and execution: Improvisation in new product development. *Journal of Marketing*, 62: 1-20.
- Moorman, C & Miner, A.S. 1998(b). Organisational improvisation and organizational memory. *Academy of Management Review*, 23(4): 698-723.
- Multiple Regression. 2003. Available from <http://www2.chass.ncsu.edu/garson/pa765/regress.htm>.

- Mumford, E. 2001. Advice for an action researcher. *Information Technology & People*, Vol. 14 (1), 12-27.
- Necco, C.R., Gordon, C.L. & Tsai, N.W. 1987. Systems analysis and design: Current practices. *MIS Quarterly*, 460-476.
- Nerur, S., Mahapatra, R. & Mangalara, G. 1998. Challenges of migrating to agile methodologies. *Communications of the ACM*, 41(7): 73-78.
- Nurminen, M.I. 2005. Work practice – Core of work informatics: Position paper for the panel on “Work practice theory as foundation for IS research”. *Proceedings of the Third International Conference on Action in Language, Organisations and Information Systems (ALOIS*2005)* 15–16 March 2005, Limerick, Ireland.
- O’Brien, R. 1998. An overview of the methodological approach of action research Available from <http://www.web.ca/~robrien/papers/arfinal.html>.
- O’Brien, J.A. 1991. *Introduction to information systems in business management*. Boston, MA: IRWIN.
- Oz, E. 2002. *Management information systems*. Canada: Course Technology.
- Peplowski, K. 1998. The process of improvisation. *Organisation Science*, 9(5): 560-561.
- Pritchett, L. & Filmer, D. 1999. What education production functions really show: A positive theory of education expenditures. *Economics of Education Review*, 18: 223-239.
- Rose, J. 2000. *Information systems development as action research - soft systems methodology and structuration theory*. Ph.D. Thesis. Available from <http://www.cs.auc.dk/~jeremy/pdf%20files/thesis.pdf>.

Rothstein R. 2001. To improve student learning: Ask the right questions. Available from http://www.ourfuture.org/projects/next_agenda/ch6_1.cfm.

Subotnik, Rena & Walberg, Herbert. 2006. *The Scientific Basis of Educational Productivity*.
Greenwich, CT: Information Age Publishing, Inc.
Available from: <http://edrev.asu.edu/essays/v9n4.pdf>

Scatterplot 2004. Available from <http://www.stat.yale.edu/Course/1997-98/101/scatter.htm>.

Schagen, I. 2002. Well, what do you know? *Education Journal* ,63: 27-28.

Serafeimidis, V. & Smithson, S. 2003. Information systems evaluation as an organisational institution – Experience from a case study. *Information Systems Journal*, 13: 251-274.

Skidmore, S. & Eva, M. 2004. *Introducing systems development*. New York: Palgrave, Macmillan.

South African Government. 1995. *White Paper on the Transformation of Public Service*.
Pretoria: Department of Education.

South African Government. 1998. *Employment Equity Act*. Government Gazette, Vol.400, No. 19370. Cape Town.

Statistics South Africa. 2001. *Census 2001*. Pretoria: Statistics South Africa.

Susman, G.I. & Evered, R.D. 1978. An assessment of the scientific merits of action research. *Administrative Science Quarterly*, Vol. 23: 582-603.

Thonstad, T. 1981. *Analysing and projecting school enrolment in developing countries: A manual of methodology*. Paris: UNESCO: Education Projections Unit, Division of Statistics on Education, UNESCO Office of Statistics.

- Truex, D., Baskerville, R. & Travis, J. 2000. A methodical systems development: The deferred meaning of systems development methods. *Accounting Management and Information Technologies*, 10: 53-79.
- UNESCO 1998. Wasted opportunities: When schools fail repetition and drop-out in primary schools. A report published by UNESCO for the International Consultative Forum on Education For All, Education for All Status and Trends 1998, Paris.
- UNESCO. 2004a. How is school wastage measured: Efficiency in education. Available from http://www.uis.unesco.org/en/stats/statistics/indicators/i_pages/indspec/efficiency.htm.
- UNESCO. 2004b. Technical notes on cohort student flow analysis. Available from http://www.uis.unesco.org/en/stats/statistics/indicators/i_pages/indspec/cohorte.html.
- UNESCO. 2006. Education Management Information System. Available from http://portal.unesco.org/education/en/ev.php-URL_ID=10202&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- Van der Berg, S. 2005a. How effective are poor schools? Poverty and educational outcomes in South Africa. Revised version of paper delivered at *SACMEQ International Invitational Research Conference*, Paris, September 2005.
- Van der Berg, S. 2005b. Apartheid's enduring legacy: Inequalities in education. Revised version of a paper presented at The Oxford University/University Stellenbosch conference on The South African Economic Policy under Democracy: A 10 year Review. 22 December 2005.
- Van der Berg, S. & Burger, R. 2003. Education and socio-economic differentials: A study of school performance in Western Cape. Development Policy Research Unit, School of Economics, University of Cape Town. Working Paper No 03/73 March 2003.

- Van der Berg, S. & Louw, M. 2006. Unravelling the mystery: Understanding South African schooling outcomes in regional context. Paper presented at the The Centre for the Study of African Economies Conference, Oxford University, 21st March 2006.
- Vegas, E. 2002. School choice, student performance, and teacher school characteristics: The Chilean case, World Bank Policy Research Working Paper 2833, April 2002. Development Research Group, World Bank, Washington D.C. Available from http://econ.worldbank.org/files/15347_wps2833.pdf.
- Von Bertalanffy, Ludwig. 1968. *General systems theory*. New York: Allen Lane Penguin Press.
- Wako, T.N. 1995. Basic indicators of education systems performance: A manual for statisticians and planners, Ministry of Education, June 1995, Addis Ababa. Available from <http://www.harare.unesco.org/documents/2005/Basic%20Indicators%20of%20Education%20Systems%20Performance.pdf#search=%22Indicators%20of%20Education%20Systems%3A%20A%20Manual%20for%20%22>
- Wako, T.N. 2003. Education Management Information Systems (EMIS): An Overview. NESIS/UNESCO. Harare, Zimbabwe. Available from <http://nesis.easynet.fr/Documents/2004/EMIS14.doc>
- Wako, T. N. (n.d.). An open secret behind internal inefficiency in first half of primary grades in Ethiopia: Paper prepared for Nazareth Conference. Available from [http://nesis.easynet.fr/Documents/The Secret.doc](http://nesis.easynet.fr/Documents/The%20Secret.doc).
- Warren, L. & Adman, P. 1999. The use of critical systems thinking in designing a system for a university information systems support service. *Information Systems Journal*, 9: 223-242.

- Wegener, T. 1993. *Applied business statistics*. Cape Town:Juta & Co. Ltd.
- Weick, K.E. 1995. *Sensemaking in organisations*. Thousands Oaks, CA: Sage Publications, Inc.
- Weick, K.E. 1998. Improvisation as a mindset for organisational analysis. *Organisation Science*, 9(5): 543-555.
- Weick, K.E. 2001. *Making sense of the organisation*. Malden, MA: Blackwell Publishers Inc.
- Windham, D. 1993. Strategies for decentralizing data use. In Chapman, David & Lars, Mählcck, (Eds.). *From data to action: Information system in educational planning*. Paris: Unesco, International Institute for Education Planning: Pergamon Press: pp. 26-146.
- Wolff, L., Schiefelbein, E. & Valenzuela, 1994. Improving the quality of primary education in Latin America and the Caribbean: Toward the 21st Century. Washington, D.C: World Bank Discussion Papers.

APPENDIX A

“The determination of the REQV is based primarily on the number of recognised prescribed full-time professional or academic years of study at an approved university, technikon or college of education and taking into account the level of school education attained” (Education Law and Policy Handbook 1999: 3C-19). Table 28 outlines the requirements of the relative education qualification value (REQV).

Table 28: Requirements in respect of REQVs

REQV	Education qualification
10	Grade 12 or lower without a teacher’s qualification
11	Grade 8,9,10 or 11 plus a teacher’s qualification of at least two years apposite training
12	Grade 12 plus one or two years apposite training
13	Grade 12 plus three years apposite training
14	Grade 12 plus four years apposite training
15	Grade 12 plus five years apposite training
16	Grade 12 plus three years apposite training. Only professional qualified educators can be classified under REQV 16, and only provided such persons are in possession of a recognized university degree.
17	Grade 12 plus three years apposite training. To be regarded as having an REQV 17, a candidate must, in addition to the requirements for classification under REQV 16, also be in possession of at least a recognized master’s degree

Source: Education Law and Policy Handbook 1999: 3C-20

APPENDIX B

Dependent Variable: Grade 12 Aggregate (Aggregate)

The grade 12 aggregate marks of a learner is determined by adding the marks of the best six¹⁸ subjects of a learner, including the compulsory languages. The grade 12 aggregate is then calculated by averaging the aggregate of the learners of a particular school.

Independent Variables

1) Poverty Index of the Community (Poverty Index)

The poverty index of the community is an overall rating of each community's socio-economic status (SES). A composite index has been constructed by using data of the 1996 census to describe the socio-economic status of the community in which the school is situated. The measure uses similar components to the Human Development Index (HDI) (Human Development Report, 2001:240). The HDI has been constructed every year since 1990 and is meant to identify a community in terms of its poverty and vulnerability. Communities high on the list are poorer and more vulnerable than communities lower down the list. The overall rating of each community's socio-economic status are based on the following indicators: (i) income, (ii) educational attainment, (iii) unemployment, and (iv) lack of water supply.

Income

A composite index for income was based on the following three indicators:

- (i) Indicator 1: the number of households with a total derived income below R18000 per annum expressed as a percentage of the total number of households in an area.

¹⁸ A learner could offer more than six subjects.

(ii) Indicator 2: the number of households with a total derived income between R18000 and R54000 per annum expressed as a percentage of the total number of households in an area.

(iii) Indicator 3: the number of households with a total derived income between R54000 and R96000 per annum expressed as a percentage of the total number of households in an area.

The results of each of these indicators were then converted into an index ranging from 0 to 1 according to the general formula:

$$\text{INDEX} = (\text{Actual } x_i \text{ value} - \text{minimum } x_i \text{ value}) / (\text{Maximum } x_i \text{ value} - \text{minimum } x_i \text{ value})$$

(1)

These indices¹⁹ were then combined to produce a weighted index for income of each community. The weightings used in the calculation were first index 0.6, second index 0.3 and third index 0.1.

Educational attainment

This index was based on the UNESCO-definition for functional literacy: “The number of persons older than 14 who have completed 7 years formal schooling”. In a South-African context it includes everybody who has completed the primary school phase. The index is calculated as the number of persons older than 14 years who did not complete primary school expressed as a percentage of total persons older than 14 years in an area.

¹⁹As this exercise is not to distinguish between rich and poor, but rather to determine the most poor and vulnerable area, one could have taken only the number of households with a total income range below R18000 as a percentage of the total households. However, in that case middle-income areas or areas just above this line, for example between R18000 and R54000, are penalized unduly. Therefore it is more appropriate to work on different weights for the indices in the income ranges and this provides also a more refined distinction than just taking the average income of the area.

Unemployment

The number of persons unemployed looking for work was expressed as a percentage of the total population over 15 within an area. Thus when the general formula was applied it formed the index for unemployment.

Lack of water supply

This index was calculated by expressing the number of all households who do not have piped water in their homes or on their dwellings site as a percentage of the total number of households. Again the general formula was used to calculate the index for water supply.

The composite index for poverty of a community was calculated as follows²⁰:

POVERTY INDEX = (functional illiteracy*0.1 + unemployment*0.1 + water supply*0.1 + income*0.7)

2) Average age of teachers (Age)

The average age of all teachers at a particular school.

3) Average experience of teachers (Experience)

The average teaching experience of all teachers at a particular school.

²⁰ The weighting factor was arbitrarily determined merely based on the important contribution of income to poverty.

4) **Attainment rate**

The number of learners who enrolled in grade 8 in 1998, expressed as a percentage of the grade 12 learners in 2002 of a particular school.

5) **Average Qualifications of Teachers (REQV)**

This indicator is calculated based on the average qualifications of all teachers at a particular school. The relative education qualification value (REQV) of a teacher is ten plus the number of recognised prescribed full-time professional or academic years of study at an approved university, technikon or college. For example a teacher with a REQV 13 means teacher qualification with 3 years academic training. Regression analysis would result into misleading estimates if the qualifications of teachers were used according to the exact REQV qualification, because no teacher could ever have qualifications lower than 10. Therefore, in the production function analysis, teacher qualification would be calculated by using the school level average REQV minus 10. A school with an average REQV of 14 would then have a value of 4 for teacher qualification, which in fact indicated the number of teacher training years.

REQV 14 is the number of teachers with relative education qualification value (REQV) of 14, expressed as a percentage of the number of teachers at a particular school.

REQV 15 is the number of teachers with REQV of 15, expressed as a percentage of the number of teachers at a particular school.

REQV 16 is the number of teachers with REQV of 16, expressed as a percentage of the number of teachers at a particular school.

REQV 17 is the number of teachers with REQV of 17, expressed as a percentage of the total number of teachers at a particular school.

6) Number of teachers paid by the community (SGB teachers)

The number of privately paid²¹ teachers per 1000 learners at the school.

7) Number of teachers paid by the state (State teachers)

The number of teachers paid by the state per 1000 learners at the school.

8) Ex Cape Education Department (CED)

A dummy variable indicating the previous education department of the school. The CED-variable include schools (predominantly white) under the old apartheid regime.

9) Ex House of Representatives (HOR)

This dummy variable includes schools predominantly from the coloured population.

10) Ex Department of Training (DET)

This dummy variable includes only learners from black schools under the previous government before 1994.

²¹ Teachers paid by the school's governing body.

11) Number of learners who repeated grade 11 (Repeaters)

The number of grade 11 repeaters in 2002, expressed as a percentage of the grade 11 enrolment in 2001. This is a good indicator to determine if weeding (skimming) takes place in a particular school.

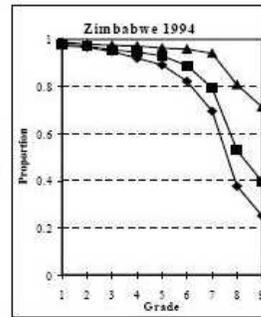
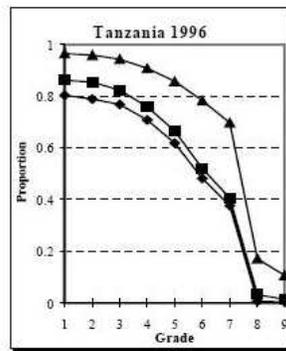
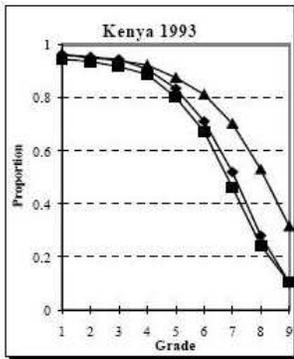
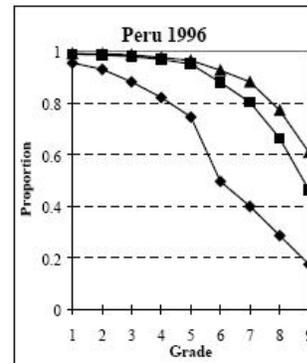
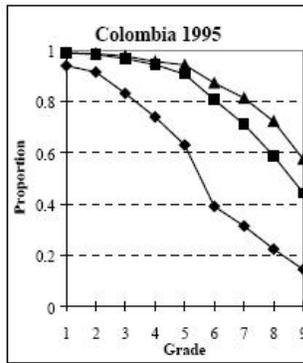
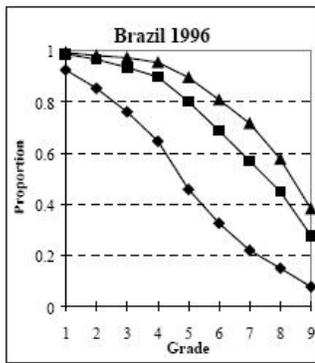
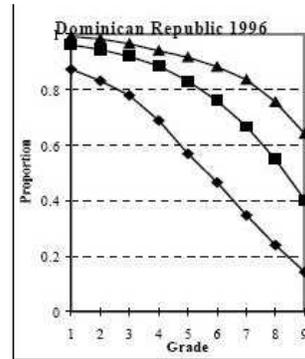
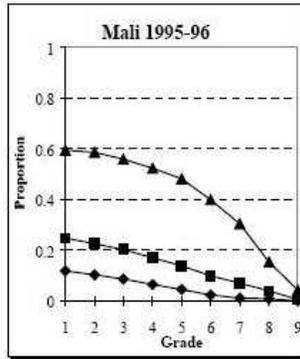
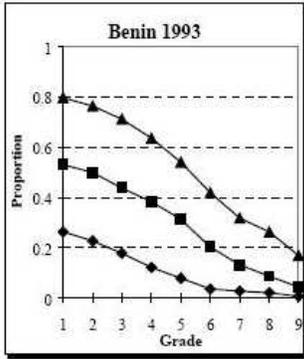
12) Dropout rate

The dropout rate of a particular school. The calculation of dropout rates per school is explained in the next chapter.

13) Metropole Central

A dummy variable for a specific education district of a particular school. The Western Cape Education Department is divided into a number of education management districts. Other district dummies were tried but found not to be significant.

APPENDIX C : Filmer and Pritchett (1998a)



Legend

