

**EFFECTIVENESS OF THE TEACHER PERFORMANCE EVALUATION SYSTEM
OF
KWEKWE DISTRICT IN ZIMBABWE.**

BY

BELINDA RINDAI MUSODZA

16013601

A thesis submitted in fulfilment of the requirements for the degree of

DOCTOR OF EDUCATION

in the

DEPARTMENT OF EDUCATIONAL MANAGEMENT

SCHOOL OF EDUCATION

UNIVERSITY OF VENDA

PROMOTER

DR M. MPETA

CO-PROMOTER

PROF. T. RUNHARE

DECLARATION

I **BELINDA RINDAI MUSODZA** declare that this thesis entitled, '**EFFECTIVNESS OF THE TEACHER PERFORMANCE EVALUATION SYSTEM IN KWEKWE DISTRICT OF ZIMBABWE**' is my own work and has not been submitted previously for a degree, at this or any other university. It is my own work in design and execution and all reference work contained therein has been duly acknowledged.

Signature.....

Date.....

EDITOR'S CERTIFICATE

DocuSign Envelope ID: AD8570BA-6558-4ED7-B62D-818531360A5E

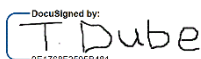
Dear Sir/Madam

Confirmation of technical and language editing plus proof reading of the Doctor of Education Report for Ms Musodza Belinda. Rindai (16013601)

This is to certify that I proof read and edited Ms M.B. Rindai's PhD Report entitled: **Effectiveness of the Teacher Performance Evaluation System in Kwekwe district of Zimbabwe.**

I advised on the acceptance and/or rejection of modifications as appropriate, and consultation with the promoter where necessary. Thereafter, the revised document can be submitted for processing.

Signed:T. Dube Dated: 15/02/2019

DocuSigned by:

9E1788F2895B481...

Tendai Dube PhD, MSc, BSc (Hon), FHEA, Dip Ed, PGCert HE, DIP IT, CMALT
Breakthrough Proofreading
Nottingham
United Kingdom
Email: tendaidube@hotmail.co.uk

DEDICATION

This thesis is dedicated to my late dear loving mum, Sipiwe Zhou Dewa. A hard worker and a woman of virtue who nurtured and supported me lovingly to the end. How proud she would have been of me for having achieved this! The thesis is also dedicated to my dear husband, Alexander, and our son, Alexander Jnr Takunda, for having had to endure my long absence during the study.

ACKNOWLEDGEMENTS

To Jehovah my God, I am thankful for giving me life and hope and shielding me from major problems that could have distracted me from this task. I held on to Him through entirely and eternal thanks are devoted to Him.

My promoters, Dr. Mamoneta Mpeta and Professor Tawanda Runhare, you were a source of inspiration at every stage of my research journey. I pay special tribute to you for the edifying comments and expert guidance throughout the research process, may Jehovah bless you and continue to endow you with the necessary wisdom and patience in your work. A special thank you to the Professor for bailing me out by providing student support for the registration process.

My heartfelt gratitude is also extended to Dr Tendai Dube for language editing of the thesis. Your encouragement is also appreciated.

To my life-coach, my late mother Sipiwe Dewa, because I owe it all to you, how I wish you could have been here to see me through. You taught me to persevere and be resolute as well as to trust in Jehovah our God. Many Thanks! To my father, a profound thank you for the support and encouragement.

My forever interested, cheering and always inquisitive son, Alexander Jr. Takunda Musodza, he was always keen to know what I was doing and how I was proceeding, although it is likely that he had at no time grasped what the research process entailed. Your time shall come. I am also grateful to my husband, the 'research assistant', for being with me for the pilot and data collection. I also thank him for typing and the preliminary editing of the report. To my three daughters, Elaine, Tendaishe and Alexis, thank you for providing moral and emotional support throughout my study. A special mention to Nqobile Muleya, my son-in law for initiating the whole process, encouragement and the moral support throughout the research journey. I thank him for the expert guidance in SPSS. Rolivhuwa!

A very special gratitude goes to the University of Venda Directorate of Research for providing the basic research project funding. My gratitude also goes to the University of Venda for the tuition and residence fees. I am also grateful to Mr Kenneth Kgomo

for being God sent in my time of distress and rescuing me out of a potentially impossible situation with financial support. May the Almighty bless him for his kindness.

To the management and staff, at the Ministry of Education in Zimbabwe, school heads and teachers of the participating schools, thank you for your co-operation and for making the study a reality, for without you, there would have been no study. A special thank you to Mrs Gudo, the then Midlands Provincial Education Director. I was humbled by your respectful, friendly and cooperative nature which made me feel welcome in the province. In the short period I was privileged to interact with you, and you were inspirational. To Mr Mazambani, the Kwekwe District Education Inspector, a special thank you to you and your team. Your support and cooperation made it enjoyable and bearable to conduct the otherwise difficult and challenging task.

I want to record my appreciation to Mr Simba Kativhu for the assistance with ATLAS ti software operations. To my friend and colleague Misozi Chiweshe, your input and support was greatly appreciated. To Dr D. Nkomo and Dr B. Dube, your unending encouragement and support was key throughout the rocky and stormy journey.

And finally, to everyone in the School of Education, it was great sharing with you in the three. Thank you for your expertise!

ABSTRACT

There has been greater demand for more accountability, results-based culture and enhanced effectiveness of programmes and services globally. The education sector has not been spared by this wave either and hence teacher effectiveness and instructional quality have risen to the top of the educational policy agenda. It is important therefore that effective teaching must be assured and teacher evaluation is a key means of providing that assurance. To date, most studies on the teacher appraisal system in Zimbabwe have focused on the perceptions and attitudes of the teachers towards the evaluation process, and at the same time the implementation challenges. Little has so far been done to determine the effectiveness of the teacher performance evaluation system in Zimbabwe. This study consequently sought to evaluate the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. The study was premised on the pragmatic philosophical worldview and hence the mixed method approach was adopted. The convergent parallel mixed method design was used. Data was collected using individual face to face semi structured interviews and a 5 point Likert scale questionnaire. Documentary review was done prior to the development of the research instruments as a way of ensuring relevance of the data collection instruments. The study was underpinned by the self-developed RADPS conceptual framework on performance evaluation system effectiveness. Stratified purposive sampling technique was used to select ten secondary schools for the quantitative strand and four for the qualitative strand. The quantitative sample was composed of 292 teachers and the qualitative sample was composed of 12 participants constituting of 4 teachers, 4 heads of departments and 4 schools heads. Quantitative data was analysed using the SPSS version 25 while the qualitative data was analysed using ATLAS ti. 8. The key findings of the study were that: the performance evaluation system was imposed and accordingly there was no buy in; there was inadequate budgetary support thereby rendering the introduction of the system mistimed; teachers as key stakeholders were excluded from the design process and hence there was no ownership; and ultimately, relevance of the system was questioned. It was also revealed that the evaluation process is merely a compliance exercise with erratic and discontinuous monitoring and supervision through the evaluation cycle. The findings from the study imply the following: policy formulation should be participatory and inclusive; readiness assessment should be

conducted before introducing a new programme; a programme should be pilot tested; evaluation process should be monitored by external officials; the process should be consequential and the system should be continuously monitored and reviewed for relevance and validity.

Key Concepts

Evaluation, performance management, results based management, teacher effectiveness,

TABLE OF CONTENTS

CHAPTER ONE.....	1
Background and Introduction to the Study	1
1.1 Chapter Overview.....	1
1.2 Background to the study	1
1.2.1 Contextual background.....	4
1.3 Statement of the Problem	9
1.4 Aim of the Study	10
1.4.1 Research Objectives.....	10
1.5 Research Questions	11
1.5.1 Main research question.....	11
1.5.2 Quantitative Research Questions	11
1.5.3 Qualitative research question.....	11
1.5.3 The Hybrid or integration research question.....	12
1.6 The Conceptual Framework	12
1.6.1 Readiness Assessment.....	12
1.6.2 Design of evaluation system.....	13
1.6.3 Process of evaluation	13
1.6.4 Significance of evaluation process	14
1.7 Definition of Key Concepts.....	14
1.8 Research Paradigm, Design and Methodology	14
1.8.1 Research Paradigm	15
1.8.2 Research design.....	16
1.8.3 Research Methodology	16
1.9 Data analysis.....	17
1.9.1 Quantitative analysis.....	17
1.9.2 Qualitative analysis	18
1.10 Sampling process.....	18
1.10.1 Study population.....	18
1.10.2 Sampling design.....	18
1.10.3 Sample size.....	20
1.11 Quality control measures	21
1.11.1 Piloting the data collection instruments	22
1.11.2 Validity and reliability of Quantitative data.....	22

1.11.3	Ensuring trustworthiness in qualitative data.....	23
1.12	Delimitation of the Study	24
1.13	Ethical considerations.....	25
1.13.1	Approval for conducting the research	25
1.13.2	Upholding participants' rights	25
1.13.3	Upholding the integrity of the research profession.....	26
1.14	Significance of Study	26
1.15	Research outline.....	27
1.15.1	Chapter 1.....	27
1.15.2	Chapter 2.....	27
1.15.3	Chapter 3.....	28
1.15.4	Chapter 4.....	28
1.15.5	Chapter 5.....	28
1.15.6	Chapter 6.....	28
1.16	Chapter summary.....	28
CHAPTER TWO		30
The RADPS Conceptual Framework on Effectiveness of Performance Evaluation System..		30
2.1	Chapter overview	30
2.2	Introduction.....	30
2.3	Conceptual framing versus theoretical framing	30
2.4	The origin of the RADPS conceptual framework.....	35
2.4.1	Components of the RADPS Framework.....	35
		44
		44
2.5	Strengths of the RADPS conceptual framework	45
2.5.1	Theoretical elements embedded in the RADPS conceptual framework.....	46
2.6	Potential limitations of the RADPS conceptual framework	49
2.7	Chapter summary.....	50
CHAPTER THREE.....		51
A review of related literature		51
3.1	Overview.....	51
3.2	Introduction.....	51
3.3	The concept of performance evaluation	52
3.4	Global perspective of evaluation.....	53
3.4.1	Teacher evaluation in South Korea	54
3.4.2	Teacher evaluation in Chile	62

3.4.3	Teacher Evaluation in China	71
3.4.4	Comparative discussion of the three country (South Korea, Chile and China) evaluation systems.....	81
3.5	Readiness Assessment.....	84
3.6	The design process of a performance evaluation system	86
3.6.1	Conditions that have to be observed in the design process	87
3.6.2	Stakeholder involvement in the design process	87
3.6.3	Steps in designing an evaluation system	89
3.7	Implementing teacher performance evaluation	102
3.7.1	Enabling environment for implementing performance evaluation.....	102
3.7.2	Implementation procedures	104
3.7.3	Challenges of implementing performance evaluation.....	106
3.8	Significance of performance evaluation process.....	109
3.8.1	Making evaluation significant by linking it to professional development activities	109
3.8.2	Making evaluation significant by ensuring accountability by the teachers	112
3.9	Evaluation in Zimbabwe	115
3.9.1	Statutory Instrument 1 of 2000.....	115
3.9.2	RBM system in Zimbabwe	118
3.10	Chapter Summary	126
CHAPTER FOUR		128
Research Design and Methodology		128
4.1	Chapter Overview.....	128
4.2	Introduction.....	128
4.3	Research Philosophical assumptions	129
4.3.1	Ontological research proposition	129
4.3.2	Epistemological research proposition	130
4.3.3	Methodological research proposition	131
4.3.4	Axiological research proposition	131
4.3.5	My philosophical assumptions on this study	132
4.4	Research Paradigm	133
4.4.1	Positivism.....	133
4.4.2	Constructivism	133
4.4.3	Pragmatism	134
4.5	The Research Methodology.....	136
4.5.1	Quantitative approach.....	136

4.5.2	Qualitative approach	137
4.5.3	Mixed methods research approach	137
4.6	The Mixed Method Research Designs.....	140
4.6.1	The Convergent parallel mixed method design	141
4.6.3	Examples of convergent parallel mixed method studies	146
4.6.4	Mixed method purpose statement	146
4.6.5	Research questions	147
4.7	Data Collection methods	149
4.7.1	Study population and sample	149
4.7.2	Selecting the research sample	150
4.7.3	The sampling design.....	150
4.8	Instrumentation	157
4.8.1	Quantitative data collection.....	158
4.8.2	Qualitative data collection	159
4.9	Data Analysis	161
4.9.1	Quantitative data analysis.....	162
4.9.2	Qualitative Data analysis.....	166
4.10	Quality Assurance Measures for the Study.....	170
4.10.1	Piloting the questionnaire.....	170
4.10.2	Piloting the interview schedule.....	171
4.10.3	Quantitative quality assurance measures	172
4.10.4	Qualitative quality assurance measures.....	175
4.11.1	Approval for conducting research.....	179
4.11.2	Participants' rights.....	180
4.11.3	Researcher's conduct.....	181
4.12	Chapter Summary	182
CHAPTER FIVE		184
Data presentation, Analysis and Discussion.....		184
5.1	Chapter Overview.....	184
5.2	Introduction.....	185
5.3	Demographic information	185
5.3.1	Quantitative sample	185
	Qualitative Sample	186
5.4	Results and discussions.....	187
5.4.1	Presentation of Multiple regression analysis results.....	188

5.5	Objective 1: To investigate the relationship between the timing of introduction and effectiveness of evaluation system.....	190
5.5.1	Quantitative Results on objective 1	190
5.5.2	Comparison of quantitative results on objective 1 according to location.....	192
5.5.3	Qualitative results on objective 1	194
5.5.4	Comparing and contrasting quantitative and qualitative results on the appropriateness of timing of introduction of the evaluation system	199
5.5.5	Discussion of quantitative and qualitative results on timing of introduction	201
5.6	Objective 2: To examine the design process of the teacher performance evaluation system.....	203
5.6.1	Quantitative results on objective 2.....	203
5.6.2	Comparison of perception of teachers on design process of the evaluation system	206
5.6.3	Qualitative results on the design process of RBM	208
5.6.4	Comparing and contrasting of quantitative and qualitative results on the design process	218
5.6.5	Discussion of quantitative and qualitative results on the design process.....	220
	inadequate stakeholder involvement in design of RBM	220
5.7	Objective 3: To appraise the implementation process of teacher performance evaluation system.....	223
5.7.1	Quantitative results on the implementation process	224
5.7.2	Comparison of quantitative results on implementation based on location	226
5.7.3	Qualitative results on the implementation process	228
5.7.4	Comparing and contrasting quantitative and qualitative results on the implementation process	238
5.7.5	Discussion of quantitative and qualitative results on the implementation process	240
5.8	Objective 4: To Assess the significance of the teacher performance evaluation process	245
5.8.1	Quantitative Results on significance of evaluation process.....	245
5.8.2	Views of urban and rural teachers on significance of evaluation process.....	248
5.8.3	Qualitative results on significance of appraisal process.....	250
5.8.4	Comparing and contrasting quantitative and qualitative results on significance of appraisal process.....	253
5.8.5	Discussion of quantitative and qualitative results on the significance of the process	255
5.9	General opinions on RBM.....	257
5.9.1	Quantitative results on teachers' general opinion on the evaluation system ..	258

5.9.2	Comparison of quantitative results on teachers' general opinion on the evaluation process	258
5.9.3	Qualitative results on educators' general opinion on an evaluation system ..	259
5.9.4	Comparing and contrasting of quantitative and qualitative results on what should be done with RBM	261
5.9.5	Discussion of quantitative and qualitative results on recommendations on RBM	262
5.10	Recommendations by the participants on how to improve RBM	263
5.10.1	Recommendations to enhance the appropriateness of the timing of the introduction.....	263
5.10.2	Recommendations to improve the design process	265
5.10.3	Recommendations to improve the implementation process.....	266
5.10.4	Recommendations to enhance the significance of the appraisal process	267
5.11	Chapter summary.....	268
CHAPTER SIX.....		271
Summary, Conclusions and Recommendations		271
6.1	Chapter Overview.....	271
6.2	Introduction.....	271
6.3	Summary of the research process.....	271
6.4	How the study responded to the research objectives.....	272
6.4.1	Objective 1: To investigate the relationship between the timing of the introduction and effectiveness of the system	272
6.4.2	Objective 2: To Examine the design process of the teacher performance evaluation system in Kwekwe District	274
6.4.3	Objective 3: To Appraise the implementation process of teacher performance evaluation system in Kwekwe district.....	275
6.4.4	Objective 4: To Assess the significance of the teacher performance evaluation process in Kwekwe District in promoting professional development and accountability	276
6.5	Overview of the major study findings	277
6.5.1	Challenges with regards to the introduction of RBM in Kwekwe	277
6.5.2	Inadequate stakeholder involvement in the design process.....	278
6.5.3	Uneven implementation process of RBM	281
6.5.4	Insignificant effect of RBM on professional teacher development and accountability	283
6.5.5	General position of educators on the way forward on RBM	285
6.6	Study conclusions	285
6.6.1	Inadequate readiness for assessment.....	286
6.6.2	Non- participatory design process	286

6.6.3	Impurities in the implementation process	287
6.6.4	Insignificance of appraisal process on teacher professional development and accountability	287
6.6.5	The status quo on teacher performance evaluation in Kwekwe district.....	288
6.7	Recommendations	290
6.7.1	Recommendations for Policy	290
6.7.2	Recommendations for practice	292
6.8	Significance of the study	295
6.8.1	Significance on research approach	295
6.8.2	Study significance on RADPS conceptual framework	295
		297
6.8.3	Significance on practice and policy	298
6.9	Limitations of study	298
6.10	Suggestions for further research.....	299
6.11	Chapter summary and conclusion	299

LIST OF APPENDICES

Appendix 1: University of Venda Higher Degrees Committee project approval.....	321
Appendix 2.1: Recommendations from ethics committee.....	322
Appendix 2.2: Ethical clearance certificate.....	323
Appendix 3.1: Permission from Ministry of Primary and Secondary Education of Zimbabwe.....	324
Appendix 3.2: Permission from Provincial office.....	325
Appendix 3.3: A letter stamped by the District Schools Inspector.....	326
Appendix 4.1: Consent form for questionnaire.....	327
Appendix 4.2 : Consent form for interview.....	329
Appendix 5.1: Questionnaire for teachers.....	331
Appendix 5.2: Interview schedule.....	335
Appendix 6: Sample of codes and quotations from ATLAS ti.....	338

LIST OF TABLES

Table 2.1:	A summary of differences and similarities between conceptual and theoretical frameworks.....	33
Table 4.1:	The outline of the sampling procedures and sample sizes.....	160
Table 5.1:	Demographic information on quantitative sample.....	189
Table 5.2 :	Demographic information on qualitative sample.....	190
Table 5.3 :	Model summary table.....	191
Table 5.4 :	ANOVA table.....	192
Table 5.5 :	Beta coefficient table.....	193
Table 5.6 :	Perceptions of teachers on appropriateness of the timing of introduction of the teacher evaluation system.....	194
Table 5.7:	Perceptions of teachers on timing based on location.....	196
Table 5.8:	Perceptions of teachers on the design process	207
Table 5.9:	Perceptions of teachers on design process based on geographical location.....	210
Table 5.10:	Perception of teachers on implementation process.....	227
Table 5.11:	Comparison of quantitative results on implementation.....	230
Table 5.12:	Perception of teachers on significance of the evaluation process...	249
Table 5.13:	Perception of teachers on the significance of the evaluation process based on location.....	251
Table 5.14:	Perception of teachers on what is to be done with RBM.....	261
Table 5.15:	Comparison of quantitative results on teachers general opinion on evaluation process based on location.....	262

LIST OF FIGURES

Figure 1.1:	The map of Zimbabwe showing the position of Kwekwe district.....	5
Figure 1.2:	The map of Kwekwe district.....	6
Figure 2.1:	The components of the RADPS conceptual framework.....	45
Figure 4.1:	Convergent parallel mixed methods design procedure.....	145
Figure 4.2:	Triangulation as an analogue.....	146
Figure 4.3:	Convergent parallel data analysis.....	165
Figure 5.1:	Concerns regarding the timing of the programme.....	198
Figure 5.2:	Capacity of evaluation instrument to be informative.....	218
Figure 5.3:	Knowledge of the statute regulating the appraisal process.....	236
Figure 6.1:	The status quo on teacher performance appraisal in Kwekwe.....	294
Figure 6.2:	Recommended RADPS conceptual framework.....	302

ACRONMYS

AEP:	Assignment for Pedagogical Training
ANOVA:	Analysis of Variance
AVDI:	Voluntary Individual Grant Assignment
CAQDAS:	Computer Assisted Qualitative Data Analysis
CDE:	Centre for Development and Enterprise
CDEIP:	Centre for Pedagogical Training, Experimentation and Research
DIPA:	Departmental Integrated Performance Agreement
ECD:	Early Childhood Development
ESSA:	Every Student Succeeds Act
HOD:	Head of Department
IRBs:	Institutional Review Boards
KRA:	Key Result Area
M & E:	Monitoring and Evaluation
MEST:	Ministry of Education Science and Technology
NCLB:	No Child Left Behind
NGOs:	Non-Governmental Organisations
NTES:	National Teacher Evaluation System
OCHA:	United Nations Office for the Coordination of Humanitarian Affairs
OME:	Office of Moral Education
OTA:	Office of the Teaching Affairs
PD:	Primary Document
PMS:	Performance Management System
PSC:	Public Service Commission
RADPS:	Readiness Assessment, Design, Process and Significance
RBM:	Results Based Management
SADC:	Southern African Development Community
SHDC:	School Higher Degrees Committee
SPSS:	Scientific Programme for Social Scientists

TALIS:	Teaching and Learning International Survey
TEPD:	Teacher Professional Development Evaluation
TNTP:	The New Teacher Project
UHDC:	University Higher Degrees Committee
UNESCO:	United Nations Educational, Scientific and Cultural Organisation
VAMs:	Value Added Models
WISE:	Visual, Integration, Serendipity and Exploration
ZIMSTAT:	Zimbabwe National Statistics Agency

CHAPTER ONE

Background and Introduction to the Study

1.1 Chapter Overview

This chapter presents an overview of the study. It gives the general background to teacher evaluation and briefly outlines the education system and teacher performance evaluation in Zimbabwe. The chapter also presents the research objectives and questions of the study and also gives a snapshot of the conceptual framework used to guide the study. Also presented in this chapter are the research design and methodology, ethical considerations, definition of key terms and the research report outline.

1.2 Background to the study

It is the general trend globally that governments are being pressured for improvements and public sector reforms in response to social, economic and political pressures (Kusek & Rist, 2004:1). The pressures emanate from factors that include diminishing public confidence in government, greater demand for more responsive services and greater accountability. There are now greater calls for a results-based culture, accountability, transparency, and enhanced effectiveness of programs from multilateral development institutions, Non-governmental organisations (NGOs), civil society, and even citizens (Kusek & Rist, 2004:1; United Nations Educational, Scientific and Cultural Organisation (UNESCO), 2015:4). The call for results and accountability in the work place has extended to all sectors of government and even to organisations. The education sector has not been spared in the call for better and more responsive services as well as accountability. This study, therefore, sought to assess the effectiveness of the teacher performance evaluation system in Kwekwe district.

One of the most important outcomes expected of all schools is providing a quality educational experience and enhancing learner achievement (Elliot, 2015:102). Teacher quality has been identified as any school's key important resource in propelling the school to achieve its mandate of educating learners effectively (Darling-

Hammond, 2000:33; Looney, 2011:41; Shinkfield & Stufflebeam, 1995:82; Stronge & Tucker, 2003:3; The New Teacher Project (TNT), 2010:1; Wright, Horn & Sanders, 1997:63). It is argued that schools can fulfil their mandate to provide quality education to learners if there is effective teaching. Effective teaching can, however, only be assured by evaluating teachers (Shinkfield & Stufflebeam, 1995:82). It is acknowledged that teachers play a critical role in enhancing student achievement (Santiago & Benavides, 2009:3; TNT, 2010:1), thereby making teacher performance evaluation key in assisting to enhance the quality of teaching if it is properly instituted.

Thus, in response to the call for instructional quality, countries have set up some form of teacher performance evaluation systems. Stronge and Tucker (2003:3) postulate that we can only know that our teachers are effective if we have high quality evaluation systems. Research indicates that a well-designed and properly implemented evaluation system for teachers is critical for a school to be effective (Stronge, 2011:2). Teacher performance evaluation can, and should be considered as a vital part of total improvement restructuring effort in education (Stronge, 2011:4). The concern on teacher quality in many countries has led to the quality of teacher evaluation systems getting under the spotlight unlike before.

However, there is general consensus that most of the current evaluation systems do not do much in terms of improving a teachers' instructional capacity or even supporting decision-making on personnel-related issues (Darling- Hammond, Amrein-Beardsley, Haertel & Rothstein, 2012:8). Donaldson (2009:1) also alludes to the fact that, in the past, most of the teacher performance evaluation systems have failed to either significantly improve instruction nor expand student learning. This is in line with assertions by Danielson and McGreal (2000:3) that teacher evaluation processes are cumbersome and yet unhelpful in spite of the well- intentioned purpose.

It is thus argued that although ideally performance evaluation systems are meant to accurately and consistently measure a teacher's strengths and weaknesses, and in turn, provide helpful feedback, they have often failed to be practically useful (Callahan & Sadeghi, 2015:48). This consequently highlights the issue of why evaluations are failing to improve teaching as well as holding the teachers to account.

Archer, Kerr and Pianta (2014:1), argue that in order for real improvement to occur, quality measurement should be effected. For teaching and learning improvement to occur, teachers should be given high quality feedback based on accurate assessments of their instruction as measured against clear standards for what is known to be effective. There is concurrence, however on the complexity of measuring teacher effectiveness (Taut & Sun, 2014:3). Researchers generally agree that given the complexity of the teaching exercise, no single measurement tool is likely to capture effectiveness of the process (Goe, Bell & Little, 2008:2; Hinchey, 2010:1). It, therefore poses a major challenge to the design and implementation of tools for measuring effective teaching. Another challenge in this domain concerns validity of the measures. Archer, Kerr and Pianta (2014:2), argue that evaluations and feedback should increase the chances that students will learn, grow and ultimately lead healthy, productive lives if they are relevant and effective.

Many factors have been attributed to the failure of the performance evaluation process. Factors that have been cited in this regard include poor infrastructure, lack of resources and incentives for the evaluation process and the lack of evaluative culture at the schools which rarely supports differentiation among teachers (Donaldson, 2009:9). The evaluation design, implementation, and its consequences interact in determining whether or not the evaluation process is successful (Donaldson, 2009:10).

Another challenge facing evaluation systems is that the systems on the ground often show no relationship between teacher effectiveness and student outcomes since almost every teacher is deemed a good performer even at schools where students fail dismally (Callahan & Sadeghi, 2015:48). The usefulness and utility of the evaluation process are assured if the exercise increases the possibility of students learning and ultimately leading productive lives otherwise it becomes an exercise in futility. The challenge with these undistinguishing evaluations is that well-performing teachers may get demotivated when they see their underperforming colleagues receive satisfactory ratings. The carefree approach that leads to the inflation of ratings can also make it difficult to reward teachers with an outstanding performance which can resultantly discourage and demoralise them (Donaldson, 2009:9). It is against this background that this study seeks to establish the effectiveness of the teacher performance

evaluation system in Kwekwe district of Zimbabwe. The contextual background to the study is given next.

1.2.1 Contextual background

In presenting the contextual background of this study, three issues are regarded. These are geographical background of the study area, structure of the education system of Zimbabwe and a background on teacher evaluation in Zimbabwe.

1.2.1.1 Geographical setting of study site

My area of focus for this study is Kwekwe district. Kwekwe district is located in the Midlands province, the province that occupies the central part Zimbabwe. Kwekwe district has a population of approximately 311 556 people (ZIMSTAT, 2012:17). The district capital is Kwekwe town which is located about 220 kilometres south west of Harare (the capital of Zimbabwe) and 230 kilometres to the north east of Bulawayo (the second largest city). The district constitutes of Kwekwe municipality (administered by Kwekwe City Council), Redcliff municipality (Redcliff Town Council), and Kwekwe rural which is administered by Zibagwe Rural District Council.

Kwekwe is located in the agro-ecological region three of the country. This region is characterised by summers that are hot and wet while the winters are warm and dry. The average annual rainfall is 650-850 mm but it is characterised by reduced effectiveness because the rain comes usually as heavy infrequent falls and also due to the high temperatures experienced in the region (OCHA, 2012). The district is endowed with rich gold and iron deposits. There are therefore a lot of artisanal miners in the district.



Figure 1.1 The map of Zimbabwe showing position of Kwekwe district

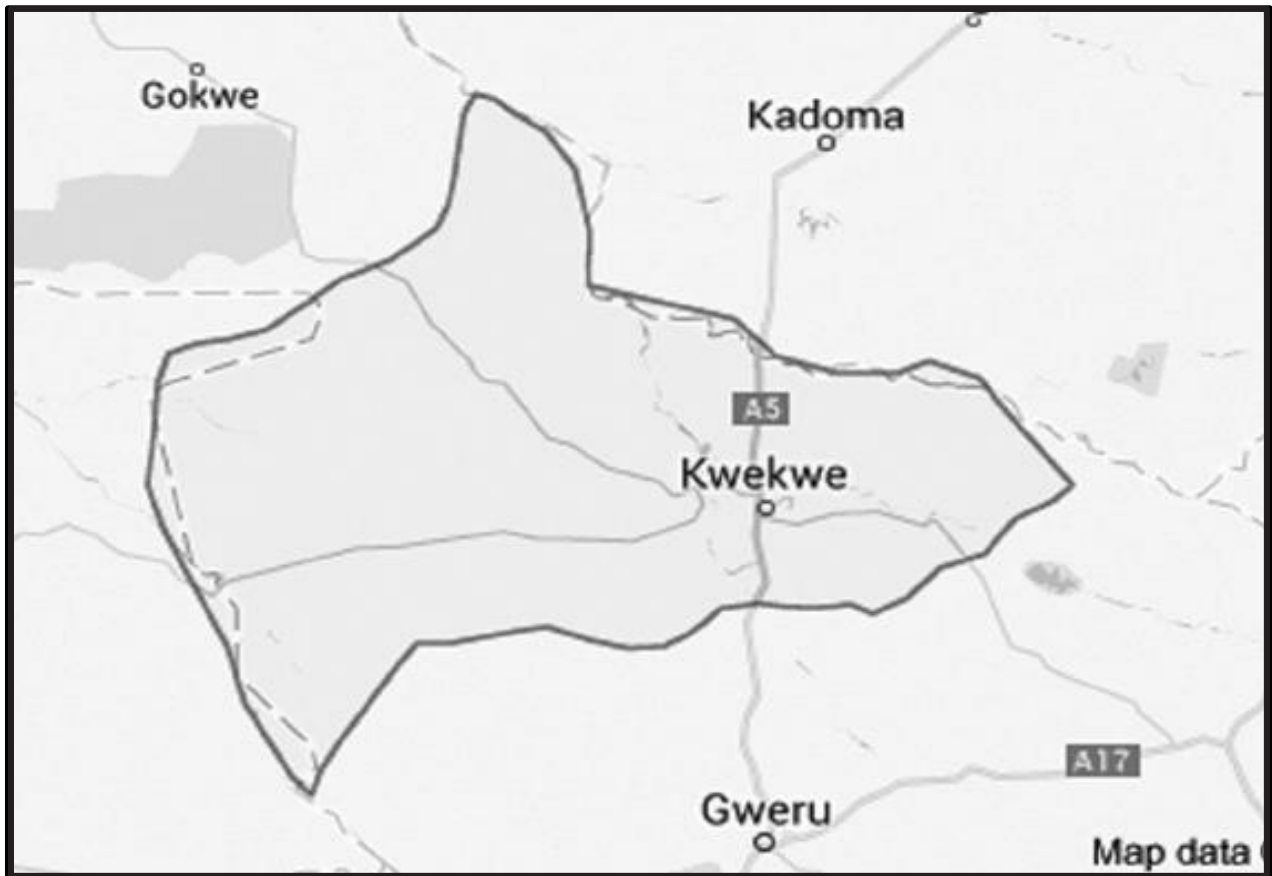


Figure 1.2 The map of Kwekwe District

1.2.1.2 The structure of Zimbabwe's education system

Zimbabwe has two ministries of education which are the Ministry of Primary and Secondary Education and the Ministry of Higher and Tertiary Education. The Ministry of Primary and Secondary education, which is the ministry of focus for this study, has under its wing, the Early Childhood Education and Care, primary and secondary education. The Ministry has a policy which attaches a minimum of two classes of early childhood learning and care per primary school and this phase caters for three to six year olds (National Action Plan of Zimbabwe, 2005:10). Early childhood learning is a two year phase level composed of Early Childhood Development (ECD) A and ECD B.

From early childhood learning, the next phase is the primary level and this has an official entry age of six years. The primary phase extends for a period of seven years at the end of which the learners write a public examination which they pay for. Primary education is compulsory and is guided by the policy of unimpeded progress

(Kanyongo, 2005:67) which makes it mandatory for learners to be promoted from one grade to the next irrespective of performance.

Although a national examination is written at grade seven level, the results of this examination are not used for promotion to the secondary phase. Promotion from primary phase to secondary phase is automatic. The grade seven results are just used for placement purposes especially for boarding facilities. The majority of government schools however take all the pupils for day scholar admission as long as they reside within their catchment zone. The official entry age for secondary education is 13years. Learners also progress automatically in this phase from Form one to Form four. The learners write the General Certificate of Education Ordinary Level examinations at the end of four years of secondary education. The learners pay to write these examinations. A candidate has to pass a minimum of five subjects with grades A, B or C to be considered to have passed this level. A candidate who passes the minimum required five subjects has an option to either proceed to advanced level or go to tertiary level. Those who proceed to tertiary level can do teacher education, nursing, agricultural education and training, polytechnic education, and industrial training and trade testing (National Action Plan for Zimbabwe, 2005:10).

The next and last phase of secondary education which is merit based is advanced level and this is a two year phase. At this level, a learner majors in a minimum of three subjects which determine the career path to be followed at university. The level culminates into an examination that qualifies a learner to enter university on passing. A pass at this level includes the following grades: A, B, C, D, and E. An F symbolises a fail while an O means a fail with 'O' level standard of work. The quality of passes at this level and the subject combinations determine the degree programme that one can pursue.

1.2.1.3 Background to teacher performance evaluation in Zimbabwe

Results based management (RBM) systems have been used by many governments as public management tools to show stakeholders their impacts and outcomes. The Zimbabwe government, likewise, in an effort to show accountability and demonstrable results, adopted RBM as a performance appraisal system for its workers in 2005. Before the adoption of RBM as an appraisal system, other systems had been used.

Zimbabwe has undergone a series of reforms since attaining its independence in 1980, all aimed at improving the civil service (Zvavahera, 2014:1). The reforms include systems such as the Public Finance Management System, Mission statements, Clients Charters and Performance Appraisal System before the introduction of Results-Based Management (RBM) which is currently in use (Matunhu & Matunhu, 2014: 64-65; Zvavahera, 2014:2). Chivore (1992:103) asserts that at independence, the new government of Zimbabwe embarked on an ambitious expansion at all levels of the education system and this impacted negatively on a number of issues inclusive of staffing, resources, and effectiveness of the system. The results of the ambitious project culminated in the Public Service Commission (PSC), which is the employer of all government workers inclusive of teachers, conducting an in-depth study and audit of the Zimbabwe Public Service in 1987 with a goal to recommend strategies to revitalise the system in order to improve the quality of service delivery and responsiveness to the needs of the public (Zigora & Chigwamba, 2000:4).

The study by the PSC resultantly led to the adoption of the Performance Management System (PMS) in 1995 in an effort to enhance the effectiveness of the education system and also make schools more accountable proffered Machingambi (2013:219). The goal of the PMS in the ministry of education was to facilitate the accomplishment of the mission of maintaining or improving the quality of public education and at the same time maintain or enhance staff satisfaction and development through training (Zvobgo, 1998 in Machingambi, 2013:219).

Different reasons have been proffered for the adoption of the Results based Management (RBM) System by the PSC. Zvavahera (2014:2) states that the former performance management system in Zimbabwe failed to yield the desired results of quality service delivery leading to the introduction of the RBM in 2005. Madhekeni (2012:122), on the other hand, argues that there has been growing concerns and pressures from internal and external stakeholders for governments to provide more demonstrable and tangible results globally. A number of international initiatives that include the Millennium Development Goals, Highly Indebted Poor Country Initiative, World Trade Organisation membership, Transparency International, and the European Union have pushed and prodded governments to adopt results-oriented management systems (Kusek & Rist, 2004:3). Madhekeni (2012) further explains that

developing countries have been pressured to adopt the result-oriented management system in an effort to improve performance and uphold accountability as a result of the success story of the RBM system in developed nations.

Following these initiatives, Zimbabwe adopted RBM in 2005. Prior to the adoption, the Zimbabwe Public Service Regulations (2000:12) had specified that all employees shall be appraised annually in an effort to improve services to the general populace. The regulations further indicated that the commission, head of ministry or department shall decide on the action if any, to be taken on the basis of the appraisal. Possible actions would include advancement or promotion; transfer to a post more in keeping with the competence of a member concerned; the granting or withholding of any performance award and demotion or discharge subject to and in accordance with disciplinary procedures provided in Part viii of the regulations. This study, therefore sought to examine the effectiveness of the teacher performance evaluation system in achieving this purpose.

1.3 Statement of the Problem

Literature indicates that job performance evaluation is controversial in terms of objectivity and impact on service delivery. Callahan and Sadeghi (2015:48) posit that, ideally, "... a teacher evaluation system should measure a teacher's strengths and weaknesses through an accurate and consistent process that provides timely and useful feedback". Evaluations should be a means by which feedback and direction for improvement are provided. However, indications are that feedback may be inadequate and not constructive while the professional growth plans may not be aligned to the evaluation objectives (Callahan & Sadeghi, 2015:48; Frase & Streshly, 1994:51 & 52). It is also corroborated that the reality or practice on the ground is that the process rarely improves what teachers do in the classroom (Marshall, 2005:731). It is further argued that the assessments are inadequate, resulting in the failure to address poor performance and failure to recognise teaching excellence. Thus, most evaluation systems have been described as being merely ritualistic and perfunctory compliance exercises that do not distinguish teacher performance nor have any impact on the teaching-learning process (Darling-Hammond, Cook, Jaquith & Hamilton, 2012:1; Hull, 2013:1; King, 2007:1; The New Teacher Project (TNTP), 2010:1). As implied in the Public Service Regulations 1 of 2000 there is a one-size-fits-all

evaluation approach used for all public service employees including teachers. This may affect its ability to accurately measure teacher performance as well as giving high-quality feedback to foster improvement. It is also argued that, in order for the evaluation process to earn sustained support from educators, it should have meaningful implications (TNTP, 2010:8). The evaluation process should be able to produce information that can be used for both formative and summative uses. Evaluation findings should inform career decisions like tenure, development, promotion, compensation and even dismissal (Santiago & Benavides, 2009:19; TNTP, 2010:8).

In view of the above scenario, this study, therefore, sought to assess the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. The plan was to ascertain the effectiveness of the performance evaluation system using the Readiness Assessment, Design, Process, and Significance (RADPS) conceptual framework on performance evaluation effectiveness. The procedure for identifying the aspects of the performance evaluation system that need improvement should be done through stakeholder involvement of the direct users of the system. The stakeholder inputs combined with contributions from literature should inform recommendations to policy makers and government on how the system can be modified or improved to enhance its effectiveness so that it's not merely perfunctory but serves its purpose of improving or maintaining the teaching and learning outcomes.

1.4 Aim of the Study

The aim of this convergent parallel mixed method study was to assess the effectiveness of the teacher performance evaluation system of Kwekwe district.

1.4.1 Research Objectives

The aim of the study was achieved through the following specific objectives that intended to:

- Investigate the relationship between the timing of the introduction of the teacher performance evaluation system and its effectiveness..
- Examine the design process of the teacher performance evaluation system in Kwekwe District.
- Appraise the implementation process of teacher performance evaluation system in Kwekwe district.

- Assess the significance of the teacher performance evaluation process in Kwekwe District in promoting accountability and professional development.

1.5 Research Questions

Since this was a mixed method study, current literature postulates that it is useful to have specific research questions for the qualitative and quantitative strands of the study (Creswell, 2015:60; Ivankova, Creswell & Plano Clark, 2016:331). I therefore had specific research questions for the two strands. The questions for the two strands however address similar aspects of the phenomenon being studied so as to enable comparison of the two data sets for either convergence or divergence as the study used the convergent parallel design. Teddlie and Tashakkori (2009:133) and Creswell (2014:149) also add that mixed method studies should have a mixed method question which illustrates the rationale for mixing of the approaches.

1.5.1 Main research question

How effective is the teacher performance evaluation system of Kwekwe district?

1.5.2 Quantitative Research Questions

I opted to use research questions for the study rather than hypotheses (Creswell 2015:70).

- What is the relationship between the timing of the introduction of the teacher performance evaluation in Kwekwe district and its effectiveness?
- How does the design process of the teacher performance evaluation system in Kwekwe district influence its effectiveness?
- How does the implementation of the teacher performance evaluation system in Kwekwe district influence its effectiveness?
- What is the significance of the teacher performance evaluation system of Kwekwe district in promoting teacher accountability and professional development?

1.5.3 Qualitative research question

- How do educators in Kwekwe district view the introduction of the performance evaluation system?

- What are the educators' views on the effect of the design process of the performance evaluation system in Kwekwe district?
- What are the perceptions of educators on the implementation of teacher performance evaluation in Kwekwe district?
- How does the teacher performance evaluation system in Kwekwe district contribute to accountability and professional development?

1.5.3 The Hybrid or integration research question

To what extent do the qualitative results on the effectiveness of teacher performance evaluation system confirm and corroborate the outcome of the quantitative survey instrument on the views and attitudes of teachers on the effectiveness of the performance evaluation system in Kwekwe district.

1.6 The Conceptual Framework

A conceptual framework was developed to inform the study. The conceptual framework was developed in line with observations by Maxwell (2008:223) that a conceptual framework is the researcher's formulation of what he/ she thinks or perceives is going on with the phenomenon which is being studied. The conceptual framework thus explains what I thought influences and affects the phenomenon under study. The concepts composing the framework were derived from theories, literature, beliefs and personal experience. In developing the framework, a number of sources that include works by Kusek and Rist (2004), Santiago and Benavides (2009), Shinkfield and Stufflebeam (1995), Stronge (1995), and TNTP (2010) were consulted. Theoretical elements from the Systems Theory, Goal Setting Theory and Expectancy Theory are embodied in the conceptual framework. The conceptual framework was named the RADPS framework which is an abbreviation of its components which are Readiness Assessment, Design, Process and Significance. The components of the conceptual framework are outlined in the following section and discussed in greater detail in Chapter two.

1.6.1 Readiness Assessment

Readiness Assessment gives an insight into a country or institutional capacity and political willingness to monitor and evaluate its goals and objectives (Kusek & Rist, 2004:39). It also considers the barriers of sustaining the system like lack of fiscal resources, political will, political champion, expertise, strategy or prior experience

(Kusek & Rist, 2004:43). In this study, the readiness assessment aspect of the framework considered the preparedness of the Zimbabwean Ministry of Education at the inception of the performance evaluation system. The study considered if there was a legislative framework in support of the evaluation process, political will, demand for the evaluation and appropriate institutional culture.

1.6.2 Design of evaluation system

This involves the planning and decision on the 'what', 'how', by 'whom' and 'why' aspects of the evaluation exercise (Santiago & Benavides, 2009:4). The 'what' aspect looks at what the evaluation exercise measures. The evaluation exercise has to be focused on the core activities of teaching (Scriven, 1994:152). Other professional responsibilities besides instructional capacity such as involvement in co-curricular activities and parental engagement are also considered.

The 'how' part involves the articulation of the teacher evaluation model by detailing the strategy used to approach the evaluation process. The mix of instruments, criteria, and standards used are explored at this phase (Santiago & Benavides, 2009:4).

The roles and responsibilities of the different stakeholders in the design and implementation of the evaluation are the by 'whom' part. The 'why' aspect involves the decision on the goals and objectives, and the uses of the evaluation information. These should be articulated in evaluation documents in the form of either policy documents or circulars. This study explored the design process and its strengths and weaknesses and used this to analyse what is done in Kwekwe district.

1.6.3 Process of evaluation

The process involves the implementation aspect of the evaluation process. It is now the enactment of the 'how' aspect. The study considered whether or not the evaluation process is implemented as is articulated in the evaluation policies and guidelines. The study also sought to identify the aspects of the implementation process that strengthened or weakened it.

1.6.4 Significance of evaluation process

Under this component, the study examined the uses of the evaluation findings or the consequences of the evaluation process. There is formative and summative use of evaluation information (Danielson, 2011:36; Stronge, 1995:134-135). The study assessed the consequentiality of the evaluation process. A detailed presentation of the RADPS conceptual framework is given in Chapter 2.

1.7 Definition of Key Concepts

Evaluation: As defined by Olatoye and Aanu (2011:164), evaluation is the act of considering or examining something in order to judge its value, quality, importance, and extent. In this study, it is used to refer to the process of assessing teacher effectiveness and it will be used interchangeably with the word 'appraisal'.

Teacher effectiveness: According to Coe, Alison, Higgins and Major (2014:2), teacher effectiveness is that which leads to student achievement using outcomes that matter to their future success. Teacher effectiveness, therefore, entails strong instruction that enables a wide range of students to learn through meeting the goals of instruction, demands of discipline and needs of students in a particular context.

Performance management: Armstrong (1996), cited in Machingambi (2013:217), describes performance management as, "a means of getting better results from the organisation, teams, and individuals by understanding and managing performance within an agreed framework of planned goals, standards and competency requirements. In this context, it is used to refer to the continuous process of monitoring, evaluating and improving individuals, team and organisational performance.

Results Based Management is described as the process of performance management and appraisal that is results-driven rather than implementation focused, according to Madhekeni (2012:123). It is therefore used in this study to refer to the process of performance monitoring and appraisal.

1.8 Research Paradigm, Design and Methodology

This section gives a general outline of the research paradigm, design and methodology. The detailed discussion of the issues is presented in Chapter 4. The section first outlines the research paradigm.

1.8.1 Research Paradigm

Quinlan, Babin, Carr, Griffin and Zikmund (2011:95) postulates that "Every project is underpinned by a philosophical framework which reveals the worldview in which the research is situated and which can be seen in every step of the research process". This research project in line with the above assertion is located in the pragmatic philosophical worldview. Pragmatism is an inclusive philosophical framework that promotes the use of both qualitative and quantitative research methods in an effort to provide comprehensive answers to the research questions (Greene, Kreider & Mayer, 2005:275). The pragmatic philosophy considers both observable phenomena, and subjective meanings as acceptable knowledge depending on the research question therefore implying that the philosophical framework accords more emphasis on answering the presenting research problems and the consequences of the research (Feilzer, 2010:7) rather than methodical issues and the type of knowledge (objective truth and relative truth) that is derived from the research. Pragmatism is, therefore, a consequence-oriented, problem-centred and pluralistic philosophical approach (Creswell, 2008:11) which acknowledges the existence of singular and multiple realities, is driven in methodological approach by the desire to address the research issues comprehensively, and, produces practical and useful findings.

The rationale behind the philosophical approach is the research questions where the use of either qualitative or quantitative approach alone does not completely address the research problem due to the complexity of the study phenomenon. This is supported by Saunders, Lewis, and Thornhill (2009:109) who argue that choosing between either post-positivism or constructivism is somewhat unrealistic. The aim of this study was not to focus on the methods, but rather, on the potential of those methods to address the research questions adequately in line with arguments by Denscombe (2008:274), Feilzer (2010:14), and Onwuegbuzie & Leech (2005:377). Researchers, therefore, concur that research methodological tools are merely there to facilitate the understanding of phenomenon but should not drive the research which is the belief and approach in this study. Pragmatism has a strong philosophical relevance for mixed method approach used in this study (Johnson & Onwuegbuzie, 2004:14).

1.8.2 Research design

A research design is a functional plan or framework which details the research methodology to be followed in order to acquire reliable and valid data, and also provide the basis for the interpretation of that data (Bertram & Christiansen, 2015:40). This study is premised on the mixed method research design which is underpinned by the pragmatic philosophy (Johnson & Onwuegbuzie, 2004:14). The mixed method approach allows the combining of methods in an effort to facilitate the production of results and findings that are closer to the 'truth' by offsetting biases of error and perspective (Greene, 2008:17). Researchers concur that the collection of multiple forms of data (numeric and text) using different strategies and methods allows the capturing of the best of both qualitative and quantitative approaches thereby cancelling out the 'method effect' leading to greater credibility being placed on the conclusions made (Creswell, 2003:23; Saunders et al., 2009:154).

Considering the universal nature of teacher performance, the mixed method approach was chosen to cater for both generality in quantitative strand and particularity in qualitative strand of the study. The use of the mixed method approach for this study was also motivated by the need for the results of the study to convey magnitude (derived from the quantitative strand) and at the same time portray contextual accounts of lived experiences and perspectives (qualitative strand) thereby giving both breadth and depth to the study (Greene, 2008:7).

The study used the parallel mixed design where the two strands of data were collected concurrently (Van Griensven, Moore & Hall, 2014:368). Data from the qualitative and quantitative strands were collected and analysed separately but at the same time. The two types of data were only merged at the interpretation stage.

1.8.3 Research Methodology

The research methodology explains the general approach to conducting the research by explaining the tools and procedures to be used in the research process (Rajasekar, Philominathan & Chinnathambi, 2013:6). The data collection process, analysis and interpretation procedures helped to provide empirical evidence which supported the conclusions drawn on the phenomenon. The next section outlines the methods that

were used to collect data in the research process. Qualitative and quantitative data collection methods were used in the study.

1.8.3.1 Quantitative data collection method

Quantitative methods make use of numbers and statistics to present data (Rajasekar et al., 2013:9). Quantitative data was collected using close-ended questionnaires in the study. The questionnaires were administered to the teachers to get numerical data to assess the magnitude and frequency of their opinions on the different aspects of the performance evaluation system.

1.8.3.2 Qualitative data collection methods

Qualitative research methods involve the exploration of attitudes, behaviours, and experiences of the subjects so as to get an in-depth perspective of the participants (Dawson, 2002:14). Onwuegbuzie and Johnson (2006:49) reiterate that qualitative methods make it possible to capture the lived experiences of people authentically. Non-numerical data in the form of text was collected so as to get meaning, perspectives, and be able to describe the situation through the lived experiences of the subjects. The qualitative strand, therefore, provided the opportunity to get closer to the participants and get detailed first-hand accounts of their own experiences and perspectives regarding the process of teacher performance evaluation.

Semi-structured interviews were conducted with key informants namely the school heads, heads of departments and teachers so as to get their experiences, perspectives and attitudes as implementers of the teacher performance evaluation process.

1.9 Data analysis

The two data sets, qualitative and quantitative data were analysed and presented separately and merged at interpretation to give breadth and depth to the study. The analysis procedures for the two strands of data are outlined below, and the detailed procedures are presented in Chapter 4.

1.9.1 Quantitative analysis

Data from the questionnaire was transformed into information using statistics to summarise the data, describe the patterns, relationships, and connections. The

Scientific Programme for Social Scientists (SPSS) version 25 was used to do the analysis.

1.9.2 Qualitative analysis

Qualitative data analysis involved the reduction of data to give impressions that shed light on the research questions (Save the Children, n.d:13). Interview data were analysed using ATLAS.ti 8 and pre-set codes that were derived from the RADPS conceptual framework were used. Emerging codes were also considered during the process and these were later used to modify and improve the RADPS framework.

1.10 Sampling process

Since it was impossible to collect data from the whole population, sampling had to be done to come up with a manageable number of participants. Sampling refers to the selection of a segment that is representative of a whole in both qualitative and quantitative studies. Sampling is critical in the research process as it determines the quality of inferences made from the findings (Onwuegbuzie & Collins, 2007:281). The study used a mixed method sampling design comprising of a qualitative and quantitative samples for the different strands of the study.

1.10.1 Study population

The study population constituted 985 educators comprising of teachers, heads of departments and school heads. The educators at different levels of appointment and different roles in the appraisal process were involved in the study in order to get the perspectives of the different stakeholders on the performance evaluation system.

1.10.2 Sampling design

Sampling design refers to the framework within which sampling procedures occur inclusive of the number and types of sampling schemes and the sample size (Onwuegbuzie & Collins, 2007:281). This study used the mixed method sampling design. Teddlie and Yu (2007:85) posit that mixed method designs choose sampling procedures that focus on generating representative samples when addressing the quantitative strand of the study and utilises sampling techniques that yield information-rich cases to address the qualitative strand of the study. A mixed method sampling design was therefore used to cater for the different strands of the study. The mixed

method sampling design used a concurrent approach. The qualitative and quantitative sampling schemes are presented next.

1.10.2.1 Quantitative sampling scheme

The random sampling technique was used to come up with the quantitative sample. The 55 schools in the district were assigned codes and the codes were cut and thrown into a hat where they were randomly picked. The same procedure was done at school level in the selection of the teachers.

1.10.2.2 Qualitative sampling scheme

A sampling scheme gives the strategies that are employed to select the units of a sample like the people, settings, groups etc (Onwuegbuzie & Collins, 2007:281). In this strand, a number of sampling techniques were used to select the different units. Stratified purposive sampling technique was used to select the schools of focus from which the participants were drawn from the district. Stratified purposive sampling entails dividing the population into subsections or groups that are relatively homogenous to one or more characteristics and randomly purposively selecting the case (Collins, Onwuegbuzie & Jiao, 2007:272). The defining classes of schools used were urban and rural. A school that is classified as urban is one that is located in an area that is under the jurisdiction of a municipality or town council whereas one classified as rural is located in an area under the jurisdiction of a rural council. The equal allocation approach (Maree & Pietersen, 2016:196) was used in selecting the elements for each stratum. This meant that each stratum had equal number of elements. I used the purposive sampling technique so as to get the information-rich cases which could provide relevant information and allow me to learn a great deal about the study phenomenon (Patton, 2002:273; Teddlie & Yu, 2007:83; Yin, 2016:93). As I decided on the sampling scheme, I also bore in mind that besides seeking information rich cases who would provide deep and detailed information on the study topic, I also had to ensure that there was breadth of information and perspectives by selecting informants that were likely to have different perspectives for maximum variation (Lincoln & Guba, 1985:233-234). The informants therefore were selected from the teachers, heads of department and school heads. This was aimed at achieving breadth in the perspectives. Defined criteria were also set for the selection of the informants. This was in line with the suggestion by Morgan (2014:127) that a

defined criteria need to be decided for the selection of information rich cases. The informants were supposed to have been in service by 2005 when the evaluation system was adopted by the government, and have knowledge and interest in the issues of performance evaluation. Heads of department and school heads had to be in substantive appointments. This was a way of ensuring that I could tap from their rich experience as evaluators who had used the system over a reasonable period of time.

1.10.3 Sample size

The sample size issue is not prescriptive. Creswell (2015:79) reasons that the different strands of data are meant to tell different stories so the size of the samples do not really have to be similar. The quantitative strand is meant to show the general trends whereas the qualitative strand is meant to give detailed perspectives on the phenomenon (Ivankova et al., 2016:311). This study, therefore, had different sample sizes for the different data sets. The samples sizes for each sample type are given under the relevant strand.

1.10.3.1 The Quantitative Sample

Large sample sizes are recommended in quantitative research as it is presumed that the larger the sample size, the less the possibility of error in reflecting the characteristics of the represented population and in turn, the more accurate the results and findings (Creswell, 2015:76; Dawson, 2002:49). The population was known so I was able to use the sample size determination table by Krejcie and Morgan (1970:608) to determine sample size for this study. The population was 985 and I had a sample size of 292 which was above the recommended sample size of 278 for a population of 1000 (Cohen, Manion & Morrison, 2011:147). This was because I had given out 310 questionnaires as a way of catering for the non-response rate which fortunately was low. This sample was large enough to allow inferential statistics to be conducted on the sample. The respondents were randomly selected, and, since the schools in the sample had different number of teachers, in order to get the correct sample size, I had to randomly select the teachers on a proportional basis relative to the school establishment.

1.10.3.2 The Qualitative Sample

It is acknowledged that qualitative studies have no formula for a defined sample size since the aim is not about representativeness but rather about information-rich cases (Crowe, Inder, & Porter, 2015:617; Yin, 2016:95). Yin (2016:98) further suggests that, instead of seeking a formulaic guidance on the sample size, one should be guided by the complexity of the study topic and the depth of data to be collected from each participant. Because of the complexity of teacher performance evaluation, I therefore considered the capacity of the participant to provide rich deep information on the phenomenon. Ten schools constituting five urban and five rural schools constituted the sample. My qualitative sample had to have a variety of informants in the form of heads of schools, heads of department and teachers for multiplicity of perspectives from the different constituents on the study phenomenon. Four teachers, four heads of departments, and four school heads were purposively selected to be part of the sample. The criterion for all the different informants was that they were supposed to have been in service by 2005 which was the time at which the RBM evaluation system was adopted by the government. The informant was also supposed to be knowledgeable and have interest in performance evaluation issues. The heads of departments and school heads had to be in substantive appointments. I took this as a measure to ensure that I had access to the most experienced people whose experience I could benefit from by getting deep and rich detailed information on RBM and its effectiveness in the district. The principles of 'saturation' or 'selection to the point of redundancy' (Creswell, 2015:77; Yin, 2016:98) and information power were used in determining the sample size. For this study, the sample size was twelve and at this figure, there was generally no new information which I was getting. This was in line with the findings by Guest, Bunce and Johnson (2006:74) that after the twelfth interview there was no new information that was coming up. I reached twelve to ensure that I could get the most possible out of the informants.

1.11 Quality control measures

Quality control measures for the quantitative and qualitative approaches had to be observed. The quality control measures for the two strands are however discussed separately except for a measure that applied to both strands. The two main instruments that were used for data collection in the quantitative and qualitative research strands were both piloted as a quality assurance measure.

1.11.1 Piloting the data collection instruments

The pilot study was conducted in two schools that were not involved in the actual study.

1.11.1.1 Piloting the questionnaire

The close-ended questionnaire was piloted to establish its reliability and validity (Teddle & Tashakkori, 2009:203). Piloting the study presented an opportunity to 'test drive' the questionnaire to allow the identification of potential problems. Two schools were used to get the participants for piloting the questionnaire and the selected schools were excluded from the final study. A total of thirty participants were involved in the pilot. Participants were told about the purpose of the exercise and so besides answering the questions, they were asked to comment on the clarity and relevance of the items, the appearance and legibility of the instrument, the comprehensiveness and length of the instrument (Dawson, 2002:95). Findings and comments from the pilot were then used to review and revise the instrument before the actual study.

1.11.1.2 Piloting the interview schedule

The semi structured interview was used as the data collection method so there was an interview schedule which had to be piloted. Two key informants were used in the process and these were purposively selected for their knowledge and experience on the RBM appraisal system whose effectiveness the study sought to establish. The purpose of piloting the interview schedule was to establish the clarity and relevance of the questions and also to determine the length of the interview to help plan for the actual study. Findings from the pilot were used to improve the interview schedule and to plan for the actual study.

1.11.2 Validity and reliability of Quantitative data

Reliability and validity of the instrument had to be ensured so as to guarantee research rigor for the quantitative stand of the study.

1.11.2.1 Reliability of the questionnaire

Reliability refers to the extent to which an instrument is free from measurement error (Scholtes, Terwee & Poolman, 2011:237). Two forms of reliability, the test-retest and internal consistency were established for this study. The test-retest reliability measures the stability of the instrument or its consistency over time and over a similar

sample (Cohen et al., 2011: 200) whereas the internal consistency reliability measures the degree of interrelatedness between the items that measure the same construct of the phenomenon (Pietersen & Maree, 2016:239). Test-retest reliability of an instrument involved administering the instrument to the same respondents on two occasions (Pietersen & Maree, 2016:238). The questionnaire was administered to the same participants twice during the pilot study and the scores of the two tests were compared to determine the correlation coefficient. The internal reliability measure showed inter-item correlation and was measured using the Cronbach's alpha coefficient.

1.11.2.2 Validity of the questionnaire

Validity refers to the degree to which the scores from the participants are indicative of the construct being measured (Creswell & Plano Clark, 2011:210). Face and content were established for the study. Face and content validity were established through expert judgement by making use of promoters, specialists in the field, and peers. The use of expert advice is supported by Teddlie and Tashakkori (2009:210) when they recommend the use of judgemental validation. They assert that the validity of an instrument can be enhanced by making use of experts in judging its validity.

1.11.3 Ensuring trustworthiness in qualitative data

Trustworthiness as measured through credibility, transferability, dependability, and confirmability is considered a key assurance measure in qualitative studies (Lincoln & Guba, 1985:290). A myriad of measures were taken to ensure trustworthiness of this study and this included the triangulation of data sources and study sites, use of thick descriptions, member checking, debriefing sessions, using the memoing technique and peer scrutiny (Shenton, 2004:64-72; Teddlie & Tashakkori, 2009:295-297). The quality control measures employed in the qualitative strand of the study are outlined next.

1.11.3.1 Credibility of study

A study is described as being credible if it provides assurance on proper and rigorous collection and interpretation of data so as to ensure the accuracy of findings and the conclusions that are drawn on the study (Yin, 2016:85). The credibility of this study

was enhanced through a number of ways which included thick descriptions and triangulation of data sources and study sites.

1.11.3.2 Transferability of study

Transferability refers to the extent to which the reader is able to make connections between elements of the study and their own experience or research process (Nieuwenhuis, 2016:124). Transferability was enhanced providing thick descriptions and selecting typical cases (educators who are using the specific performance evaluation system being studied).

1.11.3.3 Dependability of study

Dependability of qualitative study equates to reliability in quantitative study (Teddle & Tashakkori, 2009:296). In this study it was enhanced by having thick descriptions, member checks, triangulation, debriefing and memoing (Cohen et al., 2011:202)

1.11.3.4 Confirmability

Confirmability refers to the extent to which the study findings are free from researcher bias and are reflective of the participants' views (Nieuwenhuis, 2016:125). Confirmability was enhanced by using member checks, triangulation and memoing in this study.

1.12 Delimitation of the Study

The study was conducted in Kwekwe district of the Midlands Province in Zimbabwe. It focused on ten secondary schools from the urban and rural areas. Study participants were sampled from the ten secondary schools. The study confined itself to the implementers of the performance evaluation system at school level namely the teachers, heads of departments and heads of schools and did not extend to other users at district, provincial and even national levels or even the teachers' unions and parents as key stakeholders. The study could have included all the other stakeholders and districts but it was limited in coverage due to the fiscal and time constraints.

Many factors have been attributed to the failure of the performance evaluation process. This study however limited itself to the factors enunciated in the RADPS conceptual framework on performance evaluation system effectiveness. Concepts

that are constituted in this framework, *Readiness Assessment, Design, Process and Significance* were the focus of the study.

There were other possible stakeholders that could have been involved in the study like education inspectors at both district and provincial levels, parents, teachers' unions and even the learners themselves. However, it was felt that the scope of this study should be limited to the direct implementers thereby leaving room for subsequent studies that could include stakeholders at other levels.

1.13 Ethical considerations

It is acknowledged that although the aim of the study was to find credible answers to the research questions, research findings are acceptable if the participants' well-being has been preserved (Teddle & Tashakkori, 2009:198). It is in line with this position that ethical considerations had to be observed in three broad areas of getting approval to conduct the study, upholding participants' rights and maintaining high standard of researcher's conduct.

1.13.1 Approval for conducting the research

Before embarking on this study, I had to write a proposal that had to go through the School and University Higher Degrees Committees for approval. After the approval of the project, it is procedural at the University of Venda for a research project that deals with either human or animal subjects to go through the Ethical Review Committee before it can be cleared to proceed for field work. This is supported by Teddle and Tashakkori (2009:199) in their position that Institutional Review Boards (IRBs) have the mandate to review and either approve or disapprove a research proposal giving suggestions on how to ensure adequate protection of research participants. On being awarded the ethical clearance certificate, I then had to apply to the gatekeepers for permission to conduct field work starting from the Permanent Secretary of the Ministry of Primary and Secondary Education in Zimbabwe, the Provincial Education Director, the District Schools Inspector, and the Heads of the relevant schools and also get consent from the participants themselves.

1.13.2 Upholding participants' rights

The participants' rights were upheld in this study by seeking informed consent from the potential recruits and respecting their right to privacy. Informed consent refers to

the participant's agreement to take part after being made fully aware of any risks involved and understanding the rights they hold (Teddle & Tashakkori, 2009:199). Since the study was a mixed method study, I had to consider the context and demands of both the quantitative and qualitative research settings (Teddle & Tashakkori, 2009:201). I, therefore, had two consent forms, one for the quantitative strand and the other for the qualitative strand. The need to have two consent forms was also necessitated by the fact that parallel sampling was done. This project was a minimal risk project since not much stress beyond participants' normal experience was expected. Informed consent forms were however generated and given to potential participants to sign as an indication of their voluntary participation in the study.

The privacy of the participants was upheld by limiting my contact with them to the minimum possible times and also to research related issues only (Neuman, 2014:15). Confidentiality was maintained by keeping research data out of the public domain and the identity of participants was anonymised to protect their privacy.

1.13.3 Upholding the integrity of the research profession

I did my best to uphold the integrity of the research profession by using high technical standards during the research process by using appropriate design and methodology for the study. I also took measures to ensure quality by establishing reliability, validity and trustworthiness of the research process.

1.14 Significance of Study

The bedrock of education is teaching and learning, and for the teaching-learning link to subsist, there is need to have effective teachers working in the classrooms with students every day (Stronge, 2011:2). To know if we have the effective teachers, high quality evaluation systems need to be in place (Stronge & Tucker, 2003:3). It is against this background therefore, that this study sought to establish the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. It is hoped that the study findings will be considered by policy makers to enhance the effectiveness of the evaluation system which in turn will translate to the provision of quality education by producing good learning outcomes.

Although some studies have been conducted on RBM as an appraisal system for teachers in Zimbabwe, most of these studies have used the qualitative approach to

study RBM in Zimbabwe and a few have used the quantitative approach. I feel that the complexity of teacher performance requires the combination of the two approaches to adequately explore it for both depth and breadth of findings. This study therefore allowed a more comprehensive approach to the phenomenon thereby filling gaps that are open as a result of the mono-method approach by the previous studies which could somehow have limited the findings and in turn recommendations from these studies .

The other novelty about this study is that from my review of related literature, I developed a conceptual framework which underpinned the study by informing the research questions, research design and methodology and analysis of the findings. The empirical results of the study were used to further refine the RADPS conceptual framework into a model that can be used in the assessment of the effectiveness of teacher performance evaluation systems thereby contributing to the existing body of knowledge.

1.15 Research outline

The study is comprised of 6 chapters which are outlined below.

1.15.1 Chapter 1

Chapter one introduces the study and gives an overview of the thesis. The chapter gives background information to the study, followed by the problem statement before outlining the aim, objectives, and research questions. A brief presentation of the framework is given before defining the key terms of the study. An overview of the research design and methodology is then given, followed by the delimitation of study, ethical considerations, significance of the study and the research outline. The chapter is then concluded with a summary.

1.15.2 Chapter 2

Chapter two discusses the RADPS conceptual framework on performance evaluation system effectiveness. The chapter starts by justifying the appropriateness of a conceptual framework versus a theoretical framework in the context of the study. The components of the conceptual framework and its strengths are presented before explaining its limitations

1.15.3 Chapter 3

Chapter three reviews related literature and place the study within the wider body of literature. Literature was reviewed from both published and unpublished sources, and related studies on the different topics related to the study. This included steps to take before introducing a new evaluation system, the principles of designing performance evaluation systems, the implementation process of evaluation and its challenges, the uses of evaluation findings, conduct of teacher evaluation in selected countries in the developed and developing world, in SADC and in Zimbabwe.

1.15.4 Chapter 4

Chapter four discusses in detail the research design and methodology. All the procedures carried out are presented in a precise and transparent way. The ethical considerations observed during the study are also presented here.

1.15.5 Chapter 5

Chapter five presents the research results and interpretations made on the results. Quantitative findings are presented and discussed in juxtaposition with qualitative findings in this chapter to give a wholesome picture of the opinions, perspectives, and attitudes of the educators on the effectiveness of the teacher performance evaluation system in the district thereby merging the results by comparing and contrasting.

1.15.6 Chapter 6

Chapter six presents the conclusions that are informed by both the quantitative and qualitative findings. The implications of the study for practice and for future research will also be presented in this chapter.

1.16 Chapter summary

Chapter one provided an overview of the study. The background to the study was presented by giving preliminary literature that anchored and provided a useful backdrop to the problem thereby justifying the study. The purpose of the study which was to assess the effectiveness of the teacher performance evaluation system in Kwekwe district was presented and the specific objectives that were used to achieve the aim of the study were also given. Since the study is a mixed method study, the

research questions were presented separately for each of the research strands. This meant that there were separate quantitative and qualitative research questions. The two sets of the research questions however focused on similar aspects of the issue under study although they are presented in quantitative and qualitative form. A hybrid or integrated question that gave the rationale for mixing the research methods was also stated.

An overview of the RADPS conceptual framework on performance evaluation effectiveness was given. Brief outlines of the components of the RADPS conceptual framework were presented. A general overview of the research design and methodology where summaries of the research paradigm, research approach, data collection methods, and sampling procedures were presented. Also presented in the chapter is the delimitation for the study and its significance. A summary of the quality control measures and ethical considerations employed in the study were also described before presenting the research outline which specify the contents of the different chapters of the research report.

CHAPTER TWO

The RADPS Conceptual Framework on Effectiveness of Performance Evaluation System

2.1 Chapter overview

This chapter explains the Readiness Assessment, Design, Process, and Significance (RADPS) conceptual framework on which the study is premised. The first section of the chapter justifies the use of a conceptual framework in framing the study against the use of a theoretical framework. The second section of the chapter then discusses the RADPS conceptual framework by discussing the origin of the framework, the components of the framework and, the strengths of the framework. Limitations of the framework are also acknowledged. The chapter is then concluded with a summary.

2.2 Introduction

This study, which sought to ascertain the effectiveness of the teacher performance evaluation system in Kwekwe district was underpinned and informed by the RADPS conceptual framework. The scope of a study phenomenon is defined and limited through the selection of concepts and variables that depict a structure of related ideas, a conceptual framework. The conceptual framework thus serves to identify the research variables and clarify the relationships among the variables while at the same time informing the development of the specific research questions in line with the described research problem (McGaghie, Bordage & Shea, 2001:923). A study can be framed either theoretically or conceptually, and the rationale for the selection of the frame for this study is discussed next.

2.3 Conceptual framing versus theoretical framing

Theoretical and conceptual frameworks serve the same purpose of framing a research study depending on its context and research process. In this study, I however opted to use a conceptual framework due to the fact that the study was conceptualised and operationalised around four key concepts namely *readiness assessment*, *design*, *process* and *significance* (RADPS). In distinguishing between the two notions that inform research, Imenda (2014:189) describes a theoretical framework as a theory that is chosen by a researcher to guide him/ her in the research process by offering an explanation of an event, phenomena or research problem while a conceptual

framework is described as a product of related concepts that are used to explain or predict a given or give a wider understanding of given phenomenon of interest which is the effectiveness of teacher performance evaluation in this case. A conceptual framework is thus an integration of concepts derived from literature, theory or experience and support each other in providing a comprehensive understanding of occurrence or situation being studied by building up a framework-specific philosophy (Jabareen, 2009:51). A conceptual framework thus presents a researcher's understanding of how the problem was explored, the specific direction that the research took and the relationship between the variables in the study (Grant & Osanloo, 2014:17). From the aforesaid, it is noted that the two, conceptual and theoretical framework, though different in some respects, are interlinked in that concepts feed into a theory and theories also feed into concepts.

A theoretical framework on the other hand is derived from pre-existing theory that is acceptable in scholarly literature and provides a structure to guide and direct the philosophical, epistemological, methodological and analytical approach of a study as a whole (Grant & Osanloo, 2014:13). A summary of the differences between a conceptual and theoretical framework are presented in Table 2:1

Table 2:1 A summary of differences and similarities between conceptual and theoretical frameworks.

	Conceptual Framework	Theoretical Framework
Genesis	-Created by a researcher from a variety of conceptual or theoretical perspectives.	-originates from reviewed literature and /or the data collected. -Adopted/adapted from a pre-existing theory or theoretical perspective
Purpose	-Clarifies variables and concepts in a given study. -Informs research design, target population and research sample, data collection and analysis. -Guides the collection, interpretation, and explanation of data. -Guides future research.	-Clarifies the main variables and concepts in a given study. -Informs research design, target population and research sample, data collection and analysis procedures. -Guides the collection, interpretation, and explanation of data.
Conceptual meaning	-Synthesis of relevant concepts	-Application of a theory as a whole or in part.
Process underlying review of literature	-Mainly inductive.	-Mainly deductive. -can also drive social science research.
Methodological approach	-May be located in both quantitative and qualitative research paradigms and increasingly mixed method approaches. -Data mostly collected through empirical and descriptive survey instruments, interviews and the direct observations. -Strong consideration of context	-Located mainly in quantitative research paradigm. -Data collected mainly through experimental designs, empirical surveys, and tests. -Efforts made to standardise context or else ignore it.
Scope of application	-Limited to specific research problem or context.	-Wider application beyond the current research problem

(Adapted from Imenda, 2014:193)

The following discussion focuses on the strengths of the conceptual framework against the theoretical framework in the context of this study so as to justify the relevance and adoption of a conceptual framework for this particular rather than a theoretical framework. Limitations of a theoretical framework in the context of this study are given

and, on the other hand, the strengths of a conceptual framework as a research frame for this particular study are highlighted. This however does not mean a conceptual framework is uniformly stronger than a theoretical framework in all research endeavours.

One shortcoming of a theoretical framework which made me choose a conceptual framework is that it may at times be limited in addressing a given research problem. Its approach to a phenomenon is not exhaustive and comprehensive enough to allow for an extensive exploration of a complex issue like the effectiveness of a teacher performance evaluation system. This is supported by Maxwell (2008:227) who asserts that, although a theory can help to enlighten and raise awareness to some issues that might otherwise go unnoticed, it can, however, also lead a researcher to overlook certain aspects that are influential to the phenomenon or even the implication of the findings of a study. A theoretical framework thus encourages a systematic approach to research which conforms to the research agenda of the advocates of the theory (Eisenhart, 1991:205; Ngulube & Mathipa, 2015:144). This implies that a theoretical framework may lead to a biased approach to the investigation of a phenomenon in line with the assumptions of a chosen theory. It could therefore limit an inquiry on a research problem as some researchers could get influenced by their single chosen theory. On this note, Imenda (2014:189) observes that most social science studies cannot be meaningfully tackled using a single theory, but rather, different concepts from different theories.

A conceptual framework, on the other hand, allows a more comprehensive and exhaustive approach to the study of the research phenomenon. Jabareen (2009:51) supports this by proffering that a conceptual framework provides a comprehensive understanding of phenomenon through a set of interconnected concepts. The fact that conceptual frameworks are composed of a set of concepts, assumptions, theories and, beliefs as the basis to inform and support research (Casaneve & Li, 2015:107) facilitates their comprehensiveness which is somehow limited where theoretical framing is used. Conceptual frameworks are not deterministic in nature, that is, they do not allow the prediction of an outcome but rather offer an interpretative approach to social reality thereby providing understanding rather than theoretical explanations

Jabareen (2009:51). The strength of the conceptual framework is that, within it, it embodies theoretical aspects, even from a myriad of theories.

One other challenge of using a theoretical framework which made me not to apply it to this study is that it has potential to generate biased research results if not carefully handled, because theories tend to be deterministic in nature as they influence the prediction of an outcome (Jabareen, 2009:51). With a theoretical framework, a researcher can therefore be influenced to explain his/her results according theory at the expense of empirical evidence (Eisenhart, 1991:206; Lester, 2005:459; Ngulube & Mathipa, 2015:9). A conceptual framework on the other hand gives room for different or alternative ways of viewing the data. Conceptual frameworks allow the use of different concepts and theories thereby providing different perspectives on areas we investigate (Silverman, 2005:96-97). This study intends to explain results based on empirical evidence which therefore calls for a wider and more comprehensive research framework in the form of a conceptual framework which is not necessarily limiting as might be in the case of using a theoretical framework.

If not carefully handled, a theoretical framework at times can lead the data to be stripped of its context and local meaning by generalising it to serve the theory (Eisenhart, 1991:206; Ngulube & Mathipa, 2015:9). This possible disregard of context and local meaning has potential to lead to distortions of results and ultimately the conclusions that are drawn. The phenomenon of teacher performance evaluation is not such a simple and straightforward issue such that no one theory can best explain it. The study required strong consideration of the context under which teacher evaluation is understood by different stakeholders thereby making conceptual framing more appropriate to apply.

Studies which are theoretically framed are often limited for use by practitioners within a given discipline as asserted by Eisenhart (1991:206) that theoretical frameworks tend to set standards for scholarly presentations and confines the study within an academic discipline thereby limiting its use for ordinary and practical purpose or day to day use. For example, educational psychologists have their own theories that set boundaries on the research they do. Since teacher performance evaluation is more practical and less abstract, it therefore made more sense to me to frame it using a concepts rather than a theoretical frame.

I feel that the level of complexity of teacher performance evaluation calls for a multidisciplinary approach to its understanding. This resulted in me having to borrow concepts from different disciplines like education, human resources management and, monitoring and evaluation to facilitate a thorough and extensive approach to the study. Conceptual framing is, therefore, more appropriate to the study since it allows the use of concepts from a variety of sources and even across disciplines. The conceptual framing of the study using a multidisciplinary approach would hopefully enhance the usefulness of the study and allow policymakers and practitioners to derive meaning from its findings.

2.4 The origin of the RADPS conceptual framework

Literature indicates that all individuals have a conceptual framework or assumptions of how reality works and this enables a person to predict how X is related to Y or how variables are related to one another and the likely result of the interaction of the two (Svinicki,2010:5). This is in line with the assertion by Maxwell (2008:216) that a conceptual framework explains what a researcher thinks influences or affects the phenomenon under study. In line with these assertions, the RADPS conceptual framework was, therefore, developed by me borrowing concepts from literature. The RADPS conceptual framework on performance evaluation system effectiveness assumes that *Readiness Assessment*, *Design* process, *Process* of implementation and *Significance* of evaluation process are key contributors to the effectiveness of a performance evaluation system. A number of studies that include works by Kusek and Rist (2004), Santiago and Benavides (2009), Shinkfield and Stufflebeam (1995), Stronge (1995), and TNTP (2010), were consulted and concepts were borrowed from these works to come up with the framework for the study.

2.4.1 Components of the RADPS Framework

The conceptual framework developed for this study is composed of four main components, *Readiness Assessment*, *Design*, *Process* and *Significance*. The first letters of each of the components was used to come up with the acronym RADPS which is used to refer to this framework. Each of the four concepts in the framework has a number of variables under it. The concepts in the framework were derived from various sources of literature as follows: Readiness Assessment (Kusek & Rist,

2004:39-49); Design (TNTP, 2010:3-8); Process (Shinkfield & Stufflebeam, 1995:96-97) and Significance (TNTP, 2010:8). Each of the concepts is discussed as it applies to this study in the following subsections.

2.4.1.1 Readiness Assessment component of RADPS

Kusek and Rist (2004:40) highlight the importance of assessing the political, organisational and cultural factors before designing and introducing a results-based monitoring and evaluation system in what they coined 'readiness assessment'. A readiness assessment process is equated to the foundation of a building whose support is critical for all that is above it although it is hidden underground. The readiness assessment is conducted in three steps. The first step involves assessing the demand for the Results Based Management System; the second step involves the determination of the structures that are already in existence and the third step involves ascertaining the capacity building requirement for the evaluation system. The readiness assessment in this study is therefore used to determine if it was feasible to create and introduce the teacher performance evaluation system then, by considering the institutional capacity, political will, and, demand for the system. Readiness assessment involved considering the appropriateness of the timing of introduction of the Results Based Management system as an appraisal system.

In this study, readiness assessment involved the assessment and determination of whether the government and the relevant ministries were ready to build, implement and sustain the teacher performance evaluation system. In this case, the relevant ministries include the Public Service Commission which is the employer of the teachers as civil servants, and also as the ministry that introduced the Results Based Management System as a method of evaluating the performance of civil servants and the Ministry of Primary and Secondary Education as the implementing ministry. The *Readiness Assessment* aspect of the framework considers factors such as the presence of a legal statute to support evaluation, political will, institutional capacity, fiscal resources availability, technical skills and expertise, institutional culture, prior experience, and demands for such a system (Kusek & Rist, 2002:152-153; Kusek & Rist, 2004:40; Lehman, Greener & Simpson, 2002:198).

Since this study was conducted after the system had already been institute, and is already being implemented, the intention of the study was therefore simply to

determine the appropriateness of the timing of the introduction and how this could have influenced the effectiveness of the performance evaluation system. The purpose of this phase of the assessment was to get the educators' perspectives and opinions on the appropriateness of the introduction of the system and explanations of possibly any other steps that could have been taken to enhance the timeliness of the introduction for optimum effectiveness. The different variables under this concept were considered in making this determination.

2.4.1.2 Design process component of RADPS

The design phase in this study refers to the period of planning the performance evaluation system. This is a critical phase which greatly influences the process and the significance of the teacher performance evaluation process. Clear and defensible policies and guidelines should underpin the evaluation system if it is to be accepted, effectively used and legally viable (Shinkfield and Stufflebeam, 1995:59). In this study, the design aspect is assessed by considering factors such as policy formulation which enunciate the goals and objectives of the evaluation process, the purpose of evaluation, the scope of the evaluation, the roles and responsibilities of the different stakeholders and the significance of the evaluation process. This phase of the framework also considers the evaluation criteria and standards, and evaluation instruments and tools.

This study examined the extent of educators (at school level) involvement and participation in the design process of the evaluation system. According to Isoré (2009:9), a number of stakeholders should be involved in the design phase in order to promote buy-in and credibility. This is also stated by Shinkfield and Stufflebeam (1995:59) and Kusek and Rist (2004:58) who advocate for the collaborative development of an evaluation system. The Government, local authorities, school administrators, educational researchers, experienced teachers and teachers' unions can be involved in the design process. Government plays an important role in the conception of evaluation schemes since they are the ones mandated with the setting of the national objectives and learning outcomes. Isoré (2009:8) adds that the efficiency of an evaluation system is judged by the extent to which it achieves the national goals and objectives emphasizing the role of government in the design process.

Researchers and teachers may partake and contribute to the design process as experts who would know what 'good' teaching entails and so can help in the formulation of appropriate criteria and instruments to be used (Isoré:2009:9). Teachers' unions can also have an input as the representatives of teachers while in cases where the process is decentralised, then, the school administrators may also be involved. The involvement of these stakeholders in the design process is important so as to increase buy-in and ownership and at the same time lessen resistance (Heneman, Milanowski, Kimball & Odden, 2006:6); Mayne, 2007:3). This is also supported by Riordan, Lacireno-Paquet, Shakman, Bocala, and Chang (2015:6) who affirm that stakeholder support can be enhanced during implementation by involving them in the design phase.

The design phase of the evaluation should also entail the formulation of goals and objectives, the purpose of evaluation and the scope of the evaluation through an evaluation policy. The scope of evaluation should clarify if the evaluation procedure is to apply to the whole country or can vary according to province or district. Aspects like the schools and teachers to be covered under the system and whether the evaluations are mandatory or voluntary or both are also considered (Isoré, 2009:10; Santiago & Benavides, 2009:12). This study sought to find out if there is a policy that regulates the appraisal process and what the policy stipulates in terms of scope and purpose of the evaluation process. The awareness levels of the educators on the appraisal regulatory framework were also established by the study.

The design phase of the framework also entails the specification of roles and responsibilities and the consequences of the evaluation exercise. The evaluation process should produce information that is used to inform important decisions like teacher tenure, compensation, development, promotion, demotion, withholding of promotion and dismissal (TNTP, 2010:8). This study sought to establish if there are clearly stated consequences of the evaluation process and clear job descriptions as part of the regulatory framework to the evaluation process.

Since evaluation process should be consequential, it is important that evaluation findings are accurate, clear and easy to interpret. The teacher evaluation process needs evaluative criteria and standards to evaluate teachers against what is considered as 'good' teaching for the process to be regarded as fair, objective and

reliable (Ingvarson, 2002:4; Isoré, 2009:11; Santiago & Benavides, 2009:12; TNTP, 2010:4). Danielson (2011:36) supports this by advocating for a shared understanding of the definition of 'good' teaching which is enunciated through clear criteria and standards. The criteria describe the performance indicators while the standards involve the determination of levels of acceptable performance. Santiago and Benavides (2009:13) assert that the main reference standards for teacher evaluation are teacher profiles which describe competencies for teachers, job descriptions, the school development plan, internal regulations and annual activity plan. It is indicated that the process by which the standards are developed is critical for the validity, operationalisation of assessment procedures and legal defensibility (Ingvarson, 2002; 5). The tools are decided and developed during the design phase and the guidelines or manuals on how they are used. This study examined how the design process was executed and the level of stakeholders' involvement. It was one of the aims of the study to establish whether the educators at school level were involved at all in the design process and their views on how the design process could have impacted on the effectiveness of the whole system.

2.4.1.3 Process component of RADPS

The process aspect of the RADPS framework involves the scrutiny of the implementation process and the determination of effectiveness of the process from the perspectives of the educators. The study sought to establish the challenges that hinder successful implementation and possible measures that can be used to mitigate the hindrances. The process aspect encompasses the setting and agreement of goals between evaluator and evaluatee; assessment of performance against pre-set criteria and standards; and functional reporting. The evaluator and evaluatee should set and agree on objectives at the onset of the evaluation cycle. The objectives should be aligned to the job description of the evaluatee and the institutional goals and objectives. The study also sought to understand the procedures involved in the goal-setting process and the basis on which the goals are formulated.

It is important that the teachers are evaluated against clear performance standards that are based on student learning (Coe et al., 2014:39; TNTP, 2010:4). The standards should be clear so that both the teacher and evaluator interpret them in the same way.

This implies that standards should be worded clearly so that there is no room for inference (TNTP, 2010:4).

Clear evaluation criteria that are consistently applied by trained and experienced evaluators helps to ensure that the process is fair, reliable and credible. The use of trained evaluators who are able to use instruments and apply criteria consistently will enhance validity and the reliability of the evaluation process thereby increasing acceptability of evaluation findings (Danielson, 2011:36; Haefele, 1993:26; Shinkfield & Stufflebeam, 1995:55). Another aspect to consider on evaluators is the people to be involved in the evaluation process which has to be decided at the design phase. The idea of having multiple evaluators is suggested rather than relying on a single evaluator (Hinchey, 2010:11; Santiago & Benavides, 2009:17). This helps to enhance the rigor of the evaluation process.

Clear guidelines are also required to facilitate proper assessment procedures. This study had to establish the extent of usefulness of the evaluation guidelines based on their clarity and preciseness. The availability of the guidelines in the schools was also checked. Guidelines that are clear and concise coupled with evaluators with a high inter-rater agreement promote reliability and credibility of the evaluation process (Haefele, 1993:26).

Evaluation instruments and information sources should be multiple if the system is to be rigorous (Santiago & Benavides, 2009:14; Stronge, 1995:138; TNTP, 2010:6). In trying to establish the effectiveness of this evaluation system, it was also important to find out if a variety of sources are used in the process. Multiple sources of information can present a complete picture of a teacher's performance and increase the validity of the evaluation process so a rigorous system should use multiple measures to judge the teacher's performance against the set standards (Hinchey, 2010:7; Stronge 1995:138, TNTP, 2010:6). The advantage of using multiple tools is that one can offset weaknesses of another. Isoré (2009:2) indicates that types of data gathering instruments used have potential to influence results. It is therefore critical that the choice of instruments be critically considered at the design phase. The decision to choose particular evaluation instruments should be guided by their validity and reliability (Goe et al., 2008:13). Literature proffer that an evaluation instrument can be assessed against three standards namely deficiency, contamination and distortion of

measures (Haefele, 1993:24). Measurement deficiency refers the extent to which the evaluation instrument excludes some criteria considered to be important and representative of teaching. Measurement contamination refers to the inclusion of a performance criteria that is irrelevant to the teacher performance while the measurement distortion considers the extent to which the performance criteria are weighted disproportionately relative to their true importance in teaching. The instrument that was considered in this study was the reporting instrument. The educators' perspectives on the reliability and validity of the instrument were sought in the process of establishing the influence of the implementation process on the evaluation system effectiveness. There were no standard data collection instruments across the different schools therefore the study focused on the reporting instrument.

Although, literature pays particular emphasis on issues of validity and reliability, other measurement concerns that have been suggested in connection with measuring tools include comprehensiveness of the measure; generality; utility; practicality; credibility and reliability (Blanton, Sindelar, Correa, Hardman, McDonnell & Kuhel, 2003:14). The study also sought to find out the educators' opinion on the comprehensiveness of the reporting instrument. There are different domains to teaching (Danielson, 2013:4) and this study wanted to establish if the coverage of the reporting instrument is inclusive of all the aspects.

The study also examined the forms used to report the performance of a teacher to establish the multiplicity of the ratings. These forms have to be standardised and have multiple ratings (Santiago & Benavides, 2009:14). The multiplicity of ratings on the instruments determines the ability of the ratings to distinguish low performing from high performing teachers which contributes to the effectiveness of the system. TNTP (2010:7) recommends that the rating scales should be at least two levels above and two levels below the expected level of performance so as to be differentiating.

If an evaluation system intends to promote professional development, then, a culture of improvement-oriented observations should be cultivated by having regular and frequent observations and feedback. TNTP (2010:8) argues that an evaluation system should not be limited to a single rating given at the end of the year, but rather, have regular formative evaluations so that there is regular feedback and the final rating is acceptable. The main purpose of feedback is to reduce discrepancies between current

understanding, performance, and goal (Hattie & Timperley, 2007:86). This is supported by the argument advanced by TNTP (2010:8) on the need for frequent and regular feedback if the feedback is to serve a purpose. If feedback is to be constructive and facilitate improvement, then, it should be accurate, clear and timely (Frase & Streshly, 1994:51; Shinkfield & Stufflebeam, 1995:55). This study examined if there is functional reporting in the evaluation process enabled by accurate, clear and timely feedback.

2.4.1.4 Significance of the evaluation process in the RADPS framework

This aspect of the framework considers the implications or consequences of the evaluation process in gauging the significance of the evaluation exercise (TNTP, 2010:8). The study examined if there are any documented uses of the evaluation results in the evaluation policy or any other official document on the evaluation process. If they are specified uses, are these followed up or it's mere rhetoric? This study, therefore, tried to establish if the evaluation process is significant or if it's merely a perfunctory, ritualistic exercise without any consequence.

Literature states two possible uses of teacher performance evaluation results namely formative and summative uses (Centre for Development and Enterprise, 2015:4; Danielson, 2011:36; Isoré, 2009:15-17; Marzano, 2012:15; Santiago & Benavides, 2009:19; Stronge, 1995:134-135; Williams & Engel, 2013:53). The formative or improvement-oriented evaluation aims to identify the strengths and weaknesses of a teacher so that an improvement plan for the concerned teacher can be developed as necessary. The teacher improvement plan should, however, be aligned with the overall school development plan (Santiago & Benavides, 2009:19). Research findings, on the contrary, show that few countries link professional development agenda to the identified weaknesses of the teachers' reviewed performance. Lack of this important link makes the evaluation process inadequate to improve teacher performance thereby rendering the exercise meaningless (Isoré, 2009:16; Santiago & Benavides, 2009:19; Tuytens & Devos, 2014:511;). This study examined whether there is a connection between the weaknesses identified during the evaluation process and if there are any individual plans available that are linked to the identified weaknesses. It was also within the scope of this study to find out if the professional development plans are implemented or if the evaluation is done merely for compliance purposes.

Evaluation results can also be used for accountability or summative purposes to inform decisions on career advancement, performance related awards or sanctioning underperforming teachers (Danielson, 2011:36; Isoré, 2009:15; Santiago & Benavides, 2009:19). It is however noted that in most Organisation for Economic Cooperation and Development (OECD) countries, there is single salary schedule which does not cater or acknowledge good teaching practice. This has raised concern about the attractiveness of teaching as a profession and the likelihood of keeping teachers within the profession motivated (Isoré, 2009:15). Isoré also adds that there are some countries that link evaluation results with promotions to administrative positions. This has, however, been criticised for removing the good teacher from what he or she knows best, that is, by removing the teacher from the classroom. If evaluation results are consequential and used especially to inform high stakes decisions, it is important that they are accurate, clear and easy to interpret (TNTP, 2010:6). It should, however, be noted that an evaluation process that is merely ritualistic and inconsequential is not credible and therefore can be a waste of resources in terms of time and finance which therefore justifies a look into this aspect of the evaluation process. This study, therefore, examined the use of the evaluation results to determine the consequentiality or significance of the evaluation process with special focus on accountability and professional development functions.

To sum up the discussion on the components of the RADPS conceptual, a diagrammatic representation of the structure of the framework is presented in Figure 2.1.

Effectiveness of Performance evaluation system

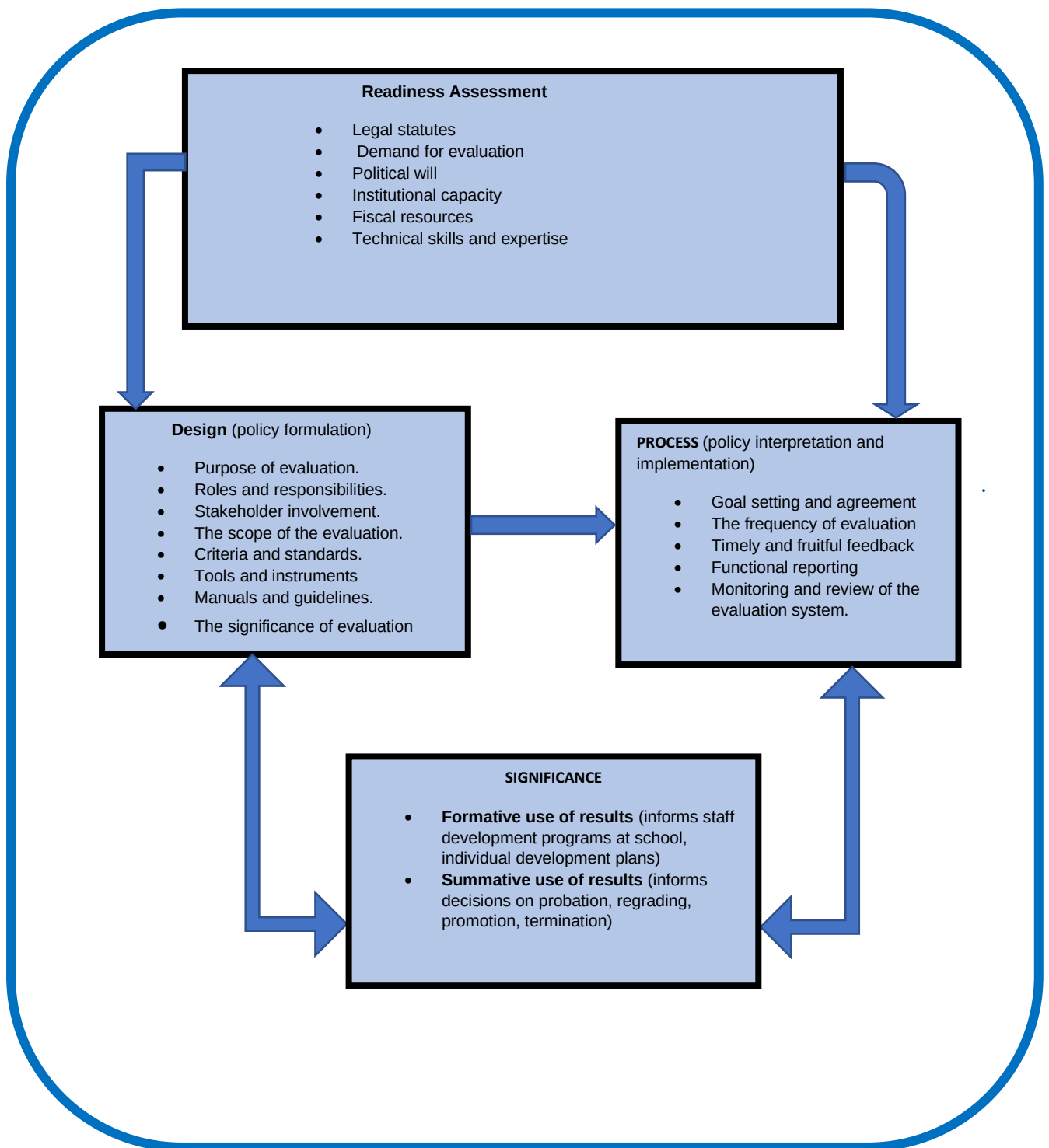


Figure 2.1 Components of the RADPS conceptual Framework (Adapted from Kusek & Rist, 2004:40 & 58; Santiago & Benavides, 2009:12-19; Shinkfield & Stufflebeam, 1995:55& 59; TNTP, 2010:4-8)

The RADPS framework shows how the Readiness Assessment influences the Design and Process components of the system. The Design affects both the Process and the Significance aspects of the system. The Process also influences the Significance of the system and Significance aspect which in turn influence both the Design and Process aspects by feeding back to them for improvement purposes. It is also wise to discuss the strengths of the RADPS framework in spite of its still being works in progress which is what the next section deals with.

2.5 Strengths of the RADPS conceptual framework

This section discusses the strengths of the RADPS conceptual framework. The major strength of the RADPS conceptual framework is its comprehensiveness. The framework provides a comprehensive perspective on the assessment of teacher performance evaluation. The conceptual framework covers aspects on readiness assessment (an assessment to determine the degree to which it is feasible to introduce the evaluation system); design (the process of planning and developing policies, guidelines, and manuals to guide the evaluation process); the actual process and the significance of the process. It describes the complex relationship between the different concepts of teacher performance evaluation. The framework, therefore, provides a holistic view of the whole system since it looks at the process of readiness assessment phase through to evaluation phase when the significance of the process is considered.

The second strength of the RADPS framework is its duties-based approach. The conceptual framework emphasises the evaluation of teachers directly on the fulfilment of their duties rather than on personality traits. This is in line with Scriven's (1994:152) proposition of duties-based approach to teacher evaluation. Scriven (1994) came up with a list of generic 'core duties' for teachers which he categorised as knowledge of subject matter, instructional competence and skills, assessment skills, professionalism and other secondary duties to the school and community. The RADPS framework in the same vein also emphasizes the importance of assessing teachers against job-related standards. For the evaluation process to judge actual performance, it is only logical that it should be based on one's core duties.

The third strength of the RADPS framework is its orientation towards both improvement and accountability functions which are in line with propositions of current

research on purposes of evaluation (Centre for Development and Enterprise (CDE), 2015:4; Santiago & Benavides, 2009:8; Stronge, 1995:131; Williams & Engel, 2013:53). Current research proposes that evaluation should serve both functions to be effective although it is acknowledged that combining the two functions has its own challenges.

The fourth strength of the RADPS framework is that it has theoretical elements embedded in it and the following theories: systems; goal setting; and, expectancy theories have been identified to have assumptions that are similar to the framework. The next section explains these theories and how their assumptions are related to the framework.

2.5.1 Theoretical elements embedded in the RADPS conceptual framework

Research literature indicates that there has been effort to provide a conceptual framework to facilitate a better understanding of the mechanisms by which performance evaluation processes are framed, influenced and produce reactions (Arvey & Murphy, 1998 cited in Ferris, Munyon, Basik & Buckley, 2008:153). Although conceptual frameworks have been useful in some respects, it is however noted that they are limited by their failure to provide discussions of the theoretical underpinnings driving the process dynamics, weighs in Ferris et al (2008:153). Ferris et al (2008) further espouse that it is important that in framing performance evaluation processes and outcomes, scholars should not only identify the key components of the context but the factors that predict the formation of such contexts. They explain that “there is need to give sufficient theoretical grounding to the conceptualisation to explain the interactive process dynamics and outcomes” (Ferris et al. 2008:153). In view of this proposition, I, therefore, decided to examine the RADPS framework for theoretical underpinnings supporting its assumptions. The RADPS conceptual framework exhibits propositions and assumptions of a number of theories including the systems, goal-setting and expectancy theories.

2.5.1.1 The Systems Theory

The assumptions of the systems tradition include holism and the environment (Dent & Umpleby, 1998:514). The systems approach proposes a holistic view rather than a reductionist view of phenomena as is exhibited by the RADPS framework. Holism assumes that phenomena should be viewed in its entirety rather than by dividing it into

separate components. The holistic perspective is similar to the approach proposed by the RADPS framework in that the framework considers the system in totality although it gives room for a step by step approach in the evaluation of a teacher performance evaluation system. A system is a set of interrelated elements where all the elements are interconnected to each other either directly or indirectly and operate as one unit (Senge, 2006 in Lunenburg, 2010:1). The system's parts rely on each other for the successful functioning of the whole system, that is, they are in a symbiotic relationship. Cook and Tonurist (2016:9) proffer that "systems are elements joined together by dynamics that produce an effect, create a whole or influence other elements and systems". A malfunction in one part will thus affect the effective and efficient functioning of the whole system. It should, however, be noted that interactions, relationships, and patterns of connectedness among the system components are not static but are dynamic and change from time to time. This assumption is similar to what is exhibited by the RADPS conceptual framework on performance evaluation system effectiveness. The four components of the framework, *Readiness Assessment*, *Design*, *Process*, and *Significance* are interconnected and a malfunction in one component of the framework just like in the systems approach will affect the effectiveness of the whole system. This means that in the case of teacher performance evaluation system as shown in the RADPS framework, for example, if the *Design* component is not properly done, then this affects the effectiveness of whole system and also influences other components within the system like the *Process* and *Significance*. This applies to the rest of the components within the framework. The symbiosis of components in the systems approach is also evident in the RADPS framework.

It is logical, therefore, that a performance evaluation system is studied in its entirety as proposed by both the systems theory and the RADPS framework. The conceptual framework identifies the readiness assessment, design, processes, and significance as the different parts of the evaluation system. These parts are interrelated and interconnected and so they have to be studied holistically for meaningful and informative insights to be construed. Systems are complex and dynamic and applying the systems' lens to complex situations is useful for understanding the systems dynamics. A systems approach just like the RADPS conceptual framework facilitates a comprehensive view of the system and allows adequate and exhaustive exploration

of the components and acknowledges the dynamism of interactions, relationships and patterns of connectedness thereby informing the better judgment of its worth and effectiveness. In the conceptual framework like in the systems theory, the environment or context is acknowledged to be significant. The context, which in this case is determined by readiness assessment, is presented to have an effect on all the other parts of the system. The RADPS framework therefore exhibits similarities to the systems theory through the interconnected components of the framework and the critical role of the environment as emphasized in both the framework and the theory.

2.5 .1.2 The Goal Setting Theory

The goal setting theory proposed by Edwin Locke and Gary Latham in the 1960s has some assumptions that are also similar to the RADPS conceptual framework. The Goal Setting theory suggests that all things being equal, people will perform better if they strive towards a definite goal than if they are merely expected to perform without specific objectives in mind. Locke and Latham (2002:705) define a goal as the aim of action expressed as a specific standard of proficiency usually within a specified timeline. This theory concurs with the RADPS conceptual framework for the study. The common positions are that enunciating goals enhances commitment to the achievement of the goal. The RADPS framework asserts that the evaluator and evaluatee should set and agree on objectives at the onset of the evaluation cycle. This resonates with the Goal Setting theory which assumes that goal setting prior to an action will help to drive an individual towards attainment of the goal. Locke and Latham in the same article postulate that goals have a directive and energising function towards performance. This is similar to the assumptions by the RADPS conceptual framework where goals are believed to steer action in a particular direction, that is, being directive, and to stimulate and motivate effort towards a target, thereby energising the individual. The RADPS framework and Goal Setting theory thus both assume that that pre-setting of goals or objectives assist in driving and motivating an individual towards a specified target thereby enhancing performance. The process of goal setting is therefore key in both the RADPS framework and the Goal Setting Theory where they give direction to action by evoking deliberate planning and development of strategies towards achievement of the set goals (Smith, Locke & Barry, 1990 in Locke & Latham, 2002:706).

The Goal setting theory like RADPS conceptual framework values continuous feedback on one's progress towards achievement of the goal. Feedback avails information on the discrepancy gap between the current performance and the expected performance (Hattie & Timperley, 2007:89). The formative aspect of the evaluation in RADPS helps to ensure that there is regular and continuous feedback on performance which promotes improvement.

2.5.1.3 The Expectancy Theory

The third theory that has propositions that concur with some of the assertions of the RADPS framework is the Expectancy Theory proposed by Victor Vroom in 1964. This theory assumes that the tendency by an individual to act in a particular manner depends on the strength of an expectation that the act will be followed by a given outcome, and on the attractiveness that the outcome has to the individual, (Bargrain, 2016:118). This links well to the Process and Significance elements of the RADPS conceptual framework, where in the Process an evaluatee is required to set goals and exert effort to achieve them being motivated by the expected outcome. The Significance aspect of the framework like the expectancy theory, emphasizes the importance of consequentiality or rewards as a way of motivating performance. Where teacher evaluation is ritualistic and of no consequential results, there will be no expectation and effort-performance relationship is negatively affected, thereby reducing the motivation to act towards the attainment of the set goals.

The RADPS conceptual framework assumes that in order for evaluations to be effective, they should be consequential (TNTP,2010:8). Evaluations that are merely ritualistic and perfunctory will diminish the reward-performance link thereby demotivating the evaluatee resulting in reduced productivity and effectiveness. Both the RADPS framework and the Expectancy Theory thus assume that linking performance to reward will maintain high effort levels in the workforce.

2.6 Potential limitations of the RADPS conceptual framework

The limitation of the espoused framework is that it is still 'works in progress' since it was developed from literature, experience, as well as belief regarding how teacher performance evaluation systems should be assessed. There is however a lot of room for improvement through validating the framework from empirical evidence. The empirical evidence from the study findings were used to improve and modify the

framework and the improved model is presented in Chapter 6 of the research report. It was however rational and appropriate to use the RADPS conceptual framework in this study in its current form based on the strengths it has, which have been mentioned earlier in this chapter, thereby justifying its use as a framework for this study.

2.7 Chapter summary

This chapter discussed the RADPS conceptual framework for effectiveness of performance evaluation system. The chapter was introduced by describing what a conceptual framework is before distinguishing a conceptual framework from a theoretical framework. The rationale for using a conceptual framing rather than a theoretical framing in this study was discussed. This was done by pointing out the strengths of using a conceptual framework compared to using a theoretical framework in assessing the effectiveness of the teacher performance evaluation system. An outline of how the conceptual framework was developed was given before presenting a detailed description of the RADPS framework and its components. The strengths of the RADPS framework and the theoretical elements embedded in it were also discussed. It is because of its strength that the framework has been used to guide the research processes. The chapter was concluded by acknowledging the limitation of the framework as a result of its status of being ‘work in progress’.

CHAPTER THREE

A review of related literature

3.1 Overview

This chapter presents a review of related literature and studies. A detailed explanation of the term ‘performance evaluation’ is given before comprehensive discussions of three selected country evaluation systems (South Korea, Chile and China) are made, to give a global perspective of teacher evaluation. Literature is then reviewed on the basis of themes derived from the RADPS conceptual framework before discussing teacher performance evaluation in Zimbabwe under the same themes. Lessons are drawn from the three countries’ evaluation systems to relate to the Zimbabwean evaluation system and the chapter is then concluded with a summary.

3.2 Introduction

Teachers are considered the single most important school-based factor in improving student achievement. However, teacher evaluation has remained a controversial topic because of the ongoing debate about what an effective teacher is, and does, and failure to reach consensus on a general method for measuring teacher effectiveness (Goe et al., 2008:2; Santiago & Benavides, 2009:3). Regardless of this debate, educational standards can be raised by ensuring that teachers are skilled, well-resourced and motivated to perform at their best. Teacher evaluation is ultimately key to continuous improvement of teaching effectiveness in schools (Santiago & Benavides, 2009:3). Evaluations give teachers feedback and promote their professional growth in spite of their teaching experience, provide schools with information they can use to build strong teaching teams, and help districts to hold school leaders accountable for the development of each teacher (TNTP, 2010:1). The evaluation processes are however often labelled as a meaningless, ritualistic and perfunctory exercise (Isoré, 2009:5). The purpose of this study was therefore to determine the effectiveness of the teacher performance evaluation system of Kwekwe district based on the RADPS conceptual framework for teacher evaluation.

3.3 The concept of performance evaluation

Evaluation, in general terms, has been defined as a systematic process of assessing the worth or merit of an object (Joint Committee, 1994 in Stufflebeam & Coryn, 2014:8). This definition implies that there is judgement to determine value by assessing worth and merit. Stufflebeam and Coryn (2014:8) explain that merit refers to the quality of an object being evaluated and it is determined by the ability of the object to achieve its intended purposes. In the case of performance evaluation, this implies the ability of the system to judge performance accurately. However, merit is limited to the excellence of the object regardless of costs (Stufflebeam & Coryn, 2014:8), while worth, on the other hand, refers to an object's combination of excellence and service in relation to need and costs. Worth considers the need of the object within a specified setting or context and time. It involves the assessment of quality in relation to contexts and costs. The definition implies that for a study to be termed an evaluation it has to report how good or bad something is.

Stufflebeam and Shinkfield (1988:4) add that some writers argue against the position of value orientation, which is accused of reducing objectivity of the process and being a negative force. Stufflebeam and Shinkfield (1988) use this point to emphasize the complexity of the evaluation process and therefore the need not only to carefully select high quality information but also to clarify and provide a defensible rationale for the value perspectives used to interpret findings. The value aspect implies that evaluation is indeterminate and difficult to do well because of the potential contradictions in the value that can be given to an object. To counter this negative assumption on the effect of 'value', Stufflebeam and Shinkfield (1988:4) propose that evaluation should be based on the shared and differential needs of the clients of a given service. This should be used as the basis for both determining the information to collect and the standards to invoke in determining the merit or worth of a service. Stufflebeam and Shinkfield (1988:5) further advanced that the value accorded to an object should be based on important sources of criteria that include:

- i) stakeholder expectations,
- ii) The merit or excellence of the service in question,
- iii) The extent of the need for the service or its potential worth.

This highlights the importance of stakeholder buy-in which can be achieved by stakeholders' involvement at every stage of the evaluation process from design, implementation through evaluation of the process and its products. The efficacy of the system and its relevance should be considered in the judgement process.

The definition above is more or less similar to the one proffered by Rogers and Badham (1992:3) which describes evaluation as a process that involves the collection and analysis of information in a systematic manner to be able to form value judgements which are evidence-based. The judgments are based on prior set targets and the extent to which these targets have been achieved. The qualities of judgements that are made are influenced by the nature and source of the data that is collected as evidence. The issues of value judgement and standards are common in the two definitions emphasising the complexity of the process.

Since this study is concerned with the teacher performance evaluation process, a definition specifically on teacher evaluation would be justified and appropriate. Shinkfield and Stufflebeam (1995:86) define teacher evaluation as a systematic process of assessing the performance of a teacher, based on a defined professional role, and the institutional goals and mission. All the three definitions highlight two important issues, that the process should be systematic and have criteria to inform judgment. The aspect of basing evaluation on criteria in the form of standards is further supported by other studies (CDE, 2015:3; Danielson & McGreal, 2000:20; Santiago & Benavides, 2009:13).

3.4 Global perspective of evaluation

The increased demand for greater accountability and real results has necessitated the need for enhanced evaluation policies (Kusek & Rist, 2004:ix). It is against this background that this study sought to determine the effectiveness of the teacher performance evaluation system in Kwekwe district in ensuring teacher quality and ultimately, quality education. To get a detailed and wide perspective on the subject, the teacher performance evaluation systems of South Korea, Chile and China were closely examined for their unique features on which lessons could be drawn and applied in other contexts such as Zimbabwe. The unique features include best practices on readiness assessment, design process, implementation process and

significance of process which could benefit other countries in designing and implementing effective teacher performance evaluation systems.

3.4.1 Teacher evaluation in South Korea

Teacher quality has become a core educational issue globally, since research has shown that the quality of students' education is highly dependent on teacher quality (Choi & Park, 2016:153). In order to provide students with the best quality of educational experience, the establishment of efficient and effective teacher evaluation systems has therefore become a priority in many countries. South Korea has not been spared either in that effort to uphold educational quality by ensuring teacher quality. Apart from the global developments, South Korea had domestic pressure emanating from the distrust of the public education system by parents. The level of public dissatisfaction reached a critical point in what was referred to by the media as the 'school or classroom collapse' of 2004 (Kang, 2013:153). This was shown by the wide engagements of private tutors by those parents who had the financial capacity to do so (Yoo, 2018:279). The engagement of the private tutors by the South Korean parents implied lack of confidence in the quality of teachers in the public education system which therefore necessitated the need to develop a valid and reliable teacher evaluation system as a measure to ensure teacher quality which in turn would help to restore parents' confidence in the system.

Since the 1960s three teacher evaluation systems were implemented namely, Teacher Performance Ratings for Promotion System, Teacher Performance-based Pay System and Teacher Evaluation for Professional Development (Choi & Park, 2016:153). The Teacher Performance Rating for Promotion was the first to be adopted in 1964 with a purpose of informing decisions on promotion (Kang, 2013:157). The system evaluated the teachers weighing their experience more than their professional capacity thereby failing to encourage teachers to seek professional development in instructional competence (Choi & Park, 2016:153; Kang, 2013:157; Yoo, 2018:280). The validity of the system was thus compromised as it focused on teaching experience and ignored the young and competent teachers thereby failing to hold teachers to account in classroom instruction as well as failing to reward instructional excellence (Yoo, 2018:280). The other limitation was that the evaluation criteria were unclear and the evaluation results were also not open for review thereby making the teachers

mistrustful of the process. The teachers were also indifferent to the process because there was automatic salary increment based on teaching experience (Kang, 2013:149).

In view of these problems, a new system, Performance-based Pay was then introduced in 2001. It aimed at generating constructive competition among the teachers and in turn offered monetary rewards commensurate with their efforts (Choi & Park, 2016:153; Yoo, 2018:280). The system however failed because the monetary rewards were insignificant and not motivating enough to teachers who already had high salaries coupled with assured job security (Yoo, 2018:280). This led to the introduction of the Teacher Professional Development Evaluation (TEPD) system in 2011 with the objective to avail teachers with corrective feedback on their instructional practice thereby developing their professional competency which would in turn improve the quality of public education (Choi & Park, 2016:153; Seo, 2012:75).

3.4.1.1 The development of the Teacher evaluation for professional development (TEPD)

TEPD was developed in response to the public call for school reform, disgruntlement with the promotion regulation and the OECD report on shortcomings of teacher policy in South Korea (Kang, 2013:159). A bottom-up approach was adopted to try and develop a new evaluation policy in 2003, where all stakeholder groups were involved in the design. However, the different stakeholder groups failed to reach consensus due to their different positions on teacher evaluations resulting in the Ministry of Education Science and Technology (MEST) drafting two versions of teacher evaluation systems that were partially agreed to by the different stakeholder groups.

MEST proceeded to pilot the two versions of the teacher evaluation system in 2006 with partial support of the teachers' unions. The development of the system went on for several years until the presidential diktat of teacher professional development regulation was amended for implementation on a national scale in 2010. It should however be noted that the new system was an addition to the already existing teacher personnel system (Kang, 2013:161). This therefore meant an increased work load for the implementers.

3.4.1.2 Features of the evaluation system

The TEPD is unique in that unlike the Evaluation for Promotion System that was conducted by the principals or vice principals and teachers, it involves all the members of a school community in that even the students and the parents also participate in the evaluation process (Seo, 2012:75; Yoo, 2018:278). The other notable feature of the evaluation system is that it was sensitive to local contexts and shifted from a national unitary approach (Kang, 2013:162).

Evaluation purpose

The purpose of the system is to diagnose professional capacity in order to be in a position to provide the relevant and appropriate support for professional capacity development to facilitate the provision of quality education (Choi & Park, 2016:164; Kang, 2013:163; Seo, 2012: 75). Improving the quality of education is intended to restore public confidence in the public education system.

Evaluation targets

TEPD targets all teachers who have taught for at least two months in schools that are governed by the Elementary and Secondary Schools Act (Kang, 2013:163). It is discretionary on the schools to decide if teachers occupying temporary positions are evaluated.

Evaluation participants

The participants in the evaluation process includes the principal or vice, master teacher, three other teachers, students in grades 4 to 12 and all the parents (Choi & Park, 2016:164; Kang, 2013:164; Seo, 2012:75). The system thus has multiple evaluations conducted by multiple evaluators.

Evaluation content

Teachers are evaluated on primarily academic instructional practice and student advising. The system is open such that other aspects like aptitude can be added. Although evaluation items are articulated, schools are free to modify them to suit their context.

Evaluation methods

An evaluation constitutes three aspects:

- i. Five scale Likert type questions
- ii. Details of excellence
- iii. Details of areas of improvement (Kang, 2013:164).

Schools have an option of giving parents online surveys to increase response rate or giving the paper surveys. Parents' surveys are conducted after parents have been provided with information on educational activities through brochures, video clips of classroom lessons or opening up the classroom for observation.

Evaluation timeline and process

Evaluations are done once a year and the process should be completed at least three months before the end of the school year (Kang, 2013:164). The three months period is to ensure that adequate time is available for reporting the results to the participants and give them time to plan for the relevant professional development in the following year.

An evaluation committee is to be established and the purpose of the committee is to develop appropriate evaluation rules and develop a clear evaluation plan (Kang, 2013:165). The evaluation procedures are introduced to the rest of the participants who are the teachers outside the evaluation committee, the students and the parents as enunciated in the evaluation plan. Detailed information on different aspects of evaluation is availed to the teachers, students and parents in brochures of different versions that are appropriate for the target group. Parents are then invited to observe different school activities and classrooms are opened up for them after the introductory procedures to allow the parents to be informed enough to complete the survey. The evaluation process determines the professional capacity of the teachers and in turn identifies their professional development needs. Every teacher produces a professional development plan based on self-assessment and integration of the peer teacher evaluation and feedback from the students and parents. The school evaluation committee aggregates the results for all the teachers in the school and produces a school report that is submitted to the local educational agencies and available for view by the teachers and parents.

Use of evaluation results

Teachers with excellent ratings can be awarded 6 to 12 months sabbatical to allow them to further their professional capacity through research. Teachers with low ratings in the peer assessments and student survey have to take 60 hours of professional training in selected teacher training institutions and those who receive ratings that are lower in peer assessment and in student surveys are mandated to take 210 hours of professional development over a period of 6 months (Kang, 2013:167; Seo, 2012:75; Yoo, 2018:278). If these teachers fail to improve their ratings the following year, they are required to take leave of the classroom for 6 months to take 730 hours of professional training at the National Training Institute of Education, Science and Technology.

3.4.1.3 Criticism of the Teacher Evaluation for Professional Development

The teachers have levelled a plethora of criticisms against the evaluation system. The first being that whilst TEPD supposedly emphasizes on professional development, it however focuses more on accountability since it identifies and rewards effective teachers whilst at the same time ‘punishing’ less competent teachers by asking them to take mandatory professional training (Seo, 2012:76; Yoo, 2018:282). Teachers argue that the system focuses more on requirements rather than support. Mandatory training is regarded as humiliating and shameful so the professional development may not have a positive outcome because teachers perceive it negatively, as a punishment. This has resultantly led to a situation where there is a tendency by the teachers to inflate scores for colleagues in an effort to avoid embarrassing them by making them go through the forced retraining. The evaluated teachers have also become unenthusiastic to openly discuss their weaknesses due to the high-stake nature of the evaluation process.

The second criticism is that there is indistinctness and lack of consensus among evaluation players on key elements of the evaluation process like a definition for ‘effective teacher’ (Seo, 2012:76; Yoo, 2018:283). This argument is that without training or guidelines that are clear and precise, the evaluation players may have varied understanding of such terms. It was also noted that since the performance indicators, standards of performance or even the evaluation criteria had simply been thrust on the stakeholders without them being involved in their development,

consensus had never been established. The ambiguity and lack of consensus led teachers to lose confidence in the evaluation process. Many of the teachers have no respect for the student and parent evaluations as they regard them to be highly subjective and mere popularity contests.

The third censure is on the unreliability of the sources of evidence which in turn compromises the validity of the process. Although the peers assessors and parents were encouraged to do classroom observations on more than one occasion, in most cases they just made a single visit. Both the evaluators and the evaluatees contend that a single observation is not sufficient enough to evaluate a teacher accurately or assist the teacher to improve instructional practice consequently rendering the evaluations unreliable and invalid (Seo, 2012:76; Yoo, 2018:284).

3.4.1.4 Cases of success with TEPD

In spite of the criticisms, there are some schools that have registered success with this particular evaluation system. The successes of these schools present possible ways to improve the national system.

Case 1: Hangang Middle School

After realising that the national evaluation questions and rubrics were failing to identify the teachers' strength and weaknesses, the school developed their own evaluation system that they used in combination with the national system. They also built a shared understanding of the standards of performance and a common definition of 'effective' teaching. This helped to elucidate what is expected of a teacher and therefore what to look for in an evaluation which in turn helped the evaluators to give more information-specific feedback that the teachers found informative and useful for improving their teaching (Seo, 2012:77; Yoo, 2018:285).

Case 2: Namsan High School

This school adopted a collective responsibility approach by using the collaborative technique to improve student learning. They developed a peer review system that is collaborative in nature where a team of teachers collaboratively diagnosed individual students' learning needs and then planned lessons together. The collaborative and collective responsibility approach by the teachers, parents and students themselves

resulted in successful modification of the new evaluation system for the teachers' professional improvement and the student's academic achievement in the national assessments (Seo, 2012:78; Yoo, 2018:285).

Case 3: Sejong Elementary School

The school realised that most parents were not participating in the evaluations process as less than half of them were involved and based the teachers' ratings on their children's opinions. Teachers therefore did not understand the rationale for the ratings they were given and what they meant (Seo, 2012:78; Yoo, 2018:286). Teachers adopted an approach to engage and involve the parents more by having more regular communications with them and enlisting their support in cultivating student learning. This resulted in an increase in parents' participation in the evaluation process and improved rating ability that provided more informative feedback on teachers' instructional practice (Seo, 2012:78; Yoo, 2018:286). The evaluation process has thus managed to build a shared vision and values between teachers and parents, cultivated a collaborative approach to student learning and developed competencies for teacher evaluation (Seo, 2012:78), leading to improved professional development and student achievement (Yoo, 2018:286).

3.4.1.5 Criticism of the three South Korea Evaluation Systems

Three evaluation systems (Performance Rating for Promotion, Performance-Based Pay System and Teacher Evaluation for Professional Development) are applied in South Korea. This occurred because every time when it was realised that the system in use had challenges, instead of reviewing , modifying or removing it, another system was added alongside it until in the end there were three separate systems operating independently. Choi & Park (2016:167-168) noted some concerns on the systems and offered suggestions on addressing these. One of the worrying issues is that the three systems have often produced differing results for an individual which resultantly raised issues of reliability and fairness eroding confidence in the teacher evaluation systems. The issue of having to satisfy the three independent systems has also been taxing on both the teachers and the administrators in terms of the time and workload. It is therefore necessary to integrate the three separate systems into a single system which has multiple functions for both formative and summative evaluations for efficiency and effectiveness.

The absence of specific rubrics for the performance standards has made it difficult for evaluators to assess the instructional practice clearly, accurately, consistently and fairly. The absence of rubrics also leaves teachers in the dark on how to improve their teaching practice and student learning. The development of detailed and tangible rubrics for each standard can enhance the effectiveness of the evaluation system.

Choi and Park (2016: 168) also suggest that establishing performance standards that are consistent with student learning is a prerequisite to setting up a new system as the universal goal of teaching is student learning. They argue that standards that are related to student learning would assist the teachers to support the students rather than focusing on their instructional practices only. A possible way to determine the effectiveness of instructional practices would be to use Value Added Models (VAMs) which can estimate the contribution of an individual teacher to students' growth. They proposed using alternative forms of student assessment other than Value Added Models like portfolios, essays, pre and post-test measures to try and establish the teacher's contribution to student growth. They also suggest inclusion of standards for teacher collaboration to foster shared responsibility for student achievement to the benefit of both the student as well as the teachers through sharing teaching strategies and skills that help to enhance their expertise.

3.4.1.6 Lessons drawn from South Korea

A number of lessons were drawn from the case of South Korea as follows:

- In a case where parents lack confidence in the public education system like what happened in the case of South Korea until around 2004 when effort was made to restore parental confidence, those parents who can afford will always seek an alternative way in the form of private tutors. This is also the situation in Zimbabwe although the government has outlawed private lessons by public teachers. This is indicative of lack of confidence which creates further inequalities on access to quality education based on social class.
- The case of South Korea proved that although stakeholder involvement in the design of a performance evaluation system can be challenging, some degree of involvement can still be achieved in order to facilitate buy-in.
- Students and parents participate in the evaluation process thereby making it all encompassing.

- Classrooms are opened up for observation by the different stakeholders like the parents and this enhances their engagement and contribution to the evaluation process.
- A militaristic approach may not be very appropriate in civilian spheres as shown by the attitudes of the teachers towards the mandatory training.

3.4.2 Teacher evaluation in Chile

The management of Chilean schools was formerly, under the military government (1973-1990) and was transferred to municipal authorities in 1980 (Isoré, 2009:35). The move meant that teachers ceased to be civil servants to become municipal employees and a change in the conditions of service. The teachers' conditions of service however failed to grow in line with those of the public servants, negatively impacting on the teachers' self-perception and that of the public opinion.

By 1990 the working conditions of the teachers dramatically deteriorated and this culminated in poor quality applicants applying for teaching until they got in the limelight of the Chilean policy makers because they were assumed to be the cause of the dissatisfying student achievement in the national standardised tests (Avalos-Bevan, 2018:298). The disappointing students results led the government to add a clause in the new Teacher Statute of 1991 that mandated teachers to be evaluated yearly (Isoré, 2009:36). Teachers however spurned the implementation of the evaluation system although they continued to call for improved working conditions and salaries. A long period of debate and negotiations then ensued.

3.4.2.1 Design and implementation of the system

The discussions on the evaluation system had started in 1999, and it culminated in the enactment in August 2004 (Avalos-Bevan, 2018:297; Isoré, 2009:36). A settlement had been reached to have a mandatory mixed formative and summative evaluation system conducted every four years unless a teacher has underperformed in the previous evaluations (Santiago, Benavides, Danielson, Goe & Nusche, 2013:45). The approach of evaluation to adopt had been one of the sticky points, the policy-makers and educational authorities favoured the accountability function which allowed them to identify underperforming teachers and remove them from teaching while on the other hand, teachers' unions favoured the formative approach which identified the strengths and weaknesses of a teacher and provided appropriate professional

development opportunities. Although the Chilean evaluation system included both functions, it however placed greater emphasis on the formative function. The aim of the system was to improve teaching and learning outcomes by motivating teachers to identify their strength and weaknesses in pursuing their growth while at the same time incorporating an accountability focus linked to the dismissal of teachers exhibiting unsatisfactory levels of competence over three consecutive evaluations (Avalos-Bevan, 2018:299). The evaluation process is premised on a shared definition of effective teaching agreed on after wide consultation with the teachers and this culminated into a clear criteria of what is to be evaluated (Avalos-Bevan, 2018:299; Isoré, 2009:36).

3.4.2.2 The actors in the system

Although the system is managed by the Centre for In-service Training in the Ministry of Education, there is a consultative committee constituted of academics, teachers' union representatives, the Chilean Association of Municipalities and the Ministry of Education (Avalos-Bevan, 2018:298; Isoré 2009:36). The consultative committee is responsible for monitoring and providing advice on the process while the implementation, production and revision of the evaluation instruments, the selection and training of the evaluators and the scorers and the analysis of the evaluation results is the jurisdiction of a contracted University Measurement Centre (Isoré , 2009:36; Santiago et al.,2013:45). The Measurement Centre is also mandated to maintain a permanent research agenda of assessing the validity of the teacher performance evaluation system (Santiago et al., 2013:45). Each district however has a committee mandated with the organisation of the evaluation procedure and the processing of the evaluation findings before sending them together with contextual information to the central processing unit at the university. The logic of having a central processing unit was the idea of teachers to enhance objectivity of the process.

3.4.2.3 Evaluation criteria

The evaluation criteria, called the Framework of good teaching, have a set of standards based on Danielson's Framework for teaching of 1996. The Framework has twenty standards based on four teaching domains namely planning, learning environment, professionalism and teaching strategies (Avalos-Bevan, 2018:300; Isoré, 2009:36). The framework is composed of statements that describe what the

teacher should do relative to each of the four teaching domains. Four levels of performance which are *unsatisfactory*, *basic*, *competent* and *outstanding* are linked to the evaluation criteria.

3.4.2.4 Evaluation instruments

The instruments that are used to gather evaluation evidence are:

- Portfolio of samples of teacher's work and a video of one lesson.
- A self-evaluation form that is structured.
- Structured interview conducted by peer evaluator with at least five years of teaching experience.
- School management and pedagogical authorities report (Isoré, 2009:36; Santiago et al., 2013:46).

The instruments have different weighting in the evaluation of a teacher's competence level. A portfolio weighs 60% of the total score, followed by a peer interview with 20% and the two other sources with a weighting of 10% each (Avalos-Bevan, 2018:300; Santiago et al., 2013:51).

3.4.2.5 Evaluators

The teacher performance evaluation system has multiple evaluators which are the Municipal Evaluation Commission, markers of portfolios, school leadership, peer evaluators and the teacher being evaluated. The Municipal Evaluation Commission is responsible for taking the decision on a teacher's final rating by reviewing and ratifying the results of the different assessment instruments as well as considering the background information of the concerned teacher (Santiago et al., 2013:51). Peer evaluators are practising teachers who have been previously rated as Competent or Outstanding and these are selected annually after lodging an application for the function with the Centre for Pedagogical Training, Experimentation and Research (CPEIP) in the Ministry of Education. They are expected to be familiar with the context of the school where the teacher being evaluated is stationed, are trained and are tested before they can be accredited.

Portfolio markers are practising teachers with at least five years of teaching experience in the same teaching area and level like the teacher being evaluated (Santiago et al., 2013:52). They are trained and are provided with detailed assessment rubrics which give descriptors for *Unsatisfactory*, *Basic*, *Competent* and *Outstanding* performance in the different areas covered by the portfolio. Teachers are expected to be active participants in their own evaluation by producing a performance portfolio and carrying out self-evaluation both of which are expected to promote a reflective approach to instructional practice in the teacher.

3.4.2.6 Consequences of the evaluation process

This is a key aspect of the evaluation process that prolonged the debate process. The evaluation system has both formative and summative functions. The summative element is however divorced from the students results and career ladder while the professional development is emphasized and differentiated on the basis of the teacher's level of performance (Isoré, 2009:37). Teachers rated at both *basic* and *unsatisfactory* levels are provided with specific professional development opportunities, with those having an unsatisfactory rating, being re-evaluated again a year later and those with a basic rating re-evaluated after two years. This means that a teacher with an *unsatisfactory* rating has up to two consecutive chances and one with a *basic* rating up to three consecutive chances before being dismissed from the municipal education system (Avalos-Bevan, 2018:301). The professional development plans are targeted for the development needs identified during the course of the evaluation and these are designed and implemented by the municipal education authorities and inspected and reviewed by the CPEIP for relevance, timeliness and effectiveness (Santiago et al., 2013:54).

Teachers with competent and outstanding rating levels are given priority in promotion opportunities and also receive non-monetary incentives (Avalos-Bevan, 2018:301; Isoré, 2009:37). They are also illegible to apply for a salary bonus which is awarded after taking the Voluntary Individual Grant Assignment (AVDI), a test that measures content and pedagogical knowledge. If the teacher is successful in the test, he or she is re-categorized in relation to the prior rating category and the monetary bonus award varies from 15% to 25 % of the basic salary depending on the performance rating. The bonus is paid over a four year period (Avalos-Bevan, 2018:304). Teachers with the

competent and outstanding ratings also have four years before the next evaluation. Another evaluation, the Assignment for Pedagogical Excellence (AEP) was introduced in 2002 and this has been used to reward teachers who have high levels of subject matter and pedagogical knowledge. Teachers in public funded schools are illegible for this evaluation and are awarded a salary bonus over 10 years depending on the results which can be classified as *outstanding*, *competent* and *sufficient*. Indications show that younger teachers have successfully taken the examination than more experienced teachers (Avalos-Bevan, 2018:304).

3.4.2.7 Evaluation Reports

Different versions of evaluation reports on the teacher performance are availed to different constituencies. These include the evaluated teacher, the Municipal Evaluation Commission, the individual schools and the municipal evaluation authorities (Santiago et al., 2013:54-55).

3.4.2.8 Strengths and shortcomings of the system

The Ministry of Education of Chile participated in the OECD's Review on Evaluation and Assessment frameworks for improving Schools' Outcomes programme whose results were released in 2013. The OECD review and information collected from different sources and actors, such as, surveys for evaluated teachers, peer evaluators, and portfolio raters, online questionnaires for local authorities have identified some strength and limitations in key aspects of the evaluation system which include the following.

The design process

One of the keys strength of the Chilean teacher evaluation system concerns its design process. The system was designed in a participatory approach where all the key stakeholders were actively involved. Although it took about ten years to reach a consensus, it was worth the while since the involvement of teachers' unions minimised teachers' resistance and promoted buy-in (Santiago & Benavides, 2009:27). The National Teacher Evaluation System policy articulated the intended purpose, uses and effects of the evaluation process. This identification and disclosing of the underlying theory of the evaluation system helped to facilitate its monitoring and evaluation as time went on which was another key strength of the evaluation system (Taut & Sun,

2014:22). It has also been noted that the final decision regarding the teacher's final performance rating is the responsibility of the local community which thus takes the process back to the local level thereby allowing them to consider the relevant contextual information which could otherwise be missed in a standardised approach (Taut & Sun, 2014:24).

The evaluation framework

The evaluation framework, the anchor of the evaluation system is regarded as valid by the teachers. An identified weakness of the framework is that it lacks distinctive descriptions of the kind of performance that would be particular for each performance level. This limits the teacher's understanding of his/her specific strengths and weaknesses that lead to their evaluation rating level (Avalos-Bevan, 2018:305). This is worsened by the fact that the rubrics that are used when assessing the teachers are not availed to them. In response to this indication, the Ministry of Education is reviewing the system and has sent out a new version of the framework for consultation with the teachers. The new version contains the performance descriptors for each of the four assessment levels (*outstanding, competent, sufficient and unsatisfactory*) for each of the standards of the four domains of the framework (Avalos-Bevan, 2018: 305).

Taut and Sun (2014:22) also note that, the fact that the National Teacher Evaluation System (NTES) is standards-based ensures that expected instructional practices are clearly spelt out thereby providing a basis for provision of detailed feedback to the teachers on their strengths and weaknesses. They further reason that a system that is based on student achievement might manage to identify the effective teachers and ineffective teachers but on the other hand fail to explain the reasons for the results which ultimately limits the impact of the evaluation process as a professional development instrument. They proceed to emphasize that in order to improve teacher quality, it is necessary that evaluation should improve classroom activities which in their view the NTES has managed to provide evidence to that effect through its informative feedback on instructional practice.

Quality of evidence provided by some instruments

The self-evaluation instrument has been criticized for not being a valid input to the overall evaluation of a teacher while the school principal's assessment has come under scrutiny for having been given limited weight in the overall evaluation of a teacher (Avalos-Bevan, 2018:306). Principals' rating had been given limited weighting because of the mistrust that teachers had in the principals because of their appointing authority which was the military government. Taut and Sun (2014:23) concur that self-assessment is not really valid. They assert that it is biased and has no reflective value in the face of the high stakes assessment. They suggest that it should rather be used for formative purposes.

The OECD review team also queried how the different parts of the portfolio contribute to the overall assessment of the teacher. The team felt that the instructions of completing the portfolio were too precise thereby limiting the range of the teacher's practices that could be included over a period of a year (Avalos-Bevan, 2018:306). Teachers also had no confidence in the ability of the portfolio to show their instructional competence. Taut and Sun (2014:24) on the other hand argue that out of all the National Teacher Evaluation System instruments, the portfolio has proven to be the most technically robust of all as evidenced by its results (the lowest among all the four instruments), its scoring process and its link to students achievement.

The OECD review team also raised concerns on the lack of live classroom observations since they supposedly provide good evidence of the teacher's performance.

Professional development opportunities

Although no direct feedback has been received from the teachers on the usefulness of the mandatory professional opportunities, it can be inferred that the low number of teachers who have been dismissed could mean that professional development has been effective. It has however been argued that the remedial approach to professional development is more correctional rather than promoting growth and diversity in teaching (Avalos-Bevan, 2018:306). Taut and Sun (2014:23) posit that the way the professional development plans are implemented in Chile exhibits a disconnection between them and successful teacher development as described in literature since

they are not integrated into the teacher's daily work, are not sustained over time and are not performed within a community that supports learning. Taut and Sun proceed to suggest that instead of professional development plans being perceived as formal, administrative and disparaging responsibility for under-performing teachers, they should rather be perceived as imperative and needful learning opportunities. The OECD team recommended a communities approach to professional development and for this to be spearheaded by the school leadership (Avalos-Bevan, 2018:306).

On the other hand, the fact that outstanding and competent teachers are free to choose professional development plans such as formal courses, seminars and workshops or internship in other countries could be indicative of a system favouring the 'accountability' concept of professional development rather than the formative approach (Avalos-Bevan, 2018:306).

Accountability policies and the teacher evaluation system

The teacher evaluation process was meant to provide information on the professional competence of teachers so that those teachers who were working hard could be identified and rewarded while at the same time identifying and dismissing the under-performing teachers (Avalos-Bevan, 2018:307). The purpose of the process was defined in response to disappointing results of students which had been attributed to the teachers. The teachers' unions had however defended the teachers by apportioning the blame to poor working conditions, inadequate initial training and inadequate professional development opportunities. The evaluation system therefore had to reflect these differing views thereby leading to the development of a system that had both the formative and summative functions. The OECD review team however noted that tension exists between the formative and summative functions of the system because of the nature of the instruments used and the processing of the information that is collected. The system has been criticised for failing to recognise the range of teaching practices that teachers engage in because of the structured nature of evidence presented in the portfolio which in turn limits the possibility to provide detailed feedback on instructional practice of the teachers. The system is also blamed for failing to create space for professional dialogue so as to promote self-reflection by the teachers on their practice.

The Chilean experience has on the other hand proved that combining formative and summative evaluations is feasible although it involves making complex and challenging trade-offs, for example, the idea of objective self-reflection may be far-fetched in a high stakes evaluation process (Taut & Sun, 2014:23). The OECD review team points out that the formative function of the evaluation system has been subsumed by the accountability function of the system (Santiago et al., 2013:170)

3.4.2.9 Development of a new teacher evaluation system

The Chilean experience has thus demonstrated that even after careful design and almost 10 years of implementation, any programme can still benefit from both internally and externally motivated research (Taut & Sun, 2014;25). A review and modification of some of the elements of the teacher evaluation system is therefore underway. This was necessitated by some issue that were raised by the Teaching and Learning International Survey (TALIS) conducted in 2013. These included the issue of low salary levels, high number of teaching hours relative to other countries, absence of relevant career structure and inadequate initial teacher preparation and this led to the Chilean government enacting the Teacher Professional Development Law in early 2016 (Avalos-Bevan, 2018: 308).

3.4.2.10 Lessons that can be drawn from Chile

Lessons that can be drawn from the case of Chile include:

- The design process of the performance evaluation system was participatory which facilitated buy-in and ownership by the stakeholders although it took a long time to reach consensus.
- The evaluation process is premised on shared understanding of 'effective teaching'.
- There are multiple evaluators and ultimately multiple evaluations which help to mitigate against bias of one evaluator.
- There is differentiated targeted professional development.
- The Chilean experience proves that combining the formative and summative functions is feasible and the two are mutually reinforcing.
- The evaluation process should not be too prescriptive to promote growth.
- The Ministry of Education contracted a University Measurement Centre mandated to maintain a permanent research agenda of assessing the validity

of the teacher performance evaluation system meaning that there is continuous review of the system.

3.4.3 Teacher Evaluation in China

The education system of China is centralised although the school management has some degree of autonomy. In the school organogram, there is the principal and vice principal followed by the directors who preside over the Office of the Teaching Affairs (OTA) and the Office of Moral Education (OME), then there is the Subject and Year department heads before the teacher (Zhang & Ng, 2017:201). The OTA is responsible for promoting research and provides a platform for the teachers to share and discuss their teaching practices through collaborative lesson observations while the OME is responsible for the supervision of the teaching of the social, political, ideological and moral values. The Subject department head oversees the teaching issues of a specific subject while the Year department head oversees teachers who teach at a given grade level irrespective of subject area and focuses on the moral education of the students. Within the subject departments, there exist lesson preparation groups where teachers who teach the same subject and grade level do collaborative lesson preparation.

3.4.3.1 Brief history of teacher evaluation system development in China

Liu and Zhao (2013:232) posit that any education system needs effective teachers and that teacher effectiveness can only be assured and ensured by having an effective evaluation system which is fair and reliable. The process of establishing a fair and reliable system is a highly debatable topic and hence the use of evaluation data for high stakes decisions is controversial. China has recently reformed its teacher evaluation system in line with the global trends where countries are trying to address the quality of education by addressing teaching quality.

The year 2009 is significant in the calendar of the education system of China because it is in that year that teacher performance pay was implemented. The implementation of the teacher performance pay evaluation in China resulted from the enactment of the 'Guidelines to teacher performance evaluation in compulsory schools' by the State Council (Liu & Zhao, 2013:232; Liu et al., 2016:793).

3.4.3.2 The introduction of the Teacher performance pay

China has had a teacher appraisal system primarily focused on the accountability function with little attention being paid to the professional development aspect (Zhang & Ng, 2017:197). There was however a focus shift in 2001 when the country embarked on the large scale New Curriculum Reform nationwide. The reform had an aim to transform the Chinese education system from focusing on knowledge transfer approach where the students were mere passive recipients of information. The Ministry of Education in turn called for a replacement of the then teacher evaluation system with one that promotes professional development. The call for a professional development oriented evaluation system gained momentum from both practitioners and researchers alike. The call however died a premature death as 2009 saw the introduction of the teacher performance pay system which meant that teachers would now be paid in relation to their performance. The salary of teachers would now be divided into basic pay that constituted 70% of the total salary and merit pay that constituted 30% of the total salary (Liu & Zhao, 2013:232; Wang, Lai & Lo, 2014:433; Zhang & Ng, 2017:197). The basic pay was similar for teachers in the same level of appointment who had similar experience and satisfied the workload in the contract while the incentive pay catered for the teacher's contribution to the school in terms of taking on extra workload, students' results and research output (Wang et al., 2014:433). The introduction of Teacher performance pay evaluation system resulted in the focus of the appraisal being reverted back to the accountability function at the expense of the development function (Zhang & Ng, 2017:197).

3.4.3.3 The characteristic features of the teacher performance pay evaluation system of China

The exploration of the Chinese teacher performance pay system in this study is mainly based on the work of Zhang and Ng (2017:203-209) because they have given a detailed and in-depth description of the system based on the study they conducted. The description includes the purpose of the appraisal process, appraisal criteria that is used, the appraisal methods, the appraisal procedure and the appraisal outcome.

The purpose of appraisal

Although the appraisal system is assumed to have solely an administrative or accountability function by virtue of its being a teacher performance pay system based

on the reward/ punishment principle, research has established that the system actually has a dual function of both accountability and professional development (Liu & Zhao, 2013:246; Zhang & Ng, 2017:202). It is proffered that Chinese researchers have undergone a transformation in the way they perceive reward and punishment evaluation which they initially regarded as purely summative without serving any professional development function. They have since acknowledged that such a system can serve both accountability and professional development purpose (Liu & Zhao, 2013:246). Wang et al. (2014:430) on the other hand assert that the performance related pay was introduced with two main aims, one of motivating teachers to partake in professional development activities that would allow them to gain skills in areas like inquiry based and cooperative learning. The second aim was to encourage teachers to work harder and inspire them to be more responsible and answerable for their individual performance. The skills in inquiry-based and cooperative learning were necessary after the New Curriculum Reform of 2001 which called for a shift from the knowledge transfer approach where the students were simple passive recipients of information who thrived on rote learning (Zhang & Ng, 2017:197).

Wang et al. (2014:433) posit that the performance related pay was introduced for three purposes. The first was to maintain and improve levels of teachers' pay so as to improve their social status. The second reason was to acknowledge personal effort by introducing the incentive scheme and removing a remuneration structure that was based on rank or appointment and experience without rewarding the hard working teacher. The third reason was to establish a centralised payment system and remove the system where schools had to pay teachers which in turn removed the problem of the best teacher migrating to the best paying schools thereby creating an imbalance in the quality of education offered by schools with differing paying capacities. The system of schools paying the teachers had created an inequality among the teachers across the different schools, but on the other hand, the introduction of the performance pay also introduced another inequality, this time within the school where teachers got different salaries based on their performance (Wang et al., 2014:436). Wang and colleagues further argue that although performance-related pay was introduced with the aim of benefitting the ordinary teacher, the opposite however happened thereby failing to encourage excellent work as had been expected.

The appraisal criteria

The evaluation criteria constitute the standards used as the basis of the evaluation of teacher performance. The standards fall into four main categories namely morality, competence, achievements and diligence (Zhang & Ng, 2017:204). The aspect of teacher morality covers issues of professional ethics and conduct while teacher competence considers the teacher's expertise in moral education, teaching practice and research. The aspect on teacher's achievement covers issues such as the students' examination results, teacher's research results, awards given to the teacher and those given to his/ her students. The teacher achievement element has the highest weighting in the evaluation criteria. Although the policy documents on the pay reform clearly stipulates that student test scores should not be used as a basis of teacher evaluation, the Bureaus of Education at district level however set an expected percentage pass for public examinations and the schools proceed to use these as the basis of evaluation (Wang et al., 2014:437).

The last aspect of the evaluation criteria which is teacher diligence focuses on the extent to which the teacher is committed to their work. It should however be noted that although teacher morality accounts for just 10% of the total score of the appraisal process, it can however still lead to the disqualification of a teacher from the appraisal process if he/ she fails to satisfy this criterion. Wang et al. (2014:433) give three main criteria namely professional ethics, teaching in the moral and subject domains and fulfilling the duties of a class teacher against the four criteria given by Zhang and Ng (2017:204). On closer examination, the criteria in the two publications basically cover similar standards although some of the standards were put in different categories.

Appraisal methods

The evaluation sources of information are multiple and include lesson observation, student evaluation, students' examination results and the teacher portfolio. Zhang and Ng (2017:205-208) give a detailed account of the different appraisal methods and appraisal procedure used in China.

Lesson observations

Lesson observations are widely used as the evaluation source of information. In China, the lesson observations are divided into public and routine lesson observations. The public lessons are designed carefully and thoroughly prepared since they are observed by a group of people that include the school administrators, the directors, peer teachers and even teachers from other schools. Public lessons are meant to facilitate professional development for both the observers and the observed through sharing the observations made which assist to enhance pedagogical competence. A teacher is expected to give at least one public lesson per year. The routine lesson observations are often done without prior notification and are the preferred means of evaluation since they tend to reflect a teacher's usual practice as there is no special preparation done for the observation.

Student evaluations of teachers

The perceptions of students on the performance of their teachers are normally gathered through the questionnaire surveys and interviews. The student survey seeks the students' perceptions on teacher effectiveness, attitude, assignment marking, classroom discipline, moral education and after class tutorials. The student evaluation extends beyond the classroom thereby making it more comprehensive relative to the lesson observation. Student evaluations also give the teacher's real and normal work tradition. Student evaluation of teachers is practised earnestly and ratings such as excellent, average and poor for each of the different evaluation items. Teachers' attitudes towards student evaluations have been evolutionary from being negative to gradual acceptance.

Analysis of students' results

Both the teachers and administrators regarded the students' results as being factual and meaningful relative to the other measures of the teacher's performance. Students' examination results are also used to identify the strengths and weaknesses of instructional practice so that the relevant remedial measures can be instituted.

Teacher portfolio

The teacher's work is inspected by looking at the teacher's lesson plans, assignment marking and notes made on the lessons that they have observed.

Appraisal procedure

Appraisal results are formulated in three sequential steps, teacher self-appraisal, then departmental level appraisal and lastly, school-level appraisal.

Self-appraisal stage

The teacher self-appraises against a criteria listed on the appraisal form. This allows the teacher to bring out his/ her view point on his/her teaching practice. Zhang and Ng (2017:207) however assert that self-evaluation does not give a genuine reflection of the teacher's performance since teachers have a tendency to inflate their scores to which the teachers admitted in the study. Self-appraisal thus does not carry significant weight in the final appraisal. It is however acknowledged that self-appraisal assists to enlighten the teachers on the school expectations and provides guidance for their future development.

The departmental level appraisal

The subject department head is responsible for evaluation at this level. The teachers are appraised on the basis of criteria similar to that used for teacher self-appraisal. The appraisal can be done by the department head alone or by a 3 or 4 member panel which includes the departmental head, lesson preparation group heads and at times ordinary teachers. The year department head also appraises the teacher in moral education and this is combined with the score awarded by the subject head.

School level appraisal

Appraisal at this stage is conducted by the school level appraisal committee composed of the principal, the vice principal and the two directors. The principal oversees the process at this stage. The appraisal is based on:

- a. Teacher's achievement (including examination scores for students)
- b. Students evaluation results.

- c. Information collected through lesson observation, inspection of teacher's daily work.
- d. Committee member's perceptions and knowledge of teachers.
- e. Departmental appraisal results.

The school level appraisal rankings are the final and official results. It should however be noted that in China, evaluation is a process and not an event, evaluation activities are conducted continually in order to fulfil the formative function of the evaluation process by facilitating the identification of strengths and weaknesses to ensure timely feedback.

Appraisal outcomes

Teachers are ranked and honoured based on the appraisal results. There are four appraisal grades: *excellent*, *qualified*, *barely qualified* and *unqualified*. Zhang and Ng (2017:208-209) give a description of the appraisal outcomes of the teacher performance pay system. Teachers are concerned about their appraisal results because they inform personnel decisions like promotion, job assignment, honour and award granting, salary and bonus payment. Teachers that are ranked as *barely qualified* or *unqualified* may be given fewer lessons, removed from teaching the graduating classes or given a non-teaching position like a librarian or even to be dismissed. A teacher who is ranked as *barely qualified* twice is taken as *unqualified*. A teacher is awarded a monthly salary if he/ she is ranked *excellent* on 3 consecutive occasions. Individual schools also have a myriad of honours and awards to acknowledge teachers' outstanding performance and achievements. Team awards like outstanding department are awarded to subject departments. Year departments that have made significant progress are also given awards.

Factors attributed to the success of teacher performance pay system

A number of factors have been listed for contributing to the success of this system. Zhang and Ng (2017:211-213) have given the following factors as the reason for the success of the system:

Complementariness of administrative and development purposes

Zhang & Ng (2017:211) assert that the formative and summative (developmental and accountability) functions can co-exist and have been found to be mutually supportive and complementary like in the case of China. Evaluation data is used for both summative and formative evaluation purposes proving that although combining formative and summative evaluation successfully is challenging, it is however possible (Danielson & McGreal, 2000:10; Hinchey, 2010:6). Zhang and Ng (2017) argue that accountability and developmental functions are mutually reinforcing rather than conflicting. Another school of thought however argues that merging the two evaluative functions makes them counter-productive while they remain splendid if kept separately (Carroll, 1997:10-11; Popham, 1988:269;). Successful combination of the two functions has however proven that they can mutually reinforce and complement each other as proven in the case of China and Chile to name a few.

Integrating teacher appraisal with other school tasks

Zhang and Ng (2017:211) established that teacher appraisal process was integrated with other school activities in the schools that they studied. Zhang and colleague proffer that an appraisal system should be linked to the school development agenda so that the development efforts are consistent and coordinated. For an appraisal system to be able to facilitate teacher development and promote educational quality, the system has to be ingrained and aligned with the other school tasks and vision (Stronge & Tucker, 2003:5).

Credibility and Fairness of teacher appraisal

Credibility and fairness have been identified as requisite ingredients for a successful appraisal system. Although teachers have acknowledged that the appraisal process is not technically perfect, they have however agreed that it is credible largely due to the existence of clear standards, validated instruments, frequent data collection processes and thorough evaluation procedures (Zhang & Ng, 2017:212). The use of multiple sources of data helps to validate the individual data sources thereby reducing the risk of bias from a single source of data. The appraisal results are evidence-based which made the teachers to generally accept the results as a true reflection of the reality. Effort has to be made to ensure that the appraisal process is perceived as fair

and equitable by all the teachers. Cardy and Dobbins (1994:54) concur that a system that is perceived as unfair and inequitable is likely to fail since teachers believe that unfairness is avoidable by having an objective appraiser. Zhang and Ng (2017:213) list measures that ensure fairness as follows:

- a. Having a panel rather than an individual finalising the appraisal results.
- b. Ensuring transparency and openness by shortlisting and displaying teachers for important awards before confirmation to give room for divergent views.
- c. Making the appraisal criteria applicable to all and at the same time accommodating the differences between teachers, for example, differences in computer proficiency between the young and the older teachers.

Liu et al. (2016:804) have however argued that the exam-driven system of China has brought unfairness by creating a situation where core-subject teachers are given more work and resultantly paid more while the non-core subject teachers are given less work thereby resulting in them getting less money than their counterparts thus creating some form of unfairness. Veteran teachers have also felt unfairly treated when the younger teachers receive more money than them in spite of their experience. Such a feeling occurs because the teachers have disregarded the issues of capability and competence considered in evaluating teacher performance (Liu et al., 2016:805). Liu et al. also submit that research has established that veteran teachers do not support incentive policies. This could be an area that could be explored in further studies in the Zimbabwean situation.

Multi-dimensionality of teacher appraisal

Literature states that a quality appraisal system should satisfy certain technical requirements which are:

- Having a clear and specific appraisal criteria
- Having appraisers that are objective
- Having appraisal methods that can accurately measure teacher performance (Zhang & Ng, 2017:213).

Zhang and colleague claim that the current appraisal system of China satisfies the listed technical requirements. The two also acknowledge that besides satisfying the technical criteria, an appraisal system needs to extend to being psychologically and

socially sensitive by taking into account the teachers' feelings and attitudes if it is to be successful. Disregarding the psychological and social impact of the appraisal process is tantamount to disaster.

Criticisms against the teacher performance pay evaluation system

Some criticisms have been levelled against the performance pay evaluation system. Wang et al. (2014:440) proffer that a closer examination of the system has established the following: compliant professionalism, a failure to improve teaching standards, and a failure to improve the quality of education. In this scenario teachers' professionalism was interpreted as 'compliant professionalism' because the teachers had to set aside their professional judgement and follow external regulations in doing their work thereby compromising their professional autonomy. They simply taught for the test so as to satisfy the requirements for the evaluation. The system is also accused of failing to improve the teaching standards and this has led to teachers being labelled as 'performative workers' (Wang et al., 2014:441). Teachers simply worked hard to satisfy the evaluation standards and shelved their professional authority. Teachers argued that instead of the school-based development effort giving room to the teachers to be innovative and creative to meet their students' needs, the development programme merely aimed at knowledge and skills development. This therefore led to the conclusion that the reform failed to motivate the teachers to work hard to enhance their teaching and professionalism but rather added pressure and stress of having to satisfy the numerous quantitative indicators and having to undergo too many evaluations.

Lastly, the reform is accused of failing to improve the quality of education due to the restrictive and regulatory approach that created tension between the demands of the school, what is desirable for the teacher and what is desirable for the student. The levelled criticism all point to a faulty implementation process which was more limiting and restrictive rather than promotive of professional development. The reform has also been criticised for the top-down approach due to the limited teacher participation and involvement especially on the design aspect of the system (Wang et al., 2014:441). Liu et al. (2016:793) on the contrary argue that although the Chinese government has always had a top-down approach, the performance pay reform to some extent reflects a bottom-up approach in the sense that it is the schools rather

than the department of education that have more power in deciding the amount of the teachers' salaries.

Lessons drawn from the case of China

- Students evaluate teachers.
- Appraisal is at three stages:- self, departmental, and school level. The school level appraisal is conducted by a committee and it is the one that gives the final result.
- Teachers that are ranked as barely qualified can have graduating classes being withdrawn from them or can be removed from teaching to non-teaching positions.
- There is collaborative lesson planning for similar subjects and grades.
- There are departments awards that further encourage team and collegial approach to teaching.
- Appraisal results inform personnel decisions.

3.4.4 Comparative discussion of the three country (South Korea, Chile and China) evaluation systems

The three teacher evaluation systems are discussed to identify their similarities and differences in terms of *readiness, design, process and significance*. In terms of readiness assessment, South Korea pilot tested two versions of the teacher evaluation system between 2006 and 2010 before full scale implementation was finally done. The piloting process helped South Korea to ascertain the teachers' readiness to implement the system and also created a forum to incorporate their input. Chile and China, on the other hand, did not have this process. Chile had however adopted a participatory approach to the design process and so the teachers' input had been incorporated implying their readiness for the process. Although South Korea had attempted a bottom-up approach at design phase by involving the key stakeholders, they had abandoned it when stakeholders failed to reach a consensus on the evaluation issues due to their different positions on teacher evaluation. The Ministry of Education, Science and Technology of South Korea however still managed to get some degree of involvement as input of the stakeholders was sought during the pilot phase. In the Chinese case , the design process was non-participatory. China adopted a top-down

approach and simply imposed the system on the implementers without even piloting it first. Thus the readiness aspect was not ascertained.

Regarding the design of the systems, the three countries differ. For instance, South Korea has three separate evaluation systems running concurrently although serving different functions. This has proved to be a burden to the teachers as they have to satisfy the demands of each of the systems. Having the three systems running at the same time has discredited the evaluation process as in most cases individuals have had conflicting results from the different evaluation systems making the process questionable. China has one teacher evaluation system which is premised on the reward/punishment principle. However, research has established that the system actually has a dual function of accountability and professional development (Liu & Zhao, 2013:246; Zhang & Ng, 2017:202). In Chile, the teacher performance evaluation system also has a dual function although it places more emphasis on the formative or professional development function like in the South Korean case. The Chinese and Chilean cases have shown that it is feasible to have an evaluation system with a dual function and that the functions can be mutually reinforcing rather than antagonising to each other.

In terms of implementation, all the three countries have multiple evaluators, thus, multiple evaluations ultimately. South Korea and China include students as part of the evaluators unlike Chile. In South Korea, even the parents are involved in the evaluation process. The Chilean evaluation process is premised on a shared understanding of 'effective teaching' which was arrived at after wide consultations had been made. This is unlike the situation that exist in South Korea where there is no consensus on the definition of 'effective teaching'. The evaluation criteria and performance standards in South Korea are also not very clear and distinct. This has led teachers not to have confidence in the evaluation process especially that conducted by the students and the parents. In the case of China, although the development process was non-participatory, the teachers regard the evaluation criteria and standards as clear and the process as thorough and therefore credible. The Chilean teachers certified the validity of the evaluation system from the drawing board since they were involved in its design. The teacher evaluation system of China has been credited for being ingrained and aligned with the other school tasks. In South

Korea and Chile, professional development programmes have been criticised for not being integrated within the teacher's daily work plan. In South Korea, professional development programmes are mandatory and teachers regard them as punitive. This has negatively affected the evaluation process as peer evaluators have had to inflate the ratings of their under-performing peers to prevent them from having to be humiliated by having to go for the mandatory programmes. Both the Chilean and Chinese professional development programmes have been criticised for being remedial and correctional rather than promoting growth and creativity. The professional development programmes in both Chile and South Korea are however differentiated and targeted in line with the evaluation process findings.

In terms of significance, the South Korean and Chilean performance teacher evaluation systems have differentiated and targeted professional development programmes. The monetary incentives of Chile have not been significant enough to motivate effort as in the case of China where thirty percent of the teacher's salary is merit award. In Chile, high performers are given sabbatical leave that they can use to conduct research which is aimed at improving their professional growth. Chilean high performing teachers are also given non-monetary incentives and are prioritised when promotion opportunities arise. In China, the performance appraisal results inform decisions on promotion, job assignment, salary and bonus payment. The Chinese, unlike the other two countries' performance systems, has team awards to promote collegiality and collaborative approach to teaching.

The three countries teacher performance evaluation systems are all high-stakes evaluation systems. The examination of the evaluation systems in these three countries has revealed differences and similarities with regard to the four components of the RADPS conceptual framework (*Readiness assessment, Design, Process, and Significance*). One of the important factors highlighted is the link between the evaluations and the teacher professional development. Although there are differences in the ways in which the evaluation results are used, the link contributes significantly towards performance evaluation effectiveness.

3.5 Readiness Assessment

In introducing a teacher evaluation system, it is prudent to assess the appropriateness of the timing of the introduction of a performance evaluation system in what Kusek and Rist (2004:40) coined 'readiness assessment'. Readiness assessment provides an analytic framework to ascertain the capacity of an organisation to monitor and evaluate its performance and determine if the organisation is ready and in a position to build, use, and sustain the system. In ascertaining readiness for a new evaluation program, there are some key issues that have to be considered (Kusek & Rist, 2004:41). These include but are not limited to having a legal framework to regulate the implementation process, political will, resources, buy-in, demand for evaluation, institutional capacity and technical skills.

One of the issues to consider when conducting a readiness assessment is to ascertain if there is a legal framework upon which the evaluation process will be anchored. The legal parameters of the evaluation process can be derived from the constitutional law, statutory law, previous rulings from the courts, and negotiated agreements (Tucker & DeSander, 2011:5). Tucker and colleague (2011:5) add that key aspects of the evaluation system that include the evaluation criteria, procedures and consequences of the evaluation process are premised on legal principles. Literature postulates that for a performance appraisal system to be legally defensible, then, it should have written policies that are communicated to all persons (Carroll & Schnier, 1982 in Haefele, 1993:23). This implies that an evaluation system that is legally defensible has to be anchored in strong legal framework which should however be made known to the relevant parties that are affected by it. Although Mackay (2006:8) acknowledges that a law or statute can be a useful way for legitimising a monitoring and evaluation system, Mackay however cautions that a law or decree alone does not guarantee that maximum effort will be directed towards development of a monitoring and evaluation system.

The second aspect in determining the readiness of the evaluation for a new evaluation system is to discern if there is any impetus and demand to build a monitoring and evaluation system. In ascertaining readiness of an organisation to build, create and sustain an evaluation system, the sources of the impetus for an evaluation system and the reasons for the demand should be established (Kusek and Rist, 2004:43). In

ascertaining the impetus and demand of building a monitoring and evaluation system, five key aspects have to be verified:

- What are the reasons for building the monitoring and evaluation system?
- Who is advocating and championing the development of the monitoring and evaluation system?
- What are the motives of the champions in building the evaluation system?
- Who is set to benefit from the development of the evaluation system?
- Who is opposed to building of the monitoring and evaluation system? (Kusek & Rist 2004:41).

Mackay (2006:5) asserts that meaningful demand is essential to institutionalise monitoring and evaluation successfully. There are some factors that have been found however to hinder demand especially in situations where key stakes need to buy-in and embrace the idea before effort can be put towards the development and funding of the monitoring and evaluation (M & E) function. The factors include lack of government demand for M & E due to inadequate understanding of the functions of M & E, lack of understanding as a result of lack of experience with M & E, and lack of experience because of weak demand (Mackay, 2006:6). Mackay (2006) suggests that demand issue can be addressed by increasing awareness of M & E which will lead to the key stakes appreciating it better by revealing to them highly cost-effective M & E activities, and giving those examples of other governments that have successfully established M & E systems. Mackay also cautions that having M & E information will not guarantee its use but that the presence of powerful incentives are key to increasing demand and in turn leading to successful institutionalisation of M & E.

Readiness assessment should also involve the identification of the champion for the system. A champion for an evaluation system should ideally be someone who is at the centre of policy making and has influence with important decision-makers if the system is to be regarded as relevant, trustworthy, usable and timely otherwise meeting the viability requirements would be difficult (Kusek & Rist, 2004:44-45). It should however be noted that although government champions have played instrumental roles in the development of some of the more successful government M & E systems, for example Chile and Australia, having a powerful champion is however no guarantee for success as in the case of Egypt where doubtful mid-level officials frustrated the efforts of these

seniors (Mackay, 2006:6). In the process of identifying the champion, it is vital to establish the motive of the champion which could include any of the following: a minister might want his/her ministry to be championing reform, doing it simply because one wants to be accountable and is results-oriented, doing it out of a sense of responsibility for the community or to find favour with donor community, parliament, or public (Kusek & Rist, 2004:45). The motive of the champion may influence the level of commitment that one holds towards the cause.

Readiness assessment should also include institutional capacity assessment in terms of technical skills, available technology, available fiscal resources and institutional experience. It is critical to identify existing gaps that have to be built to sustain the evaluation system and this is done by identifying both the current and possible hurdles (Kusek & Rist, 2004: 43). Possible impediments can include financial limitations, political will, political champion, expertise or prior experience. Key factors to consider in determining institutional capacity include:

- Ascertain available skills and expertise.
- Ascertain the availability of technical assistance, capacity building or training in M & E
- Verify the availability of any institutes, universities or research centres that can avail technical assistance and trainings for all civil servants (Kusek & Rist, 2004:43).

Determination of institutional capacity during the readiness assessment phase allows the designers to assess the feasibility of the intended programme which thus increases the chances of programme success.

3. 6 The design process of a performance evaluation system

The design process is a procedure that involves the conversion of a vision or abstract ideology into physical reality. In this scenario, the design process involved the translation of the idea of performance evaluation into reality by developing the regulatory framework, the job descriptions, standards, evaluation criteria, evaluation guidelines, evaluation tools, reporting instruments and any other relevant materials for the evaluation process.

3.6.1 Conditions that have to be observed in the design process

Danielson and McGreal (2000:64) outline four conditions that must be observed in the process of designing a teacher performance evaluation system that is comprehensive. The first condition is that the system must be framed in line with the general objectives of the schooling process. It must be aligned to the mandate of the national education system although the approach to the development of the evaluation framework can differ depending on how centralised or decentralised the education system is. Other researchers concur to this when they proffer that it is critical for the performance evaluation system goals to become an integral part and embedded in the other system initiatives for coherence rather than being disconnected and isolated activity (Goe, Holdheide & Miller, 2014:9; Stronge & Tucker, 1999:347).

The second condition is that the goal or intentions of the evaluation process should be articulated from the outset (Danielson & McGreal, 2000:64; Marzano, 2012:19). The third observation to be made is to ensure that the roles of the educators at the different hierarchical levels are clarified within the evaluation framework and lastly, the teacher evaluation process needs to be in sync with the student assessment, school and system level evaluations. This implies that the teacher evaluation process must complement and be consistent in objectives with the (student) assessment and (school and system level) evaluation frameworks.

3.6.2 Stakeholder involvement in the design process

The design process must be as inclusive as possible so as to enhance its acceptance, facilitate successful implementation and sustainable use. A system that is borne out of the contributions of all key stakeholders is better respected and embraced compared to one that is imposed on the users and is likely to be perceived as responsive, useful and fair (Goe et al., 2014:14; Stronge and Tucker, 1999:353). An evaluation system 's validity and relevance can only be certified by the direct users so their involvement from the drawing table helps to validate the system before it has even gotten off the ground (Danielson & McGreal, 2000:64).

Isoré (2009:8-9) advances that different stakeholders which include the Government, local authorities, school administrators, educational researchers, experienced teachers and the teachers' unions should be involved in the design process of a

teacher performance evaluation system. The Government has the legal mandate to set the national learning outcome objectives, of which if an evaluation system is to be judged as efficient, then, it must be directed towards the attainment of those very goals. It is rational that the Government is involved from the onset since it often plays a direct role in the implementation and monitoring of the teacher evaluation processes particularly in countries where the teachers are civil servants. In some countries however, local authorities are responsible for ensuring the educational quality of the schools under their responsibility and accountable for the achievement of the national objectives which therefore justifies the involvement of their representatives in the design process of the performance evaluation system. Educational researchers and teachers are key in the design process because of the expertise that they hold in terms of 'good teaching' practices, as well as the instruments and criteria that can be used in the process. Besides their expertise and being better placed to tell what can work in the teaching process, teacher involvement in both the design and implementation processes is purported to assist in garnering teachers' collaboration, support and promotion of the state and districts effort in provision of quality education (Goe et al., 2014:14). Research assert that teacher involvement has to be overemphasized because their exclusion from the design process promotes division and discord between them and the administrators which is fatal and a recipe for disaster to the teacher evaluation process (McLaughlin, 1990:406) while on the other hand, their participation is critical for them to develop a sense of ownership(Stronge & Tucker, 1999:341). The unions for the teachers may also be consulted for their contribution in the design and implementation procedures because they represent the interests of the teachers and what they approve will be acceptable to their constituents.

Johnson (1997:85) however cautions that the different stakeholders may have differing conceptions of 'effective teaching' which have to be addressed early in the process of policy development so that the stakeholders progress with a common perspective. Strong and Tucker (1999:339) also weigh in submitting that expectations of different stakeholders differ and are often conflicting and therefore political skill is necessary to gain their input and political support for the new evaluation system. It is important to ensure that the perspectives and interests of the different stakeholders are addressed in the evaluation system that is eventually created. Johnson (1997:85) proceeds to emphasize that for any evaluation system to be viable, it has to address the concerns

of all constituents which is possible only with a reasonable amount of negotiation and compromise by all stakeholders.

Stakeholder participation allows political processes of expression, discussion of conflicting ideas and negotiation before policy implementation which enhances the final product (Stronge & Tucker, 1999:341). This is reiterated by Heneman et al. (2006:6) when they proffer that for an evaluation system to survive and contribute to performance improvement, it must be accepted by those who use it.

3.6.3 Steps in designing an evaluation system

The discussion of the design process of an evaluation system is mainly informed by Danielson and McGreal's (2000:64-75) step by step procedure. A few other researchers were also used to enrich the discussion in the likes of Choi and Park (2016), Goe et al. (2014), Isoré (2009), Heneman et al. (2006), Marzano (2012), Peterson (2000), Santiago and Benavides (2009), Zhang and Ng (2017) and others.

3.6.3.1 Step 1: Determining the process and evaluation goals

The initial step to be adopted involves deciding on the general process to be followed in the design procedure and this normally includes the formation of the evaluation committee, coming up with a schedule of committee meetings and deciding on how decisions will be made. It is also critical that the evaluation goals are decided at the out-set of the design process.

Determining the evaluation committee

As indicated earlier on, it is important that the committee is as representative as possible by including members from different constituencies, for example, teachers, school administrators, central office administrators and even union representatives. This is important in that it will help to ensure that the perspectives and interests of the different groups can be accommodated during the process.

The different constituencies serve different purposes which help to enhance the validity of the evaluation system. Teachers are experts in their practice and therefore can offer valuable contributions on the more technical aspects like the evaluative criteria while at the same time giving a more accurate perspective on time demands of the system. School administrators on the other hand will assist to ensure that the

evaluation system is workable given the available resources since they are more concerned with the practical aspects and issues and will strive to ensure that the process results in professional learning as instructional leaders.

Members of the central office administrators are cognisant with the state laws and have extensive and valuable content expertise and therefore will assist to make sure that the system is in line with the national regulatory framework governing performance evaluations. Teachers' unions will ensure that the process promotes the professional learning and development and that the teachers are active participants in the process. Representation of the different constituencies of the stakeholders will thus make the evaluation system more valid, relevant and acceptable (Heneman et al, 2006:6; Isoré, 2009:8). It is also recommended that the committee should try to have the different level schools (primary, secondary, rural and urban) being represented in the committee so that their interests are also catered for in the evaluation system.

Desirable attributes for evaluation committee members

The task of designing a new teacher evaluation system is challenging and taxing therefore the members should have interest, commitment, willingness to participate, open-mindedness, leadership and communication skills (Danielson & McGreal, 2000:66). Committee members need to have enthusiasm for the task and be committed to its success. Members also need to be active participants so that the perspectives of the constituencies that they represent are also regarded and inculcated while at the same time being receptive to new ideas and perspectives. The members need to possess good communication skills so that they can manage to serve as liaison between the committee and the constituent they represent and be able to justify the positions taken by the committee in their reporting.

Determining the evaluation goals

A fundamental aspect of the design process is to decide on the goals or purpose of the evaluation process. Goe et al. (2014:9) assert that it is imperative that the first step in designing a teacher evaluation system should be to establish well defined goals and purposes of the system. They add that there is a general tendency to simplify this crucial step which is however the basis of all other decisions throughout the design process namely, the methods of evaluation that are selected, the weighting for the

different components of the criteria and the consequences of the evaluation process. It is important therefore that teacher evaluation goals are selected with caution and their selection should be guided by the general objective of teacher evaluation which is to improve teaching and learning.

Two major purposes of performance evaluation have been identified namely the formative/improvement/developmental and summative/accountability/administrative functions (CDE, 2015:15; Danielson & McGreal, 2000:8; Marzano, 2012:19; Santiago & Benavides, 2009:7).

Formative / Improvement function

Formative evaluation is targeted at acquiring information to help teachers to improve and expand their teaching abilities. The formative function of a performance evaluation system can be achieved by a number of activities which include the identification of training needs of a teacher and determination of the teacher's strengths and weaknesses for further professional development (Danielson & McGreal, 2000:8; Haefele, 1993:22; Hinchey, 2010:6; Santiago & Benavides, 2009:8). Formative evaluation aims to provide useful feedback for the improvement of teaching practices through professional development. This type of evaluation aims to shape a teacher's performance and makes it more effective by encouraging the use of different approaches in a bid to identify the ones that work better (Popham, 1988:270). Formative evaluation enhances development through counselling in order to motivate the teacher and also through training to develop the professional skills.

For formative evaluation to have an effect on instructional practice, an environment that is non-threatening is required in order to encourage teachers to open up and readily admit their weaknesses (Hinchey, 2010:6; Popham, 1988:269). Formative evaluation requires that the evaluator relates to the evaluatee as an ally by providing help and assistance to improve performance. It requires that an evaluator be able to diagnose the needs of the teacher in terms of what he/she should learn, suggest or recommend ways in which the learning should occur, monitor and direct the learning process, and take responsibility for the teacher's professional growth (Hunter, 1988:278). Formative evaluation should facilitate professional conversation in a collaborative setting which promotes reflection on own practice. The discussions

should be genuine without any undercurrents of point scoring if they are to succeed in promoting self-assessment and reflection. It focuses on helping teachers improve their instructional performance rather than penalise them for their shortcomings.

Summative/ Accountability function

Summative evaluation unlike formative evaluation refers to the process of making judgement aimed at making high stake decisions such as awarding merit pay, promotion or tenure (CDE, 2015:15; Danielson & McGreal, 2000:7; Hinchey, 2010:6). This type of evaluation focuses on holding teachers to account for their performances and associates their performance to a range of consequences for their career. It is aimed at making sure that teachers perform at their best to enhance student learning and also works as a means to provide recognition to those teachers that excel in their performance (Santiago & Benavides, 2009:8).

3.6.3.2 Step 2: Deciding the procedure (the 'How')

The committee must decide on issues such as whether there will be differentiated procedures for the different groups of people, the timelines for the evaluation activities, roles of the different stakeholders in the evaluation process, and measures to ensure fairness and impartiality in the process (Danielson & McGreal, 2000:68). If the process is to be differentiated for the different teacher groups, for example novice and veteran teachers, then the committee must decide the activities to be used by the different groups and the stages at which they will be used. Process could be differentiated for the novice and veteran teachers in terms of the criteria (Danielson & McGreal, 2000:71; TNTP, 2010:4) used and/or the frequency of evaluation (Goe et al., 2014:34).

Santiago and Benavides (2009:12) also add that for the teachers that are to be assessed, it must be decided, whether it is all teachers or a subset, for example permanent versus contract teachers, novice versus experienced teachers, or primary versus secondary teachers. The nature and character of the evaluation system which is the 'how' aspect also must articulate whether evaluation will be mandatory or voluntary, the frequency of the evaluations (yearly or every two or more years), and whether it will be school-based or externally-driven teacher evaluation (Santiago & Benavides, 2009:12). The externally-driven evaluation system has the aspects that are assessed, the evaluation instruments and criteria being common across the

schools whereas in the school-based system, the school has the responsibility to design the instruments and criteria used in the process and is responsible for all other associated evaluation procedures. The advantage of the school-based system is there is school ownership of the evaluation processes and these are customised to the school context although the quality might be compromised due to limited expertise in both the design and implementation aspects.

In terms of frequency of evaluations, it is suggested that every teacher should be evaluated at least every year so that feedback on performance is availed to every teacher annually irrespective of their performance and experience (Santiago & Benavides, 2009:12).

3.6.3.3 Deciding the evaluative standards and criteria (the what)

For a teacher evaluation system to be regarded as fair and reliable, it needs to assess the teachers based on a clear criterion and reference standards of what is considered as 'good' teaching (Isoré, 2009:11; Santiago & Benavides, 2009:12). An evaluation criterion refers to the explicit guideline of what constitutes the expected code of professional responsibility (Isoré, 2009:14). The standards that are used to judge and evaluate the teacher are the evaluation criteria (Zhang & Ng, 2017:198). Standards assist in defining the teacher's work and articulating what is to be assessed in terms of actions and performances. (Celik, 2011:74; Smith, 2005:180). Smith (2005:190) however cautions that standards should be treated with care lest they inhibit professional autonomy and foil creativity and development and should thus not be too detailed but rather give room to an explorative approach in achieving professional competence. The use of performance standards has been criticised in cases where the standards are originated by people from without the teaching profession which tends to violate and demean the teachers as professional (Celik, 2011:74) and for the fact that they are blamed for disregarding the 'complexity and unpredictability of teaching and learning (Korthagen, 2004:79).

In establishing evaluation standards, care should be taken to ensure that these are aligned to the evaluation system goals and purpose (Goe et al., 2014:11). Marzano (2012:19) reasons out that, the difference in purpose will lead to a difference in design in the sense that, in cases where accountability is the main purpose, a small number of elements in the evaluation criteria is adequate enough to determine the level of the

teacher's skills in the classroom whereas if the evaluation model is developmental oriented, then the evaluation criteria has to be more comprehensive, specific and should focus on the growth of the teacher in the teaching or instructional practices.

The importance of evaluation criteria cannot be overemphasized. Some researchers have described it as the very 'heart' of the evaluation system (Conley, 1987:62). Quality evaluation criteria are thus key to the establishment of an effective teacher appraisal system (Danielson & McGreal, 2000:70; Shinkfield & Stufflebeam, 1995:43-62). It is imperative that as part of the design process, clear evaluation criteria to be used in measuring the level of performance of the teachers are developed.

TNTP (2010:4) proffers that evaluation criteria should be clear, rigorous and based mainly on evidence of student learning rather than teacher behaviours while at the same time being precise to minimise inference so that both the teachers and the evaluators interpret them in the same way. Goe et al. (2014:11) reiterate that standards of performance should go beyond merely stating what is valued in a teacher to defining the practices that can be measured reliably and correlated to student growth. The criteria should be a live document that is refined regularly especially during the first few years that they are in use.

Evaluation criteria can be expressed either in input or output form (Danielson & McGreal, 2000:71). The input format describes the expected tasks of teaching whereas the output measures the level of student learning or a combination of both the teaching tasks and student learning. In my opinion, a measure that encompasses the two dimensions would be more valid in giving a correct reflection of the performance of the teacher.

When defining the evaluation criteria, the relative weight of each criterion should be stated, the description levels of performance on each criterion and the expected performance levels for the teachers at the different stages of their career (Danielson & McGreal, 2000:71). A criterion that is graduated in its approach acknowledges the variations in competency level due to experience and is more ideal because it may not be very real to expect the same level of competency from a newly graduated teacher and an experienced teacher which therefore justifies a graduated criterion for the different stages of careers. The TNTP (2010:4) concur with this in their assertion that

in spite of it being noble to hold high expectations of teachers from their early days in the profession, it is however unrealistic to expect even the talented ones to be at the same level of competency with the experienced educators. An example of such an approach is in Coventry, Rhode Island in America, where the first-year teachers are evaluated on only 7 of the 22 components of the criteria and another 5 are added in the second year of teaching and the teacher is expected to exhibit all the 22 components of the framework from the third year of teaching.

A common evaluation framework that has influenced a number of appraisal systems around the world is Danielson's (2013) Framework for teaching. The framework consists of four domains composed of 22 components. The four major areas covered by the framework are *planning and preparation*, the *classroom environment*, *instruction* and *professional responsibilities* (Danielson, 2013:4). The planning and preparation component constitutes a demonstration of knowledge of subject content and pedagogy, knowledge of the students, setting of instructional goals, developing instruction that is logical and consistent and the capacity to assess student learning.

Under the classroom environment component, the teacher is expected to create an environment of respect and rapport, institute a culture of learning, control classroom procedures and student behaviour and maintain physical order. The instruction component considers the ability of a teacher to communicate clearly and accurately, the ability of the teacher to use a wide range of instructional strategies, the ability to engage students in the learning process, ability to give fruitful feedback and a flexible and responsive approach (Isoré, 2009:11; Zhang & Ng, 2017:198).

As a show of professional responsibilities, the teacher is expected to have a reflective approach on his or her practice, ability to communicate with the students' families, contribute to the school and district development and exhibit professional growth and development (Isoré, 2009:11; Zhang & Ng, 2017:198). In Danielson's Framework for Teaching, each component is measured against four levels of performance: 'unsatisfactory', 'basic', 'proficient' and 'distinguished' (Isoré, 2009:11).

Score combining

As part of the evaluation criteria, it has to be decided if the scores from the different aspects of teaching should be combined. Combining the scores implies that a

deficiency in one aspect of teaching can be compensated for by good performance in another area (Danielson & McGreal, 2000:37).

Rating Scales

It is important that criteria include a rating scale with descriptions of different levels of performance so that both evaluator and evaluatee have a common understanding on a level of performance (Danielson & McGreal, 2000:35). Literature indicates that ratings should be multiple and should include at least two levels above and below expectations with the number of categories being large enough to give teachers a clear picture of their current level of performance yet on the other hand small enough to allow for clear distinction between the levels of performance (TNTP, 2010:7). The levels that meet expectations should be clearly stated and there should be no ambiguity on that.

Selecting points for action

Decisions on when any action that will be taken as result of the evaluation process will be actuated should be determined and specified clearly from the outset (Goe et al., 2014:43). For example, if evaluation results are linked to teacher advancement, how many outstanding ratings will a teacher need before the promotion; if there is under-performance; in how many evaluation cycles will improvement be expected?

3.6.3.4 Evaluation information sources

The selection of evaluation measures is informed by the evaluation purpose and teaching standards, for example, if the evaluation purpose is to improve teacher competency, then a number of the measures should provide evidence of instructional competence (Goe et al., 2014:19). Teaching is a complex exercise and no single data source can present a complete and accurate picture of a teacher's performance. It is therefore important that evaluation systems make use of multiple evaluation measures to determine the level of performance of teachers in an accurate and objective way (Peterson, 2000:91-102; Santiago & Benavides, 2009:14; TNTP, 2010:6). Multiple measures of teacher performance should be used to facilitate a broad view of the teacher's effectiveness on the basis of varied evidence which strengthens the process by making the results to be accepted with greater confidence (Goe et al., 2014:19)

and allows for compensation of weakness of the information sources (Hinchey, 2010:7).

Evaluation measures should yield reliable information so selection of relevant measures is a critical aspect of the design process. Goe et al. (2014:19) suggest that in selecting evaluation measures, certain factors should be considered including:

- The purpose of the evaluation.
- The strength of the measure.
- The relevance of the measure to all student populations and teaching contexts.
- The human and resources capacity strengths and limitations of the evaluation measure. This is done by considering the demands that the measure places on the implementers and the expenses associated with the measure.

Possible Evaluation measures

There are many possible sources of information which include classroom observations, teacher self-evaluation, teacher portfolio, teacher testing, students' results, and surveys of students and parents.

Classroom observations

Coe et al. (2014:25) posit that classroom observations are a common source of teacher performance evaluation information in many countries. Literature indicates that teacher interaction with the students in the classroom is likely to show accurately the level of professional performance and distinguish effective teaching from ineffective teaching (Santiago & Benavides, 2009:14; TNTP, 2010:2). Classroom observations have thus been merited for facilitating the collection of first-hand information by allowing the appraiser to observe teaching in action (Zhang & Ng, 2017:205). Hinchey (2010:8) indicates that the effectiveness of classroom observation is dependent on a number of factors which include availability of adequate time to do a thorough classroom assessment, familiarity of observer with subject content and observer training on the use of the instrument which should be on going. Hinchey however points out that the use of protocols have been criticised for being prescriptive and limiting. There is however an argument that classroom observation is more

effective when it is used mainly for formative purposes rather than for high stakes summative purposes and when it is collegial in nature (Coe et al., 2014:25).

Teacher self-evaluation

This is an instrument which allows the teacher to express their own opinion on their performance. It allows the teacher to reflect on his or her performance and to consider factors that could have impacted on their performance. Self-evaluation promotes stakeholder buy-in by showing respect for teacher's contextual knowledge through the incorporation of their input (Hinchey, 2010:9).

Teacher portfolio

A teacher portfolio constitutes different teaching materials ranging from teaching materials, lesson plans, samples of students work and commentaries on student assessment, reflection notes and self-reported questionnaires. Teacher portfolios can be used to complement teacher self-evaluation (Santiago & Benavides, 2009:15). Portfolios are built over time and their comprehensiveness and ability to provide insight into multiple facets of teachers' performance have increased their popularity although their assessment is challenging due to their complexity (Hinchey, 2010:9).

Teacher testing

Teacher testing can be used to establish a teacher's professional and subject content knowledge and it is often used to access entry into the profession or a permanent position although the results of such tests can also be used for evaluation purposes.

Student results

The use of students results for evaluative purposes is controversial because of the argument that there are many factors that impact on students results besides teacher input (Choi & Park, 2016:168). There is also another argument that since the primary responsibility of a teacher is to assist the student to learn, then measures of student learning have to be considered during the evaluation of a teacher's performance (Santiago & Benavides, 2009:16; TNTP, 2010:2). Santiago and Benavides (2009:16) point out that although the evidence of student learning is basic, it is however more relevant for the evaluation at school level rather than at individual teacher level. It

however appears that there is growing consensus that evaluations should consider a teacher's contribution to student learning in determining teacher performance (Darling-Hammond et al., 2012:8). This has led to the use of value-added models gaining ground as an acceptable measure of teacher performance. Value added models (VAMs) are developed to allow the use of statistical methods in measuring teacher input in student test score gains by isolating student characteristics and other factors that influence achievement like home, school, and community factors. VAMs have however been criticised for lack of consistence and the fact that they are affected by the type of students that a teacher has (Darling-Hammond et al., 2012:9). Their lack of consistence occurs because a teacher's value-added score differs based on the test, class, and even statistical model that is used. VAMs are designed to be used to determine the teacher's effect on students that have been randomly assigned but then they are limited in that if students are selectively assigned they cannot be adjusted to pick out the accurate input of the teacher.

Students and parents surveys

Student evaluation of teachers involves the collection of students' perceptions of their teachers' performance often through questionnaire surveys (Zhang & Ng, 2017:205). Zhang and Ng (2017: 205) assert that comparative to the classroom observations, student evaluations of teachers are more comprehensive in the scope of their coverage since they extend beyond the classroom thereby providing information that is otherwise inaccessible about a teacher's genuine daily work practice. Santiago and Benavides (2009:16) however indicate that students and parents surveys are not usually used for individual teacher evaluations but more for the whole school evaluations. The quality of survey instrument can determine the reliability and validity of results of student surveys so their use is recommended as part of a comprehensive system of assessment (Hinchey, 2010:10).

Weighting of Evaluation measures

The different measures should be weighted so that both the evaluator and evaluatee are clear on how each measure will factor into the final rating (TNTP, 2010:6). Danielson and McGreal (2000:37) however argue that it is discretionary to weigh the measures if it has been decided that they are not equally important. TNTP (2010:6)

explains that data sources should not be weighted equally as some are more reliable than other measures, for example teacher designed essay are not as reliable as district-set tests that are more rigorous and allow for comparisons between teachers. Goe et al. (2014:37) concur that the measures differ in their degree of reliability and usefulness so the weighting of the measure should be determined by its validity, impact on student achievement and how informative the measure is on instructional practice. The purpose and goal of the evaluation system should be reflected in the weighting of the evaluation measures, for example, if teacher collaboration is the goal of the system, then a rubric measuring teacher's degree of collaboration with colleagues would have more weight.

3.6.3.5 Step 5: Determining the plan for implementation

There are different approaches that can be used by the evaluation committee to introduce the new evaluation system. It is common that a new evaluation system may be flawed and defective and therefore may need to be corrected (Danielson & McGreal, 2000:73). The implication of this observation is that it is therefore not wise to have full scale implementation with a new evaluation system because of the likely chances of it being defective somehow. Possible options that may be used to introduce a new system are:

- Have volunteers to use the new system and give feedback.
- Sample schools where all the teachers use the system.
- Select few teachers in the district to use the system.
- Using system with teachers that are on probation only.

3.6.3.6 Step 6: Deciding on the evaluators and designing their training programme

The responsibility of teacher evaluation should not be left in the hands of the school principals and administrators alone but rather should involve even the teachers themselves and other concerned parties. The involvement of all the stakeholders in the design and even the implementation of the evaluation process increases the understanding by all the relevant parties and promotes ownership of the evaluation system (Zhang & Ng, 2017:198). Teachers can be involved in the evaluation process by having them as peer evaluators or through self-evaluation procedures. It is argued that although certain researchers have criticised self-evaluation, self-evaluation and

peer evaluation on the other hand besides creating an opportunity for teacher involvement, also creates a platform to reflect on own practice thereby nurturing teacher professional development (Beerens, 2000:41-45).

It is important that the teachers should have high regard for the evaluators by perceiving them to be credible and trustworthy (Stiggins & Duke, 1988:84). It is therefore important that there is effective selection of evaluators and their training to ensure credibility of the evaluation system and ensure that the evaluation results are fair and justifiable. Goe et al. (2014:37) therefore suggest that heads of departments with expertise in the subject areas could be used as evaluators to ensure that appropriate frequency and duration of evaluation is done plus relevant feedback for the subject area is given. A high level of integrity for the evaluators is achievable by providing training combined with feedback and support (Goe et al., 2014:37). Evaluators can earn credibility if they are competent and possess the expertise for their role. They need to be capacitated to make appropriate judgements and concrete recommendations to the teachers (Conley, 1987:62). The task of evaluating teaching practice is complex and requires cognitive application to make consistent judgement. Danielson and McGreal (2000:74) proffer that in any evaluation situation, whether it is a classroom observation or a portfolio, a substantial amount of data is presented and the evaluator has to make a determination on the following factors:

- Which points in the data are relevant for the evaluative criteria?
- Is the evidence collected for a criterion representative or skewed in a certain direction (positive or negative evidence only)

After making the determinations, the evaluator has to make an interpretation on the evidence that has been presented for each criterion and there is often more than one interpretation which is why the evaluator and evaluatee need to have a discussion. After an interpretation has been made, a judgement against the descriptions of the levels of performance has to be made. It is therefore recommended that before introducing a new system, evaluators should be trained so that they can be able to make reliable judgements based on presenting evidence. Goe et al. (2014: 37) posit that the amount of training that is required is dependent on the measures that are used in a particular evaluation model.

Evaluators have to be given a chance to practise observing and judging as part of training. An orientation programme has to be conducted with the teachers who will implement the new system and this should focus on the similar issues to those that are covered in the evaluator training with special emphasis on understanding and demonstrating the evaluative criteria (Danielson & McGreal, 2000:77). Evaluators should be multiple so as to reduce load on the principals and avail more time to them for effectiveness and at the same time improve accuracy of ratings by reducing effects of evaluator bias (White, Cowhy, Stevens & Sporte, 2012:24).

3.7 Implementing teacher performance evaluation

The design of an evaluation system may be commendable but its success is to a large extent dependent on how well it is implemented (TNTP, 2010:9).

3.7.1 Enabling environment for implementing performance evaluation

The successful implementation of an evaluation system is however dependent on the creation of an enabling environment which is composed of various factors. One of the important factors that have to be considered for successful implementation is the development of an evaluative culture in the organisation. Mayne (2010:2) posit that it is fundamental to develop and maintain an evaluative culture in order to be able to set up functional and successful evaluation approaches. Mayne further avows that an evaluative culture values, seeks out, and regards empirical information as fundamental to good management. In discussing organisational change in government, Kim (2002:2) asserts that a change of organisational culture is pivotal for the successful implementation of improvement strategies thereby emphasising the importance of fostering an evaluative culture to promote effective evaluation. An evaluative culture is created by having strong leadership, having the right incentives for people to collect and use performance information, developing and supporting a learning culture and valuing evidence-based information.

Institutional commitment is another requisite factor in creating an enabling environment for successful implementation of evaluation. In support of this, McLaughlin (1990:403) affirms that institutional commitment is key to the implementation and sustenance of a meaningful teacher evaluation program. Conley (1987: 64) asserts that commitment is shown through:

- Setting teaching excellence and student learning as primary purposes of the evaluation process.
- Devoting time to the evaluation procedures.
- Allocating available resources like professional development funds to support the evaluation process.

There is need to direct more resources and personnel towards the teacher performance evaluation process for implementation to occur more smoothly (TNTP, 2010:9). Resources are required to ensure that evaluators receive ongoing training and support on the technicalities of the system and pertinent issues like performance management and on how to have appropriate communication with the evaluatees so that the evaluative process is beneficial. Teachers also need to be availed with the requisite infrastructural and material requirements so that they are able to conduct their duties in a conducive and enabling environment which thus requires more resources.

Evaluator training is also important in setting an enabling environment. Evaluators need to be trained in the processes and uses of the evaluation system so that they are familiar with data collection methods used, can make constructive observations, can analyse the collected data, can conduct pre- and post-observation conferences, can write informative reports and can assist teachers to improve their instructional practice (Conley, 1987:62). Overlooking the training aspect could undermine the entire evaluation effort (Stronge & Tucker, 1999:347). Evaluators need to be credible and trusted by teachers so that the feedback they give to the teachers is respected and accepted (White et al., 2012:13). White et al., (2012:6) proffer that teachers should be trained before a new policy is introduced. Training helps to ensure that teachers understand the system which in turn helps them to appreciate the importance of the system whilst at the same time getting the technical skills necessary for the implementation process. Training should be provided on aspects such as the standards of evaluation to be used, rubrics and the meaning of the various level of performance. White et al. (2012:8) aver that training should be on-going and ideally scheduled for the beginning of the year for the benefit of new teachers and at the same time providing a platform for old teachers to clarify any issues that arise during the implementation process and policy modification.

It is important that new evaluation policy is piloted over a period of time so that the introduction of the system does not bring too much anxiety on the stakeholders and also gives room for the requisite modifications and refinements made in response to the pilot feedback (White et al., 2012:9). Research established that piloting also increases stakeholders' buy-in once they learn from those involved in the pilot that the product is satisfactory. The period of piloting has to be long enough to allow the incorporation of feedback while at the same time minimising an extended and complicated period when implementers have to manage two parallel evaluation systems (White et al., 2012:9).

3.7.2 Implementation procedures

3.7.2.1 Objective setting

Individual teachers are required to set and agree with the school administrators on performance objectives for specified timelines in most teacher performance evaluation models. Systems that allow each teacher to set individualised goals allow the educator to work towards goals that are specific to his/ her classroom scenario. Haefele (1993:28) posits that the goal setting process assists in directing people on what to do and help to focus the task at hand in a particular direction. The setting of the performance objectives is done at the beginning of an appraisal cycle and involves a discussion between the evaluatee and evaluator thereby inducing critical reflection by the two parties (Santiago & Benavides, 2009:14).

3.7.2.2 Performance evaluation feedback

Feedback refers to information given by an agent concerning aspects of an individual's performance or understanding (Hattie & Timperley, 2007:81) and aimed at reducing the variance between current performance and the goal (Hattie & Timperley, 2007:86). Feedback should not be compliance-driven but rather an opportunity for the evaluator and evaluatee to openly and constructively discuss instruction (TNTP, 2010:8).

Hattie and Timperley (2007:90-97) assert that there are four levels of feedback which are, corrective, processing, self-regulation and self. Corrective feedback is directed on the task and avails information on how well the task is being performed. Processing feedback is targeted at processes that underlie a task and self-regulation feedback targets issues of self-efficacy, self-regulation proficiencies and self-belief (it's

motivational). Feedback can also be directed at the 'self' and this is the least effective feedback, for example, 'you are a good teacher'. Feedback about self-regulation and processing are powerful in terms of deep processing and mastery of tasks whereas corrective feedback is only powerful only if it enhances strategy processing and self-regulation which it rarely does. This therefore implies that the most effective feedback is that which is directed at self-regulation and processing levels. Hattie and colleague add that, for feedback to be effective, it has to provide precise information of task at hand, emphasize on the strengths rather than the weaknesses, and also build on previous feedback. This means that feedback has to be continuous and regular and this is only achievable if evaluation activities are continuous and not a one-off event.

3.7.2.3 Monitoring the performance evaluation process

It is acknowledged that even a well-designed evaluation system needs to be evaluated regularly in terms of its goals and results, its structure and processes to ensure system efficacy and sustainability (Goe et al., 2014:47). It is recommended that requisite adjustments should be made on the design or implementation aspects in response to the observations made (TNTP, 2010:9). Goe et al. (2014:47) assert that the system can allow the determination of a number of key issues:

- Whether stakeholders understand and value the system.
- Whether student performance is improving.
- Whether the evaluation process has any impact on instructional practice.
- If teacher retention is improved.
- If the system is implemented accurately as it should be done.

The TNTP (2010:9) further recommends specific metrics to guide the monitoring process as follows:

- Are the evaluators evaluating teachers accurately? The logical position is that evaluation ratings should be reflective of patterns of students' academic development and growth.
- Are teachers improving their performance over time? A working system is expected to assist the teachers especially the inexperienced teachers to improve in competence. External evaluators can be used to verify this so that the risk of rating inflation is minimised or eliminated.

- Are top-performing teachers being retained better than consistently low-performing teachers? The system should have measures that promote retention of the high performing teachers yet discouraging the retention of consistently low-performing teachers.
- Are teachers receiving useful feedback? Teachers should regularly be given a platform to express their views on whether they are getting useful feedback from the evaluation process and administrators of schools where teachers consistently express dissatisfaction should be monitored more closely.
- Do teachers believe that they are being evaluated fairly? Surveys should be conducted regularly to establish if the teachers are confident in the fairness and consistency of the evaluation process and in cases where lack of confidence is expressed, investigations should be conducted.
- Are the school leaders getting the necessary support to conduct accurate evaluations? Surveys of evaluators should be conducted to establish whether they have adequate training, time and resources to implement the evaluation system smoothly. District leaders and human resources department should take responsibility for achieving this goal.

3.7.3 Challenges of implementing performance evaluation

A myriad of challenges have been cited for hindering the smooth implementation of the evaluation process. Some of the challenges bedevilling the implementation process of evaluations are:

Firstly, the absence of an evaluation culture where the stakeholders view the process with suspicion and wariness. A culture of evaluation has to be cultivated where it becomes a norm and tradition for even peers to evaluate each other and share teaching practices. Mayne (2010:6) proffers that a lack of evaluative culture is characterised by:

- Limiting the use of evaluation information to external reporting only and not utilising the information internally.
- Failure to prioritise and avail time for learning in-spite of acknowledging the need for learning.
- Discouraging an experimental approach yet on the other hand emphasising the importance of achieving results.

- Maintaining the status quo and yet claiming to be evidence-seeking.
- Highlighting the importance of achieving results yet on the other hand emphasizing the need to observe rules, processes and output.

An evaluative culture needs to be developed and cultivated for the successful implementation of an evaluation programme (Mayne, 2010:2; Kim, 2002:2).

Secondly, the presence of technical challenges to implementation can present in different forms which include limited evaluation expertise, inadequate preparation for the schools to implement evaluation, limited understanding and appreciation of the purpose and uses of evaluation (Santiago & Benavides, 2009:21). In support of this assertion, Derrington and Campbell (2015:314) in their study established that principals did not know everything which they needed to know at the onset of the evaluation program. This implied that they may not have been given adequate training and induction on the evaluation system before it was introduced. This challenge can be addressed by building evaluator capacity by providing extensive training on the observation process, standards, evaluation tools, the different teacher performance levels, providing formative feedback, and distinguishing the different performance levels (White et al., 2012:13). Evaluators should be competent enough to accurately distinguish the competence levels and provide informative and helpful feedback on instructional practice.

Thirdly, they may be inadequate resources in terms of the time required for the various evaluation processes like observation procedures, feedback and the required developmental work (Derrington & Campbell, 2015:312). Observations made are that implementers can lack adequate time to familiarise themselves with the requirements, rubrics and procedures of a new system and find it difficult to incorporate the processes effectively into the existing work schedule.

Derrington and Campbell, 2015:314) in their study established that principals of schools were forced to reapportion and reprioritise their usual duties as they tried to fit in the new mandatory evaluation responsibilities. The time limitations and increased responsibilities that are often suddenly added to a load that is usually already large can be overwhelming. A new evaluation system brings additional workload to people who are already overwhelmed and this has the tendency to make people overlook the benefits to be derived from its implementation. The challenge of time can be minimised

by ensuring that the evaluation reform initiatives are aligned and become an integral part to the existing institutional initiatives and agenda rather being a disconnected and isolated activity (Stronge & Tucker, 1999:347; White et al., 2012:8). White et al. (2012:24) also add that the evaluation system can be streamlined by decreasing logistical demands on evaluators by investing in technology to decrease the amount of paperwork thereby freeing more time for the evaluators to conduct more observations, be more efficient and provide instructional leadership.

In their study, Derrington and Campbell (2015:314) found out that the principals in the schools complained about continued revisions and clarifications of procedures of the new evaluation system which led to continuous changes of procedures thereby eroding the confidence of the implementers in the system. Numerous revisions in the early stages of a newly introduced system could be indicative of a system that had been prematurely introduced before adequate piloting had been done to allow the processes and procedures to be refined. White et al. (2012:9) assert that piloting a new evaluation system before full scale implementation is important in order to give room for the necessary modifications and refinements which will in turn minimise the revisions that need to be made especially during the early phases.

Another challenge that has been identified concerns issues to do with inconsistency and subjectivity of the evaluation process during implementation (Derrington & Campbell, 2015:317). Subjectivity could be attributable to inadequate training on the part of the evaluators resulting in low inter-rater reliability or it could mean that the instrument is not reliable. Inter-rater reliability can be increased by providing evaluators with hours of training on aspects such as the standards, rubrics and the meaning of the different levels of performance (White et al., 2012:6). The reliability of the instrument could also be addressed by piloting the instrument prior to full implementation so as to establish its reliability and taking the necessary measures to increase reliability if necessary.

An important observation made is that support of a new programme should be ongoing, rather than a one-time event and should be responsive to the needs on the ground as the needs are more likely to evolve through the different phases of implementation (Derrington & Campbell, 2015:322). Santiago and Benavides (2009:22) suggest that one way to address implementation challenges is to have an

advisory group composed of educational researchers and experienced teachers to monitor as well as guide the implementation of teacher evaluation. Since evaluation is more of a technical process, such a group can tap from global best policies and practices generated from empirical evidence so that they improve the evaluation process. It has also been suggested that the evaluation committee that is set at design phase can be maintained so that it can monitor and evaluate the implementation process, gathering feedback and communicating the concerns as a way of ensuring fidelity (White et al., 2012:6). Failure to support the change can be frustrating to the implementers as well as distracting them from their key roles whereas on the other hand support can facilitate a faster rate of progression to advanced levels of implementation (Derrington & Campbell, 2015:322).

3.8 Significance of performance evaluation process

An evaluation process should have meaningful implications for it to earn support from the implementers (TNTP, 2010:8). This implies that the information that is produced from the evaluation process must be used in ways that can enhance performance and make the teachers accountable. One of the key objectives of teacher evaluation is to identify the weaknesses and shortcomings of a teacher so that targeted improvement plans and professional development can be organised as a response to the findings of the evaluation process (Santiago & Benavides, 2009:19). It is the responsibility of the instructional leaders who can be the school heads or the heads of departments to ensure that the link between the evaluation and professional development programs are created.

3.8.1 Making evaluation significant by linking it to professional development activities

A significant body of research has demonstrated that the most important factor that affects student achievement at school level is teacher quality (Choi & Park, 2016:153; Darling-Hammond, 2000:39; Looney, 2011:440). In order to remain relevant to the ever changing curricula and knowledge, it is imperative that teachers also continue to upgrade their knowledge and skills as they progress in their career. Literature indicates that teacher evaluation systems that are well-designed and have professional development programs that are aligned to evaluation findings can go a long way in enhancing the quality of teaching and in turn raise student achievement (Looney,

2011:440). On the contrary, literature states that while evaluation systems have been good at identifying teachers' weaknesses, they have however been limited in transforming the identified weaknesses into strengths (White et al. 2012:11). Indications are that most teacher evaluation systems are not effective and this is attributed to a myriad of factors that include approaches that are haphazard, having evaluators that are not trained, having ineffective tools and not aligning professional development programs to the evaluation findings.

Kennedy (2010:4) observes that the current trend is that whenever a problem is identified with the teachers, one of the solutions often proposed is the introduction of a new assessment system. Research has however not produced enough evidence to support this logic. This has been disputed by Murphy, Hallinger and Heck (2013:350) in their argument that recent teacher evaluation systems have not shown a capacity to promote school improvement in relation to student learning implying that the introduction of a new evaluation system may not necessarily be the answer to the exhibited problems. A myriad of factors have been blamed for this lack of impact and one of the factors which has been consistently cited is the troublesome relationship between evaluation and professional development (Smylie, 2014:98).

Professional development refers to the opportunities that are availed to teachers to enable them to learn and improve their practice on the basis of evaluation findings. Smylie (2014:98) has come up with two perspectives to the issue of professional development and teacher evaluation. The two issues that have been identified are the 'weak link problem' where the importance of professional development is not being emphasized in the evaluation process as a way of facilitating teacher improvement. The 'weak link problem' is evident where vague and weak provisions for professional development are made. Policy might not have a clear cut standing on the aspect of professional development thereby leaving it to the schools or districts to make the linkages themselves. The other dimension to the professional development aspect is the 'weak quality problem' where the development opportunities, although they may be clearly stipulated in the evaluation system policy, may however be of poor quality, and therefore not really be of any value to aiding the improvement of the actual practice. The argument that is advanced here is that while the issue of evaluation and professional development is not only problematic in the link of evaluation to

professional development opportunities, it is also problematic in the quality of the opportunities that are availed on the basis of the evaluation process.

America has advanced different policy approaches to professional development advanced through the No Child Left Behind (NCLB) and Every Student Succeeds Act (ESSA). Although both the acts touch on professional development, there was a policy shift on how to approach professional development of teachers. Whereas the NCLB which was replaced by ESSA defined professional development as a way to basically transfer knowledge to the teachers, ESSA on the other hand promotes a continuous approach to professional development (Rosen & Parise, 2017:2). The NCBL had an intermittent or sporadic approach to the issue of professional development but the ESSA has a job-embedded approach which is a continuous approach. .

Goe et al. (2014:43) concur to the importance of linking professional development to evaluation findings through their assertion that the most significant phase of the evaluation cycle is the use of evaluation results to support professional development. Evaluation should be linked to professional development to advance the improvement function and to focus support for all teachers at various levels of performance (White et al., 2012:24). Darling–Hammond (2013:99) also supports this notion by positing that evaluation can only be most effective if it is linked to both formal professional development and job-embedded learning opportunities. She further emphasizes that,

Evaluation alone will not improve practice. Productive feedback must be accompanied by opportunities to learn. Evaluations should trigger continuous goal-setting for areas teachers want to work on, specific professional development support and coaching, and opportunities to share expertise as part of recognising teachers' strengths and needs.

All this points to the importance of ensuring that evaluation findings should inform professional development activities if the evaluation process is to have impact in improving the teaching and learning process. Evaluation ratings coupled with staff surveys can be used to inform targeted professional development programs (White et al., 2012:13). Murphy et al. (2013:352) support this argument by their assertion that instructional improvement can only be achieved if evaluation and other bundles of action, which include giving actionable teachers feedback, creating contexts for

collaborative work and developing a system where teachers have the opportunity to develop and refine their skills. Individual, school and even district wide development needs can be identified through evaluation results and resultantly, targeted professional development can be implemented (Goe et al., 2014:43). Care should however be taken to ensure that professional development programs are evaluated by stakeholders to determine their effectiveness on improving teaching practice through a continual analysis of teacher participation, support and outcomes related to student achievement (Goe et al., 2014:43). This position resonates with the ESS), signed into law on 10 December 2015 in America as a replacement to the NCLB Act of 2002, as it calls for professional development that is job-embedded while at the same time being aligned to the school or district improvement plan. ESSA also calls for continuous professional development that is targeted towards identified needs, professional development that has teacher input from active teacher participation, and professional development that should be evaluated on a regular basis to assess its influence on effectiveness of teachers (Rosen & Parise, 2017:1).

Research has however established that in spite of the continual improvement approach as envisioned by ESSA, districts can still do more to improve their existing teacher evaluation systems to inform and guide the teachers' professional development (Rosen & Parise, 2017:6). Rosen and colleague explain that this could be achieved by ensuring that districts assist the school leaders to conduct evaluations that provide tangible suggestions on how teachers can improve their practice, leaders that are equipped enough to identify the professional needy areas of teachers, and leaders who avail appropriate learning opportunities to the teachers.

3.8.2 Making evaluation significant by ensuring accountability by the teachers

If the performance of teachers matter, it is important that findings made on this very aspect should lead to significant decisions that influence the quality of instruction that is availed to the learner. Good teaching should be acknowledged and valued if it is to be maintained (TNTP, 2010:2). Evaluation information can therefore be used for high stakes decisions like the tenure, compensation, promotion or dismissal, thereby making it critical that the results are clear, accurate and easy to interpret (TNTP, 2010:8). Santiago and Benavides (2009:19) weigh in by proffering that, to retain effective teachers and make teaching an attractive career choice, then, findings from

teacher evaluations can be used to inform decisions on career advancement, performance rewards or establishing sanctions on a teacher. Such decisions however do not have to be made on the spur of the moment where a single observation leads to such high stake decision, but rather should be underpinned by well-informed professional judgement based on multiple observations.

Santiago and Benavides (2009:19-20) posit that teacher evaluation consequences might be at following levels:

- a) **Key points in career-** Tenure decisions at the end of probation or contract renewal decisions can be informed by teacher evaluation findings.
- b) **Career advancement-**It has been noticed that most countries link evaluation results to career progression rather than directly to salary. Evaluation results can determine the rate at which a teacher progresses through the career ladder with the top rating leading to faster progression and lower rating—stalling progression.
- c) **Performance rewards-** Rewards may be given to acknowledge effective teaching established through evaluations in form of a bonus one off monetary prize, support for post graduate study, time allowances or opportunities for in-service education. Rewards can be given at departmental or even school levels rather than at individual levels. It is however cautioned that linking performance to monetary rewards needs to be approached with extra care so that it is not contentious or become divisive.
- d) **Sanctions for ineffective teachers-** Ineffective teachers may be sanctioned for persistent under-performance and this can include the removal from teaching duties to other duties within the school or education system or even the termination of service.

There is however a school of thought that is against high-stake evaluations. In their argument against the use of material incentives as a way of motivating teachers and stimulating desirable behaviour, Finnigan and Gross (2007:596) refer to the three categories of incentives proposed by Clark and Wilson (1961). According to the two, incentives can be categorised into material, purposive and solidarity incentives.

Material incentives refer to the tangible monetary rewards for desirable performance and penalties for undesirable performance. Purposive incentives on the other hand

involve the satisfaction that an individual gets by achieving valued goals which in the case of a teacher is the student succeeding. Solidarity incentives refer to the motivation that one gets from his/her professional status (satisfaction derived from belonging to a professional community). It is argued that in the effort by an organisation to induce desirable behaviour, it is important to understand the interaction of individual internal motive with the external incentives so as to understand the effect of incentives on the individual's motivation which in turn translates to their effort towards a task (Finnigan & Gross, 2007:596; Muralidharan & Sundararaman, 2011:396). Psychological literature indicates that if workers perceive extrinsic incentives as a form of control and as likely to interfere with their professional behaviour, then intrinsic motivation will likely be forced out, while on the other hand, if the extrinsic incentives are regarded as strengthening professional behaviour, then they can enhance intrinsic motivation. This therefore implies that caution should be observed in designing and framing an incentive programme so that its efficacy and perceptions that teachers have about it are positive.

Finnigan and Gross (2007:596) established that although teachers increased their effort by trying new strategies and even by engaging in professional development activities, the teachers however responded less to the material incentives built in the Chicago Probation Policy and more to the value that they placed on their professional status and the individual goals that they had for the students to succeed. This is contrary to the use of material incentives as a way of motivating teachers to improve their performance as encapsulated in the theories of personnel economics that postulate that compensation can:

- a) Serve as a strong motivational incentive to boost organisational performance by stimulating employees to improve their practice.
- b) Improve the composition of the workforce by luring and retaining high performers while at the same time discouraging low performers from entering or staying in the system (Pham, Nguyen & Springer, 2017:3).

Incentives like merit pay besides motivating teachers to improve their performance and in turn that of the students, also provide a theoretically appropriate solution to the problem of limited monitoring of the teachers' performance by administrators and concerns attributable to qualification and experience (Pham et al., 2017:3). Also

supporting the position are study findings by Muralidharan and Sundararaman (2011:402) which established that linking a part of the teacher's pay to objective measures of performance is likely to be popular with the teachers and also likely to improve student learning outcomes. Research has thus established that merit pay in fact encourages teachers to work harder as promulgated by the principal agent theory (Pham et al., 2017:29).

3.9 Evaluation in Zimbabwe

This section presents a review of the Statutory Instrument 1 of 2000 which regulates the appraisal process in Zimbabwe before giving an overview of related studies on the Zimbabwean evaluation system. Although a number of documents that are linked to evaluation were reviewed to get the necessary background information to develop the data collection instruments, I will confine myself to the statutory instrument. I however noted from that review that the Ministry of Primary and Secondary Education seem to be running two parallel performance evaluation systems, one originated by the Public Service Commission and the other by the Ministry of Primary and Secondary Education itself. This study is however limited to the Results based management (RBM) system.

3.9.1 Statutory Instrument 1 of 2000

Statutory instrument 1 of 2000 was released on 3 January of 2000 as a supplement to the Zimbabwe Government Gazette Extraordinary dated 3 January 2000 (2000: 72). The document was originated by the Public Commission of Zimbabwe and it was created to regulate the conduct of the public servants serving under the Commission. Statutory Instrument 1 of 2000 contains regulations on different issues that affect the conduct of civil servants including performance appraisal.

3.9.1.1 Purpose of evaluation

The purpose of the evaluation process is twofold, to continually review and improve the quality of service provided to the public and the performance, potential and development needs of all members (Zimbabwe Public Service Regulations, 2000:12). The evaluation process therefore has both an accountability and professional development function since it aims to improve service and performance.

3.9.1.2 Character and nature of evaluation

According to the Zimbabwe Public Service Regulations (2000:12), all members have to be assessed continually and final reports have to be submitted once a year. This implies that every member including those who may be employed on contract basis are also supposed to be appraised since the regulations does not single them out but simply states that all members. What can be noted is that the process is supposed to be continual although a final report is to be submitted once every year.

3.9.1.3 Methods of evaluation

The regulations are not specific about the sources of evaluation information but simply states that regular interviews have to be conducted with the member concerned and reported in a form prescribed by the Commission (Zimbabwe Public Service Regulations, 2000:12).

3.9.1.4 Consequences of evaluation

The action to be taken on the basis of the appraisal is discretionary to the Commission, Head of Ministry or Department but includes:

- Advancement or promotion.
- Transfer to a post commensurate with the member's competence.
- Participation in targeted skills development program.
- Granting or withholding of any performance award.
- Demotion or discharge in line with disciplinary procedures in Part VIII of the regulation (Zimbabwe Public Service Regulations, 2000:12-13).

3.9.1.5 Critique of Public Service Regulations

A number of issues were noted during the review of the statutory instrument. It should be noted that the Zimbabwe Public Service Regulations apply to all member of the civil servants irrespective of their ministries or departments. The regulations are therefore not particular to teachers which might make them rather weak in regularising the teacher performance appraisal process of teachers considering the complexity of the teaching task.

The other issue is that although possible consequences of the appraisal findings were stated, the statute does not, however, specify the performance level and the resultant

consequence. There is, therefore, no clarity on the consequence of the different levels of performance leaving it open and discretionary upon Commission, Ministry or Department since they are the ones that will decide the action to be taken on the basis of the appraisal results at the end of an appraisal cycle. This lack of clarity makes it impossible for a member to make any claims on the basis of performance. This makes the statute potentially biased on this aspect in favour of the employer since the resultant action is decided by the employer, a position which can be abused in favour of the employer. This contradicts the suggestion that indications of when action to be taken on the basis of the evaluation findings will be actuated should be specified from the outset (Goe et al., 2014: 43). This lack of clarity can also affect the consequentiality of the appraisal process thereby rendering it insignificant. The statute however lists the possible consequences of the evaluation processes which is in line with indications in literature that appraisal processes should have meaningful implications if they are to gain continued support from stakeholders and contribute to the improvement of the workforce (Santiago & Benavides, 2009:19; TNTP, 2010:8).

Another issue noted is that the criteria to be used in the appraisal are not mentioned in the statute and neither is there a referral document on this aspect. It is also silent on the sources of appraisal information to be used in the process. It simply talks of interviews being conducted but is silent on the sources of information on which the interviews are to be based on. There is no other document that the statute makes reference to for these pertinent issues.

The Statutory Instrument does not stipulate the frequency of appraisal interviews. It refers to regular interviews being conducted between appraiser and appraisee. This vagueness may compromise the quality of appraisal data if the frequency of contacts between appraiser and appraisee are inadequate.

In spite of the weaknesses or omissions noted above, it is however acknowledged that the statutory instrument has a laid out grievance procedure to be pursued in the event of being aggrieved (Zimbabwe Public Service Regulations, 2000:50-52). This is commendable for such critical information to be availed openly to all stakeholders to promote fairness and justice.

The statutory instrument also indicates that circulars or notices governing both content and conduct of performance appraisals may be issued by the Commission when necessary (Zimbabwe Public Service Regulations, 2000:13). This is positive as it points to the realisation that the statutory instrument may have omitted or failed to adequately cover key aspects that may need to be addressed through circulars or notices. This is why in any teacher evaluation system, periodic reviews should be done to address omissions at design stage as well as to realign the evaluation process to any new needs that crop up with changing times.

3.9.2 RBM system in Zimbabwe

Since independence in 1980, Zimbabwe has undergone several reform programs aimed at improving public service delivery and is currently using Results Based Management (RBM) system to measure the performance of its workers (Zvavahera, 2014:1). Performance measurement involves assessing the extent to which set goals and objectives have been met against set standards, that is, the actual performance relative to the desired performance (Zvavahera, 2014:4).

3.9.2.1 What triggered adoption of RBM in Zimbabwe?

Madhekeni (2012:124) established that in Zimbabwe, RBM was triggered by the following:

- Search for better quality and responsive service delivery by the public.
- Resource constraints.
- Call for politicians to be more service oriented and people sensitive.
- Arguments for more effective resource allocation by financial controllers.
- Insistence by private sector for improved services, infrastructure development, and growth.
- Demands by donors for efficient and effective use limited funds.
- Increased challenges of the competitive and globalised world.
- Call for increased accountability, transparency and results from government.

Literature indicates that a number of benefits are derived from using RBM as a performance evaluation system. Zvavahera (2014:4) asserts that the benefits of using RBM are:

- The system allows employees to set their own goals and objective. At the beginning of each appraisal cycle, performance targets for the particular appraisal period are discussed and agreed on by the evaluator and evaluatee.
- Employees develop their own work plan which enhances their responsibility and commitment to the achievement of the plan.
- Employees are recognised and rewarded for outstanding performance.
- Approach is results oriented and aims to enhance employee motivation and retention.
- A fundamental aspect of RBM is staff development and training. The appraisal process is expected to identify staff weaknesses and training programmes are to be developed in line with the identified needs.
- Improves evaluator-evaluatee relationship due to constant communication.

3.9.2.2. Achievements that have so far been derived from using RBM

Although there has been some controversy as regards the issues of applicability, benefits, and drawbacks in the Zimbabwe context that is fraught with a number of institutional, organisational, and systemic weaknesses nullifying government efforts, research on the other hand has shown that if implemented properly, then it can improve government performance. Madhekeni (2012:125) has listed the achievements so far derived from the use of RBM as a performance evaluation system as follows:

- Improved funding opportunities from bilateral institutions and donors due to demonstration of commitment to the accountability principle and well-being of the people.
- Monitoring and evaluation is now done in a systematic and logical manner and producing more objective findings.
- Uniform reporting by the different structures.
- Brought decentralisation where results are assessed at provincial and district levels.
- Decentralisation allows results to be monitored closer to the field where they originated from.

A number of related studies were reviewed and these came up with a number of findings, conclusions and recommendations pertaining to the areas of focus for this

study which are: appropriateness of timing of the introduction of the system, the design process, implementation process and significance of the process.

3.9.2.3 Readiness Assessment on introducing RBM system

Research has indicated the importance of appropriate timing in introducing a new performance evaluation system (Kusek & Rist, 2004:40). Timing is an important factor to consider because hurried implementation may increase insecurity and fears whilst on the other hand, slow implementation may result in the process being overtaken by events (Madhekeni, 2012:126). These assertions emphasise the importance of appropriate timing in introducing a new performance evaluation system.

On the aspect of timing, a number of findings were made. Mavhiki, Nyamwanza and Dhorro (2013:137) posit that government workers' work culture has affected buy-in of RBM. The government workers are not results-oriented and not keen to account for their work which is contrary to RBM principles of accountability, transparency and demonstrable results. There is still a need to change the mind-set of stakeholders so that they can understand and appreciate the rationale for managing results and improving service delivery (Madhekeni 2012:125). The issue of work culture resonates with the assertion by Mayne (2010:2) that developing and maintaining an evaluation culture is key to setting up of functional and successful evaluation approaches which in the Zimbabwean context was missing when RBM was introduced as per the research findings. Readiness assessment of institutional evaluative culture was important to ascertain the demand for the program.

A second issue that has been raised pertaining to timing is that of lack of buy-in by both top management and low level workers. Machingambi (2013:221) established that the performance evaluation system was not communicated properly to most teachers before its introduction. It emerged that the system had simply been imposed which resultantly led to teachers not having a sense of ownership for the programme and also developing a negative attitude towards the system thereby affecting the outcome of the system. The position that teachers neither understand nor appreciate the objectives of the system and thus have not embraced it and still regard it as belonging to the authorities is reiterated by Gutuza (2016:395) and Madhekeni (2012:125).

Indications are that even the top management in the form of ministers, parliamentarians and even commissioners have poor buy-in due to poor understanding and appreciation of results management and therefore lack the requisite political will (Madhekeni, 2012:125). Research indicates that top leadership commitment which is necessary for a change process to survive resistance was missing and the fact that the leadership neither understood nor agreed with the system objectives is blamed for their lack of commitment and ultimately lack of political will (Mavhiki et al., 2013:136; Musingafi, Dumbu & Chadamoyo, 2013:260). The lack of buy-in by the key stakeholders could imply that the introduction was done hurriedly or prematurely before adequate marketing of the system had been done for appreciation and acceptance of the system.

The third issue that could also have negatively affected the appropriateness of the timing of the introduction is the issue of resource availability. Resource constraints led to most of the employees not getting the requisite training to allow them to be conversant with the new system (Machingambi, 2013:223; Madhekeni, 2012:125). Research established that teachers were incapacitated, lacked skills due to the inadequate training that they had received and had no motivation to implement the performance evaluation system (Gutuza, 2016:395; Machingambi, 2013:223). Besides the training being hampered by the lack of resources, other aspects like marketing, procurement of the necessary material resources and fiscal resources for supporting the system through provisions of incentives and funding of targeted staff development programme was also affected thereby affecting the appropriateness of the timing of the introduction of the system. Pretorius and Ngwenya (2008:156) emphasize the point that a programme has to be marketed to the implementers so that they are made conscious of its benefits if the stakeholders are to embrace and accept it. Ascertaining fiscal capacity of the government as readiness assessment would have helped to determine the appropriateness of the introduction of RBM program.

3.9.2.4 Design Process of RBM system

Research has established that teachers are more likely to buy into a system if they are involved somehow in the design process rather than if they are excluded (Goe et al., 2014:14; White et al., 2012:12). In a study conducted by Gutuza (2016:394), 81% of

the teachers felt that they had been side-lined during the policy formulation phase. The lack of teacher involvement in RBM formulation had the effect of discouraging teachers from embracing and accepting the system.

Goe et al. (2014: 14) proffer that teacher involvement is fundamental since they are the experts and are therefore better placed to tell what can work or not work in the teaching-learning process. This implies that excluding the teachers as the experts from the design process could therefore lead to a system that is poorly designed and thus ineffective in achieving its purpose. This is especially correct in the case of the education sector which is a specialised sector. Mavhiki et al. (2013:137) concur to this by asserting that the design of RBM in Zimbabwe has been criticised on performance indicators of certain job categories and departments like the education sector where teachers have challenged the use of pass rate as a performance measure. Such a scenario could have been averted by enlisting teacher participation in the design aspect. Machingambi (2013:223) asserts that, teachers have to participate and be involved in the design and development of the instruments and procedures that are used to assess their performance so that they can be motivated, empowered and be cognisant with the evaluation procedures. Adopting a participatory approach from the design phase helps to ensure that the stakeholders identify with the system and may even become its advocates. It is however evident from a number of studies that in the Zimbabwean situation, stakeholder involvement has been very limited if any (Gutuza, 2016:394; Machingambi, 2013:223) and this has ultimately led to lack of ownership and buy-in from the stakeholders resultantly compromising the effectiveness of the system. It is declared that imposing a programme on people is a recipe for disaster and a guarantee for failure (Musingafi et al., 2013:26).

3.9.2.5 Implementing RBM

According to the TNTP (2010:9), “The success of any evaluation system—no matter how solid its design—ultimately depends on how well it is implemented”. The implementation process is the process by which the design is put into action, so no matter how good a design, if it is not well actioned, then it will come to naught. An evaluation design is similar to the plan of a house, which can be very good on paper and yet if the builder fails to follow the plan properly, then the house will not be as good

as it is on the plan. The implementation process is therefore critical too, since it can ruin a good job done at design phase.

Gutuza (2016:394) established a number of challenges hounding the implementation process:

- a. Teachers complained about RBM having too much paperwork which in turn compromised their service provision as they were forced to do a lot of writing at the expense of the actual teaching. The time taken in completing the paper work eats a lot into the time that could otherwise be used in preparing for teaching and related activities. White et al. (2012:24) allude to this fact and suggest that to decrease paperwork and increase efficiency, it is critical to invest in technology so that less time is spent on writing.
- b. Teachers also voiced concern on the fact that school heads did not give them adequate support on RBM as they appeared to be pre-occupied with other administrative issues excluding RBM and only reminded teachers to fill the forms at the end of the evaluation cycle. This implies therefore that the evaluation is just an 'event' issue since there appears to be no activity in between the development of the work plan and the rating process done at the end. The implementation process is therefore doomed due to the fact that the Heads who are supposed to spearhead the implementation process have no zeal and zest for the task. This tends to make the whole exercise a fallacy. Mavhiki et al. (2013:137) concur to that by citing lack of enthusiasm by senior leaders as having affected the implementation process negatively thereby leading to implementation failure.

Other researchers have also established other impediments to the implementation process. Lack of financial, human and physical resources have been pointed as hindrances to the implementation of RBM across the different sectors in the civil service of Zimbabwe (Mavhiki et al., 2013:138). In the case of Zimbabwe, this was supposedly worsened by the fact that the funding partner had withdrawn its support and this ultimately affected the training programme and capacity of the driving ministry to impart the requisite skills. Machingambi (2013:223) also established that most teachers had not been trained on the system except a few who had been given one-off training which thus affected their understanding of how the system works and in

turn their ability to implement it. The lack of training and skills on the part of the teachers and the heads greatly hampers the effectiveness of the evaluation system (Machingambi, 2013:223).

Machingambi (2013:224) established that lack of resources impeded the success of the implementation process. Financial resources are required to procure materials for teaching and to support developmental programs which assist the teachers to improve their instructional competence. The situation was worsened by the hyperinflationary environment that existed and reached its peak in 2007-2008 (Madhekeni, 2012:126).

3.9.2.6 Significance of the RBM evaluation process

Literature indicates that performance information should be utilised and be seen to be utilised to give the system credibility (Madhekeni, 2012:126). An evaluation process should be significant and consequential for it to earn sustained support and respect from teachers and school leaders (TNTP, 2010:6). Zvavahera (2014:6-7) established a number of findings pertaining to the consequentiality of the performance evaluation system in Zimbabwe as follows:

- Employees are not appreciating the system as it is not linked to any concrete benefits.
- Although a number of employees are undertaking a number of courses, most of the courses were however not linked to performance gaps identified during the appraisal cycle which thus indicated a weak link between the evaluation process and professional development activities. The evaluation system is therefore not seeking to identify performance gaps so that it provides targeted development training so as to enhance efficiency and effectiveness as enshrined in Zimbabwe Public Service Regulation (2000:13). This is contrary to the assertion that effective teacher evaluation allows teachers to gain more knowledge and skills which in turn enhances their confidence and instructional competence (Musingafi et al., 2013:438).
- Although Zimbabwe Public Service Regulations (2000:12-13) indicates that performance awards, advancement and promotions may be awarded on the discretion of the Commission, Head of Ministry or Department, Zvavahera established that there is no meaningful relationship between the listed aspects. Indications are that salary increments, promotions and advancements are done

haphazardly with no connection to performance thereby leading to the employees feeling disgruntled. Mavhiki et al. (2013:137) have corroborated the issue of lack of performance rewards as a challenge by indicating that both drivers and implementers have highlighted that the issue of incentives has affected implementation especially as any salary rise that has been awarded has been given to all members irrespective of the RBM rating.

Although the RBM drivers have attributed the lack of incentives to the financial constraints currently being experienced by government (Mavhiki et al., 2013:137), indications are that the chances of service delivery improving is next to zero unless workers are rewarded according to their performance (Zvavahera, 2014:7).

3.9.2.7 Ways to enhance RBM

A number of studies that have been conducted have come up with different ways that can be used to enhance the RBM as an evaluation system. These recommendations include:

- Situational analysis should be conducted before adopting a new system (Mavhiki et al., 2013:138; Musingafi et al., 2013:26).
- The Minister of Labour and Public Service should market the system so that stakeholders are made aware of its benefits and merits (Gutuza, 2016:395).
- Government to mobilise resources to support the programme (Mavhiki et al., 2013:138).
- To adopt a selective approach to RBM to cut down on implementation costs (Mavhiki et al., 2013:138).
- Government and policy makers to increase political will and commitment to the programme (Machingambi, 2013:224; Musingafi et al., 2013:26).
- To develop an RBM model customised to the services sector (Gutuza, 2016:395; Machingambi, 2013:224; Mavhiki et al., 2013:138).
- A participatory approach to be adopted at all phases from design phase so that stakeholders identify with the system (ownership) and can even become its advocates (Machingambi, 2013:224; Pretorius & Ngwenya, 2008:162).
- Piloting should be done before full implementation (Madhekeni, 2012:126; Mavhiki et al., 2013:138).

- Training of all employees at all levels to promote understanding and appreciation of the system (Gutuza, 2016:395; Machingambi, 2013:224; Madhekeni, 2012:126; Mavhiki et al., 2013:138; Zvavahera, 2014:7).
- Train school heads on proper use of RBM as a supervision tool and how they can use it to enhance their role as instructional leaders (Gutuza, 2016:395).
- Utilise performance information (Madhekeni, 2012:126).
- Incentivise the system (Gutuza, 2016:395; Madhekeni, 2012:126). Madhekeni however also cautions that care should be taken when giving incentives so that they are not perceived as rewarding good performance and penalising poor performance but rather regarded as a learning opportunity.

3.10 Chapter Summary

This chapter presented a review of related literature and perspective with the aim of setting the reader and study into perspective. The chapter was introduced by explaining the concept of performance evaluation. A global perspective of evaluation was presented by giving a detailed review of three selected countries namely South Korea, Chile and China. Notable features of these country evaluation systems included aspects on readiness assessment, design processes, the implementation processes and the significance or consequentiality of the process. An interesting feature of the teacher evaluation system of South Korea was the attempt that it made on having a participatory design approach and then abandoning it after stakeholders failed to reach consensus. South Korea then proceeded to pilot two versions of the evaluation system in order to get input from the stakeholders which was factored in and also used to promote a sense of ownership among the stakeholders. The interesting aspects of the teacher performance evaluation system of Chile was its design process which was participatory. The evaluation systems of South Korea and Chile also have differentiated and targeted professional development programmes. A thought-provoking feature of the teacher evaluation system of China was the bottom up design approach, the use of students as evaluators and team awards. Literature, mostly based on American literature was also reviewed based on the four major themes derived from the RADPS conceptual framework, that is, appropriateness of timing of the performance evaluation system, the design process, the implementation process and the significance of the evaluation process. Related literature on RBM

system was also presented along the same themes that are used in the main body of the review but with specific focus on the Zimbabwe case.

CHAPTER FOUR

Research Design and Methodology

4.1 Chapter Overview

This chapter presents and discusses the design and methodology used in the study. The epistemological and ontological positions of the researcher which outline the position from which the research process is entered into and managed is also outlined herein. The positions lead us to the research paradigm of the study before a discussion of the research methodology. The methodology section presents a detail of the research design, the sampling processes and procedures used, data collection methods, data analysis, and interpretation procedures. The measures that were adopted to ensure study rigor and validity plus the ethical considerations observed throughout the study are also presented in this chapter. The chapter is concluded with a summary which sums up the chapter's key issues.

4.2 Introduction

This study sought to assess the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. The aim of the study was achieved through the following specific objectives which sought to:

- Investigate the relationship between the timing of the introduction of the teacher performance evaluation system and its effectiveness.
- Examine the design process of the teacher performance evaluation system in Kwekwe district.
- Appraise the implementation process of teacher performance evaluation system of Kwekwe district.
- Assess the significance of the teacher performance evaluation system of Kwekwe district in promoting teacher accountability and professional development.

A research study is guided by a research approach. A research approach refers to a set of plans and procedures for a study starting from the general assumptions to the detailed methods of data collection, analysis and interpretation (Creswell, 2014:3). In this study, the research approach includes the philosophical assumptions that I bring

to the study as the researcher, the research design, the specific data collection methods and instruments, data analysis and interpretation.

In deciding on the research paradigm and methodology for the study, I took a cue from the suggestion that it is prudent to conduct ‘question-led’ research rather than ‘method-led’ inquiry (Grix, 2002:180; Krauss, 2005:761). The research questions should guide the research approach that is adopted for an inquiry depending on the philosophical assumptions that the research holds on entering the research landscape. The philosophical assumptions which are the ontological and epistemological perspectives of the researcher guide the research approach that is adopted to conduct and manage a study. The ontological and epistemological assumptions define the inquiry paradigm (Mack, 2010:5). The research paradigm in turn shapes the methodological approach which includes the research design, research methods and analysis procedures that are adopted for a study (Tuli, 2010:105).

4.3 Research Philosophical assumptions

According to Quinlan et al. (2011:95), it is a norm, that every research study is underpinned by a philosophical framework that shows the worldview in which the research is located, and, is thus exhibited throughout the research process. As I enter the research landscape, it is important to reflect on the philosophical assumptions that I bring with me which shape the way I approach the inquiry process. Literature indicates that ontological, epistemological and methodological questions should be key features in a discussion about the nature of social science research. These are critical in shaping and defining the conduct of an inquiry (Popkewitz, Tabachnick & Zeichner, 1979:52). These three questions types define a paradigm and are connected in such a way that an answer that is given to one of the questions automatically directs the answers that can be given to the other questions. Answers to the ontological question determine the answers to the epistemological question which in turn influence the answers to the methodological question (Guba & Lincoln, 1994:108). This discussion on research design and methodology would therefore not be complete without mentioning my position as regards these key issues.

4.3.1 Ontological research proposition

The purpose of any enquiry is to search for knowledge and the reality of that which is sought to be known. In my bid to find out about the effectiveness of the teacher

evaluation system of Kwekwe district, I had to reflect on what I consider to be the reality of what is to be known, which my ontological assumption is.

Blaikie (2000:8) proffers that ontological claims are:

“claims and assumptions that are made about the nature of social reality, claims about what exists, what it looks like, what units make it up and how these units interact with each other. In short, ontological assumptions are concerned with what we believe constitutes social reality”.

Ontological claims are concerned with the nature of the social phenomenon being studied. These claims consider whether the social reality is external to the individuals, that is, whether it is objective or it's constructed and a result of a creation of meaning by one's mind (Cohen et al., 2011:3). There are two opposing ontological perspectives, positivism and constructivism. Positivism asserts that reality exists independent of the knower, that is, it is 'out there', it is real, absolute and singular whereas constructivism assumes that there are multiple realities that are socially constructed, and also informed by personal experiences (Cohen, Manion & Morrison 2007:7; Creswell & Plano Clark, 2011:41; Krauss, 2005:760). Bryman (2001:16-18) also adds that social realities are not static but are continuously revised thereby making them dynamic and contextual.

4.3.2 Epistemological research proposition

Epistemological claims relate to the relationship between the researcher and that which is being researched (Creswell & Plano Clark, 2011:42). The epistemological assumptions involve answering questions on how the researcher should relate to the researched, that is, how a person can acquire knowledge (Guba & Lincoln, 1994:108; Krauss, 2005:759). It is in short, the knowledge theory that one holds. There are again two opposing epistemological positions, that is, objectivism and subjectivism. Positivists believe that knowledge is 'hard, objective and tangible' (Cohen et al., 2011:6; Cohen & Manion, 1989:7). This position holds that the researcher is independent from the research process and that knowledge should be gained from direct observations or measurements where a unitary approach to inquiry is adopted by examining the component parts of the research phenomenon (Krauss, 2005:759). Constructivists/interpretivists on the other hand claim that knowledge is personal,

subjective and unique (Cohen & Manion, 1989:7; Cohen et al., 2011:6). This position asserts that knowledge is established by constructing meanings that are attached to the phenomenon under the study thereby making it dynamic. Knowledge is supposedly affected by context and time since it is socially constructed. Constructivists/interpretivists assert that the researcher interacts with the subject to get the knowledge thereby becoming involved or immersed in the research process (Krauss, 2005:759). The differences between positivists and constructivists is that the former believe that knowledge is objective and that the researcher should remain detached from the research process whereas the latter believe that knowledge is subjective and that the researcher becomes immersed and therefore part of the research process.

4.3.3 Methodological research proposition

The other aspect to the research process involves the methodological question which deals with the issue of how the researcher can go about finding out whatever he/she believes can be known (Guba & Lincoln, 1994:108). The methodological approach that a researcher adopts is underpinned by the ontological and epistemological assumptions that one holds (Grix, 2002:179; Guba & Lincoln, 1994:108). Positivists have a deductive approach to study which involves testing theory or a hypothesis whereas the constructivists use an inductive approach which involves the use of participants' views to build themes and generate theory (Creswell & Plano Clark, 2011:41).

4.3.4 Axiological research proposition

Other schools of thought suggest a fourth perspective to the research process, axiology. Axiology refers to the role of values in the research process (Cohen et al., 2011:3; Creswell & Plano Clark, 2011:42). Positivists believe that values have no place in the research process which should remain unbiased by instituting checks and balances to eliminate biases. Constructivists on the other hand accept that the research process is influenced by the values of the researcher and therefore biased in a way. The researcher is however expected to talk about these biases and their effect on the interpretation of the findings. Since I did not adopt a purist position, to reduce the effect of biases I put in place checks in the form of quality control measures

and had a standardised questionnaire and I acknowledged how my values could have influenced my interpretation of the data that I collected from the interviews.

4.3.5 My philosophical assumptions on this study

As I embarked on this research journey, I brought with me some ontological, epistemological, methodological, and axiological assumptions which were informed by the nature of the research questions. For the research questions on the effectiveness of the teacher performance evaluation system to be answered exhaustively, I could not adopt a purist position but rather took a middle position. In seeking the reality of what exists in terms of effectiveness of the teacher evaluation system, I do not only believe that there is an objective reality out there, but that, there also exist multiple realities from the participants who are the educators. The objective reality was acquired through quantitative methodologies while the multiple realities were established by getting people's views and experiences on the effectiveness of the teacher performance evaluation system through qualitative methodologies.

My epistemological assumptions were shaped by the ontological assumptions in line with the assertions by Guba and Lincoln (1994:108), that, answers to the epistemological question are determined by the answers that are given to the ontological question. In seeking to understand the effectiveness of the teacher performance evaluation system in Kwekwe district, I needed both the objective and subjective knowledge for complete understanding. The objective knowledge was acquired by using quantitative data that sought to establish the patterns and trends of the attitude of the educators on the effectiveness of the evaluation system and the relationship between the variables derived from the RADPS framework. In order for me to understand the patterns and trends from the objective data, I had to get the deep and detailed explanations from interviewing key informants for their personal experiences and perspectives on the effectiveness of the evaluation system. This therefore made me adopt a middle position where assumptions from both the positivist and constructivist assumptions were brought onto the research landscape.

In terms of my methodological position, I adopted the deductive-inductive approach by testing the relationships between the variables in the RADPS conceptual framework and using the participants' views to generate themes that were later used to improve the original conceptual framework.

4.4 Research Paradigm

Bertram and Christiansen (2015:22) define a research paradigm as a worldview or belief by a community of researchers, of the nature of the world, what can be known about the world, and the methods that can be employed to learn about the world. The research paradigm of this study is, therefore, concerned with the way I as the researcher view things in the world in terms of my position on nature of reality (ontology), knowledge (epistemology), how to acquire the knowledge (methodology), and the role of values in the research process (axiology). These assumptions underpin my choice of paradigm for the study. The beliefs or assumptions we have about the world influence the way in which we research about the world (Betram & Christiansen, 2015:22). There are different paradigmatic positions that differ in their ontological, epistemological, methodological and axiological assumptions. Four paradigms, positivism, constructivism, transformative and pragmatism exist. I am however going to briefly outline two of these, positivism and constructivism to provide background to the rationale and logic for my adoption of pragmatism as the paradigm underpinning this study.

4.4.1 Positivism

The positivist paradigm is a world view hinged on beliefs of cause and effect (determinism) and the interrelation of selected variables (reductionism) (Creswell & Plano Clark, 2011:40). Positivists assume that reality is single and that a researcher should remain detached and impartial in the research process thereby making knowledge objective. They believe in value free research process that is based on empirical observation and measurement (Cohen et al., 2011:7-15; Creswell, 2003:6-8; Creswell & Plano Clark, 2011:40-42). This paradigmatic position is associated with the quantitative methods of research.

4.4.2 Constructivism

The constructivists' paradigm is opposite to the positivist position. This worldview contrary to the positivist position assumes that reality is multiple and that it is socially constructed and at the same time value-laden (Cohen et al., 2011:15-21; Creswell & Plano Clark, 2011:40-42). Constructivists believe that a researcher has to get immersed in the research process and will thus be influenced and in turn also influence the participants when getting data (Creswell, 2003:8-9). Constructivism is inductive in

nature and uses participants' views to generate themes and theory. Constructivism paradigm is associated with the qualitative research methods.

4.4.3 Pragmatism

Based on my ontological and epistemological positions which I presented earlier, neither positivism nor constructivism alone were appropriate for this study since my assumptions are not purist but rather, more of eccentric in nature thereby making pragmatism more appropriate in this case. Pragmatism arises out of actions, situations, and consequences and is not committed to any one system of philosophy and reality but rather draws liberally from both positivist and constructivist assumptions (Creswell, 2014:10-11). Pragmatism assumes that the world is both objective and socially constructed (Morgan, 2014:39). I believe that multiple realities exist in any given situation and therefore am not purist to anyone philosophical perspective. The pragmatic approach believes not only in the uniqueness of the knowledge of individuals but also on the role of society in the construction of this knowledge. In line with this assertion, I, therefore, sought to get the rich and deep individual experiences of the educators on the effectiveness of the teacher performance evaluation system to help me understand better the patterns and trends that were shown by the objective data.

Pragmatism is concerned more with the application of what works in relation to the presenting problems (Creswell, 2014:11). It does not focus on methods, but rather, on the research problem and uses all the possible approaches to understand the problem (Creswell, 2013:28; Feilzer, 2010:14). Pragmatism therefore assumes the mutual relevance of knowledge produced by both positivism and constructivism (Johnson, Onwuegbuzie & Turner, 2007:113; Morgan, 2014:40; Onwuegbuzie & Leech, 2005:377). The approach accepts both observable phenomena and subjective meanings as acceptable and balanced knowledge dependent upon the research questions (Teddlie & Tashakkori, 2009:88). Pragmatism emphasizes an inquiry process that aims to combine and exploit the different strengths of both qualitative and quantitative methods (Morgan, 2014:41). This position allowed me in the process of inquiry to use both the observable patterns and trends elicited from the quantitative data and the subjective meanings from the qualitative data in drawing conclusions on the effectiveness of the teacher performance evaluation system. The use of the two forms of data in a single research helped to provide me with a more comprehensive

and complete picture of how effective the teacher performance evaluation appraisal process is in Kwekwe district. The statistics generated from the quantitative strand of the study helped to give breadth to the study by allowing generalisation of findings over the population while the qualitative strand gave depth to the study by offering rich and deeply personal experiences on the effectiveness of the appraisal process. The use of both observable trends and patterns and subjective meanings in this study made it both scientifically and practically relevant. As a study which is evaluative in nature, the production of statistical data and rich textual data of personal perspectives and opinions on the study phenomenon expanded my understanding and enhanced the judgment process on the effectiveness of the teacher performance evaluation system.

Pragmatism thus acknowledges the value of both qualitative and quantitative methods in producing knowledge that can facilitate a better understanding of society and social phenomenon. It, therefore, provided for the use of both qualitative and quantitative data collection methods to gather information and to make inquiry into the complex phenomenon of teacher performance evaluation being studied and reported herein.

Pragmatist ontology aims at actions and change Goldkuhl (2012:2), and this study, similarly, aimed to inform action in terms of enhancement of teacher performance evaluation effectiveness through its recommendations, which is hoped will, in turn, translate to better teacher accountability and professional development, and, positive learning outcomes. This study intended to bring out the importance of creating evaluation systems that can accurately, fairly and objectively measure performance while highlighting the need for the system to be consequential. This can hopefully influence a relook at the system and in turn, bring positive change if some of the recommendations are adopted.

The knowledge character embodied in pragmatism is not restricted to explanations as in positivism or to understanding as in interpretivism. Pragmatism has other forms of knowledge such as prescriptive knowledge which provides guidelines, the normative knowledge that exhibits values and prospective knowledge that suggests possibilities. This enhanced its suitability for this research. The study intended to bring out the values of the different stakeholders in relation to the evaluation of teacher performance and at the same time assist to produce guidelines on how to enhance the teacher

performance evaluation system. This study, similar to the assertions by Goldkuhl (2012:9), was not limited to bringing local improvement but extended to be instrumental in creating knowledge that may be useful for local as well as general practices. The major aim of the study was thus, to create knowledge that can be used to improve the general practice of teaching and learning.

4.5 The Research Methodology

Before embarking on a research journey, a researcher has to decide on the plans and procedures about how the research will be conducted. The plans and procedures of conducting a research involve the philosophical assumptions that a researcher holds, the procedures of conducting the inquiry, the data collection methods, analysis and interpretation (Creswell, 2014:3). These plans and procedures collectively refer to what is called the research methodological approach.

When conducting research, there are three possible approaches that can be used namely, qualitative, quantitative and mixed method approaches. It is prudent that a researcher is cognisant of the alternative research approaches to make the relevant and appropriate decisions about the approach to use before commencing on a research journey (Yilmaz, 2013:311). I therefore decided to outline the possible approaches before zeroing in on the approach that I used for this study.

4.5.1 Quantitative approach

The other approach to inquiry is the quantitative approach. The quantitative approach is defined as an approach that uses numerical data that is analysed statistically to explain phenomena (Yilmaz, 2013:312). Creswell (2014:155) adds that the quantitative approach produces numeric descriptions of trends, attitude, or opinions of a population from a sample of the population and these are generalised to the population. The quantitative approach is informed by the objectivist epistemology and so it tries to explain social behaviour by measuring what it assumes to be static knowledge (Yilmaz, 2013:312). This approach explores relationships among two or more variables that are identified as characterising the observed phenomenon (Leedy & Ormrod, 2010:182). The quantitative approach assumes that social and psychological phenomena have an objective reality that is not influenced by the subjects being studied so the researcher is expected to detach him/herself from the study and its subjects to be able to study the reality objectively.

4.5.2 Qualitative approach

The qualitative approach is defined as an approach that is emergent, inductive, and naturalistic to the study of people, cases, and phenomena to facilitate the revelations of the meaning that people connect to their experiences of the world through descriptive means. Crowe et al. (2015:616) proffer that a qualitative research approach is an approach that is oriented towards understanding meanings and experiences of informants. The approach assumes that there is no absolute reality but that reality is socially constructed which therefore makes it subjective and dynamic. It tries to establish how social experience is created and how meaning is ascribed to the social experience (Yilmaz, 2013:312). This approach to research provides new insights and knowledge in areas that are complex and poorly understood like teacher performance evaluation. Leedy and Ormrod (2010:135) add that this approach to research focuses on phenomenon in its natural setting, studies phenomenon in its complexity by considering the multidimensionality of issues and also places value in the researcher's ability to interpret and make sense of what he or she sees as critical to allow understanding of a phenomenon thereby placing the researcher as an instrument to research.

4.5.3 Mixed methods research approach

Due to the nature of the research questions, neither the qualitative nor the quantitative approach alone could adequately address the research problem. Therefore, I adopted the mixed method research approach to address this complicated and compounded issue. The selection of the mixed method research approach was guided and influenced by the research problem and research questions (Bertram & Christiansen, 2015:41; Creswell, 2014:20). I was guided by the philosophy that the research question determines the research style and data collection methods used in what is referred to as the 'fitness for purpose' approach where the methods of data collection have to match the type of data that is required (Bertram & Christiansen, 2015:41). This is supported by Leedy and Ormrod (2014:269) who posit that a research design should be driven by a research problem and its sub-problems as some call for either quantitative or qualitative data, and only some call for both types of data.

A number of definitions have been proffered by different researchers for mixed method research approach. Johnson et al. (2007) addressed the issue of the definition by

analysing different definitions by experts in the field of mixed method research. They scrutinised nineteen different definitions and noticed that they are some similarities and differences among them. Informed by the analysis of these numerous definitions, Johnson et al. (2007:123) coined a broad definition of mixed method research as follows:

Mixed method research is the type of research in which a researcher or a team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration.

This definition was found to be relevant and appropriate for use in this study since it covers more or less all the basic issues involved in mixed method research.

The complexity of judging the effectiveness of teacher performance evaluation system calls for the use of both the quantitative and qualitative approaches since neither of the two approaches on their own would adequately address the problem in a way that would provide a fuller picture and deeper understanding of the phenomenon. This is supported by the assertion that mixed method research is appropriate where neither quantitative nor qualitative approach alone can provide the best understanding of the problem being studied (Creswell, 2014:20; Creswell, 2015:2; Teddlie & Tashakkori, 2009:15;). The core characteristics of mixed method research as summed up by Creswell (2015:3) are that the approach involves the collection and analysis of both the quantitative and qualitative data in response to the research questions thereby making meticulous and thorough use of both approaches. Creswell further points out that mixed method research uses a specific design, in this case the parallel convergent design to interpret and integrate the quantitative and qualitative data.

Based on the characteristic features of mixed method research as described by Creswell (2015:3), I believe this approach would provide an exhaustive way to tackle the research questions for this study. The assumption is that combining the statistical trends generated from the quantitative data with stories and personal experiences from the qualitative data helps to provide a better understanding of the research problem as compared to what either one of the two forms of data would do on their own (Creswell, 2015:2; Freshwater & Cahill, 2014:35). In judging the effectiveness of the teacher performance evaluation system, the statistical trends of the educators'

attitudes and views were not adequate enough on their own without the explanations and narratives generated from the qualitative data to explain them. The integration helped to enhance the meaning of the statistics thereby creating a better, deeper and richer understanding of the problem at hand.

In choosing to use the mixed method research approach, a number of factors were considered that point to the advantage of using the approach. The first advantage that I considered was the ability of the mixed method approach to enhance the validity of the study. The use of multiple methods, perspectives, and data sets to examine the phenomenon in question helped to enhance the validity of the study through the triangulation of methods and data sources (Dawson, 2002:20; Freshwater & Cahill, 2014:36; Teddlie & Tashakkori, 2009:13). The use of questionnaires and interview schedules as data collection instruments coupled with multiple data sources that included teachers, heads of departments and heads of schools facilitated triangulation of data collection methods and sources. Different and complementary data on the research topic was collected so as to mitigate bias introduced by a mono-method approach since the mixed method approach allowed the cancelling out of methodical biases of one approach and at the same time build on the strengths of both approaches (Creswell, 2003:23; Creswell, 2015:15; Leedy & Ormrod, 2014:269; Johnson & Onwuegbuzie, 2004:21; Saunders et al., 2009:154). Greene (2008:17) supports this assertion in her argument that the mixing of methods allows us to get results that are closer to the truth due to the offsetting of biases of error and perspective.

The second reason why I used the mixed method approach is the assumption or notion that collection, analysis and interpretation of both quantitative and qualitative data would provide a more comprehensive and complete understanding of the effectiveness of the teacher performance evaluation system that neither qualitative nor quantitative approach alone could provide (Creswell, 2015:15; Greene, 2008:24; Johnson & Onwuegbuzie, 2004:21; Leedy & Ormrod, 2014:269). The narratives from the interviews helped to explain the numbers and patterns that had been generated through the quantitative approach by explaining and revealing the underlying meanings thereby clarifying the meanings of the numbers and patterns. The quantitative aspect of the study allowed the measuring of the views of the educators

on the effectiveness of the teacher performance evaluation process and gave precision to the study whereas the qualitative strand provided in-depth experience from an individual perspective thereby giving depth to the study. The quantitative strand also provided statistical power and generalisability whereas the qualitative strand provided meaning, context, and depth (Creswell 2015:15). The use of both the qualitative and quantitative data to inform the conclusions further enhanced the credibility and trustworthiness of the conclusions. Corroboration and convergence from the different approaches provided stronger evidence (Johnson & Onwuegbuzie (2004:21). This was especially important in this case as the findings and conclusions of the study were meant to help to inform policy so their validity and trustworthiness were critical to avoid misleading decision-makers.

4.6 The Mixed Method Research Designs

Mixed method research as a research approach has different possible designs that can be used. The designs differ in the purpose for which they can be used. The design that I chose to use in conducting the mixed method research was determined by the motivation that I had for mixing the research approaches. This is supported by Ivankova et al. (2016:315) who proffer that the choice of mixed method design is informed by the purpose of the study. There are basically three main designs that can be used in mixed method designs. These are the *convergent*, *sequential* and *embedded* designs.

The designs are distinguished by timing, weighting, and mixing of the quantitative and qualitative elements (Plano-Clark, Huddleston-Casas, Churchill, O'Neil Green, Garrett, 2008:1555-1556). The timing refers to when the data is collected, analysed and interpreted. The first option for timing is concurrent where the activities for the two approaches are conducted at the same time or in a single phase. The second option is sequential which involves the activities for the two approaches being conducted in a definite sequence of different phases. Furthermore, weighting in the mixed method refers to the relative importance of the quantitative and qualitative methods in addressing the study purpose. Weighting can be equal or unequal in favour of one approach. The mixing aspect refers to how the quantitative data and results are related to the qualitative data and results with the aim of enriching understanding of what is learned from one strand of the study. Mixing can be through *merging*, *connection* or *embedding*. Merging involves comparing results to see if they corroborate with each

other and can be done at analysis phase through transformation by quantifying qualitative data to facilitate relating of the two data sets or at interpretation and discussion. Studies that are connected have one strand building upon the findings of the other study. Embedded mixing occurs at design level by using one type of data within the context of a design based on the other data type.

The three main designs also differ in the purposes that they serve. According to Morgan (2014:67), there are three main motivations for mixing and integrating methods which ultimately lead to a specific design. These motivations established by Morgan (2014) are:

- i. **Convergence-** Convergence is done when there is a need to produce greater confidence and certainty in the results by showing that the different methods with their different strengths lead to similar conclusions. This serves the purpose of ruling out methodical effects on the results that have been produced thereby corroborating, confirming and validating the results.
- ii. **Sequential contributions-** This motivation for mixing is used when the strengths of one method can serve as an input or as a follow-up method on another method. This helps to facilitate the achievement of more comprehensive detail than would be possible if only one method had been used.
- iii. **Additional coverage-** This is done to achieve additional coverage in terms of purpose which is accomplished when different methods provide different kinds of data for a study.

4.6.1 The Convergent parallel mixed method design

The purpose of this study was to determine the effectiveness of the teacher performance evaluation system in Kwekwe district. To make this determination, it was important to establish the general trends of the views, attitudes, and opinions of educators regarding the effectiveness of the teacher performance evaluation system. General causal relationships between variables within the RADPS conceptual framework had to be established but I felt that what came out of this inquiry alone could not justifiably be used as a basis of judgement of the system on its own since quantitative approach uses predetermined response categories where participants' different perspectives and experiences are fitted (Yilmaz, 2013:313). To rule out

methodical effects on the findings, the qualitative approach had to be used to corroborate and confirm the results (Morgan, 2014:70).

Simply establishing the educators' perspectives and attitudes on the system through the patterns and trends did not give me adequate confidence and certainty in the results. The quantitative data had to be confirmed with the qualitative data where the subjects were able to provide detailed rich and deep narratives of their experiences and perspectives without being limited to predetermined responses. This confirmation and corroboration were critical for the research since it is expected to have recommendations that can influence issues of policy and practice. This is supported by Morgan (2014: 72) who proffers that when conducting research that has a potential impact on policy issues, then, the convergent findings from the different methods provide more confidence in the results thereby justifying their use in making policy decisions. The need to have certainty and confidence in the results and ultimately the conclusions that were drawn led me to adopt the convergent parallel design. This design is also known as the concurrent or triangulation design.

The convergent parallel mixed method designs is a design that involves the collection and analysis of quantitative and qualitative data separately with the aim of integrating the results of the two databases through merging (Creswell, 2014:219; Creswell, 2015:36; Teddlie & Tashakkori, 2009:152). The purpose of the design is to compare the two types of data to allow a better understanding of the research problem (Ivankova et al., 2016:324) by gaining multiple pictures of the same problem from different perspectives (Creswell, 2015:37). The convergent parallel design assumes that the quantitative and qualitative strands collect different types of information (Creswell 2014:219) and this helps to answer both confirmatory and exploratory questions within the same study (Teddlie & Tashakkori, 2009:152).The Convergent parallel design is illustrated in Figure 4.1

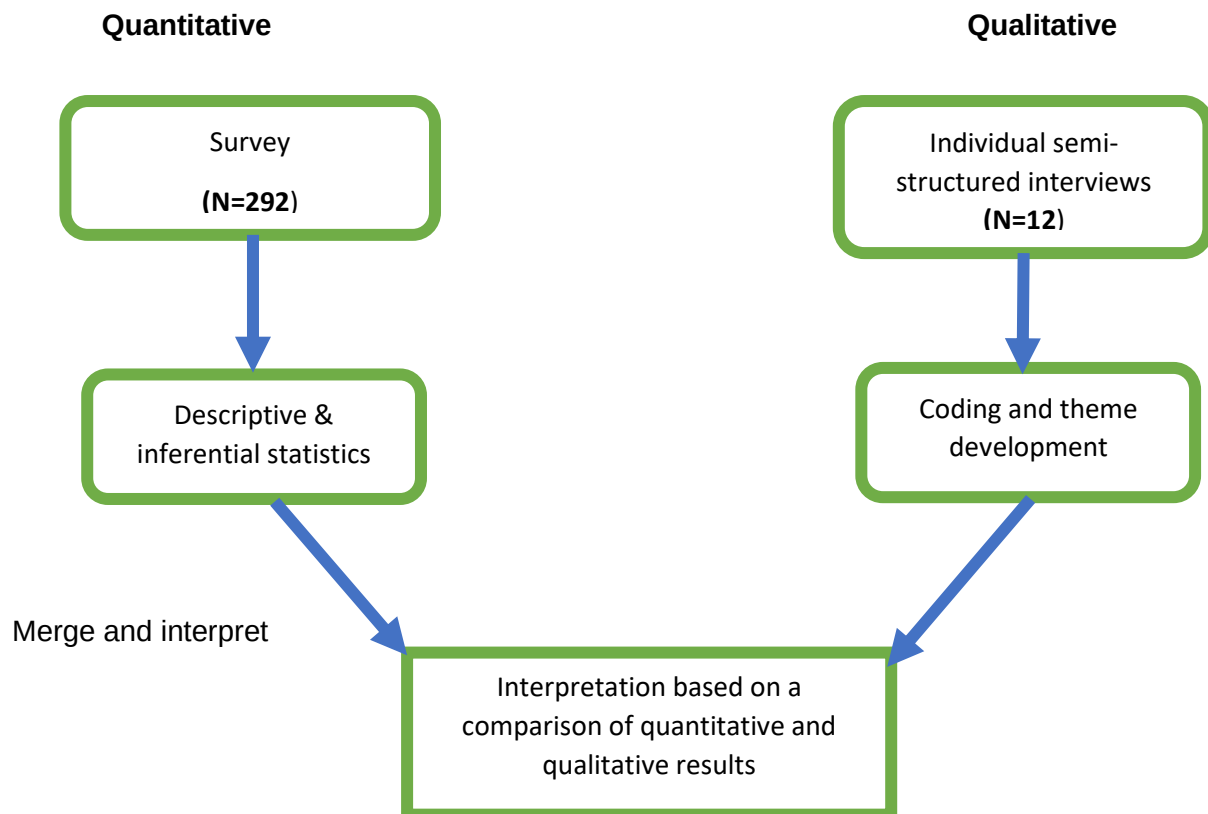


Figure 4.1: Convergent parallel mixed methods design procedures (Adapted from: Ivankova et al 2016:326)

The idea of seeking convergence of results was originated by Donald Campbell and his colleagues in the 1950s from the concept of triangulation. Morgan (2014:68) explains the idea of triangulation as being based on navigation and surveying procedures wherein if an object is seen from a particular point, one can draw a straight line from the point of observation in the direction of the object thereby providing a blinkered view which is somehow limited. The addition of a second sighting on the same object from different positions generates two intersecting lines with the object being located at the point of intersection, which is the tip of a triangle, hence the term, triangulation. The effect of this is to allow multiple perspectives or views on the object under scrutiny which helps to enhance the credibility of the view.

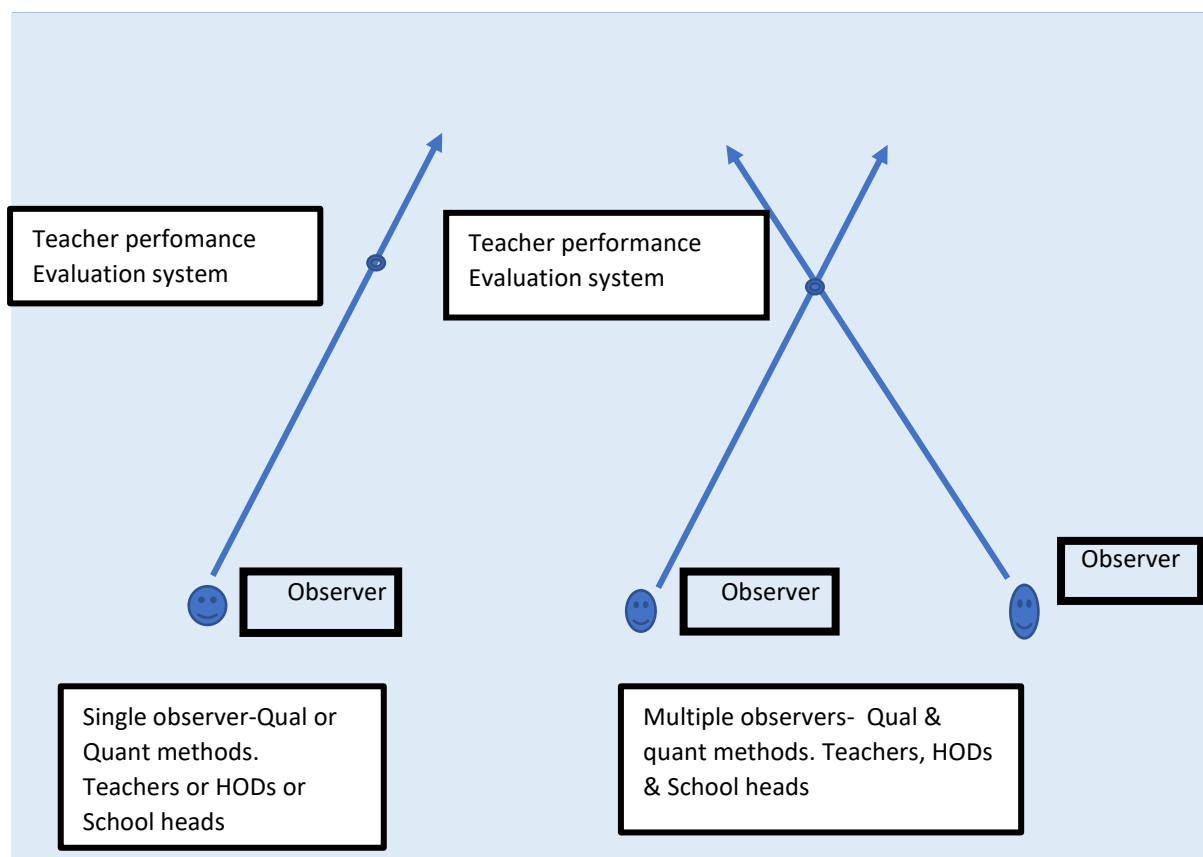


Figure 4.2: Triangulation as an analogy (adapted from: Morgan, 2014:69)

Morgan (2014:69) explains this analogy to mean that a single method limits one's ability to observe a phenomenon, but if the use of different methods produces similar observations, then the convergence can increase the confidence in the conclusions that are drawn. The independence of the two methods is therefore emphasised to facilitate comparison of results for convergence. In this case, as illustrated in the diagram and in the analogy above, if the use of both the quantitative and qualitative methods coupled with the use of multiple sources of information (teachers, heads of departments and school heads) produces similar or corroborating findings on effectiveness of the teacher performance evaluation system, then this will enhance the credibility of the study by increasing confidence in the conclusions drawn from the findings.

In this study, I collected both the qualitative and quantitative data concurrently and analysed the two databases separately. The data that I collected in the two strands were on the same variables or concepts as proffered by Creswell, (2014:222) and

Leedy and Ormrod, (2014:269, 270). The key features of this design are that the collection and analysis of data for the two strands of approaches are done independently or parallel; it is collected and analysed concurrently, and with respect to the same questions (Leedy & Ormrod, 2014:269). The results from the two databases were then merged by comparing them so as to determine the degree to which they agreed or disagreed with each other (Ivankova et al., 2016:319).

Merging is a process that involves bringing the two databases together and this can be done using different approaches that include data transformation, joint displays, and side-by-side comparison. Data transformation involves the conversion of one of the databases into the other form to facilitate easy comparison (Creswell, 2015:36). I did not use this approach to merging because although the approach allows qualitative data to be quantified to allow direct comparison with the quantitative data set, it however, sacrifices a lot of information that is attainable only through a full-scale qualitative research. This loss of some qualitative data by quantifying would in my view, in turn, partly defeat the purpose of mixing. Joint displays involve the merging of the two forms of data in either a table or a graph by jointly displaying both forms of data in one visual display for easier comparison. I, however, opted to use the side-by-side comparison. I did the merging by presenting and discussing the quantitative statistical results first followed by the qualitative findings before comparing them to either confirm or disconfirm the quantitative results. A good example of a study that used the side by side discussion approach to merging is one by Classen, Lopez, Winter, Awadzi, Ferree, and Garvan (2007:682:691). This study was on population-based health promotion perspective for older driver safety. The study discusses quantitative findings first and then the qualitative findings before comparing or integrating the two.

The convergent parallel design had the advantage of allowing me to collect the data for the two strands simultaneously or concurrently which meant that I spent less time in collecting the data than I would have spent if I had used the sequential design where the data collection is done in phases (Ivankova et al., 2016:319). Teddlie and Tashakkori (2009:153) acknowledge that convergent parallel designs are powerful but are however challenging to conduct due to the complexity of having to run the different strands at the same time. They, however, add that numerous dissertations with this

design have been conducted by solo researchers. I was able to mitigate this challenge by negotiating to administer the questionnaires in the mornings at the schools and then conducting the interviews in the afternoons.

4.6.3 Examples of convergent parallel mixed method studies

A popular study that used the convergent parallel design is the World Bank Guatemala Poverty Assessment conducted by Rao & Woolcock, 2003:173). In this study, two separate teams collected qualitative and quantitative data separately and mixing occurred at inference stage. According to the researchers, the different data sources helped to provide a more accurate picture of the spatial and demographic diversity of the poor as well as the context within which poverty was experienced by the different ethnic groups. Information on the local mechanisms that excluded them from participation in the main economic activities and the barriers that hindered their interests and aspirations was also identified as a result of the different data sources. This study, however, differed from my study in the sense that instead of using the team approach at data collection stage, I used the solo approach and was able to get a more accurate picture from the insights provided by the different datasets.

Another example of a study that used the convergent parallel design was conducted by Lopez and Tashakkori (2006:125-142). The study compared the effect of two different bilingual education programmes on attitude and academic achievement of Grade 5 learners. The quantitative strand used a Likert-type scale to measure the self-perceptions and self-beliefs related to bilingualism, administered standardised tests in the different academic subjects and linguistic competence in English and Spanish whereas the qualitative strand used the interviewing data collection technique. The two data sets were analysed separately and integration was done by comparing and contrasting the conclusions and by trying to produce a more comprehensive understanding of how the two programmes impacted on the learners.

4.6.4 Mixed method purpose statement

In line with the recommendations advanced in the steps of designing a mixed method study by Ivankova et al. (2016:330), a mixed method purpose statement should be presented after justifying the use of the mixed method approach. The mixed method purpose statement for this study outlines the intent of the study, the design used, the data collected and the rationale for the mixed method study (Creswell, 2015:68).

The purpose of this convergent parallel mixed method was to assess the effectiveness of the teacher performance evaluation system in Kwekwe district by obtaining quantitative results from a survey of 292 educators and concurrently getting deeper and richer details from purposively selected 15 informants to compare the qualitative and quantitative data sets so as to come out with well-validated conclusions about the complex issue of the effectiveness of teacher performance evaluation. In this study, the survey questionnaire was used to test the RADPS conceptual framework that assumes that readiness assessment, design, process, and significance influences the effectiveness of teacher performance evaluation in Kwekwe district using participants that were randomly selected from ten secondary schools. The interviews explored personal experiences, perspectives, and attitudes of the educators that had been purposively selected from selected secondary schools in Kwekwe district to explain the results from the quantitative strand on the effectiveness of the teacher performance evaluation system.

4.6.5 Research questions

Current literature recommends that a mixed method study should have specific research questions for the quantitative and qualitative components of the study and a mixed method research question (Ivankova et al., 2016:331; Creswell, 2015:60). In line with this recommendation, I therefore developed separate qualitative and quantitative questions. Although the questions were separate, they however focused on similar aspects of the phenomenon under study. The separate quantitative and qualitative research questions were subsidiary questions to the main research question as presented next:

4.6.5.1 Main Research Question

How effective is the teacher performance evaluation system in Kwekwe district of Zimbabwe?

4.6.5.2 Quantitative Research Questions

I opted to use research questions rather than hypotheses for the quantitative strand of this study. Creswell (2015:70) states that most mixed method studies use research questions rather than hypotheses. Creswell adds that in stating the quantitative research questions, it is important to identify the dependent and independent variables of the study. A variable is the aspect that is measured in research and an independent

variable is that which is not related to the purpose of the study which however may affect the outcome (Khotari & Garg, 2014:32). The components of the RADPS conceptual framework are *Readiness Assessment* which deals with the appropriateness of the timing of the introduction of the system, *Design* which refers to the planning aspect, *Process* which relates to the implementation aspect of the system, and *Significance* which involves the impact of the evaluation process in promoting accountability and professional development. The independent variables were thus timing, design process, implementation process and significance of process while the dependent variable is effectiveness of performance evaluation system. The quantitative questions are:

- What is the relationship between the timing of the introduction of the teacher performance evaluation in Kwekwe district and its effectiveness?
- How does the design process of the teacher performance evaluation system in Kwekwe district influence its effectiveness?
- How does the implementation of teacher performance evaluation system in Kwekwe district influence its effectiveness?
- What is the significance of the teacher performance evaluation system of Kwekwe district in promoting teacher accountability and professional development?

4.6.5.3 Qualitative research question

In this study, I used exploratory verbs in stating the qualitative research questions in line with the suggestion by Creswell (2015:71).

- How do educators in Kwekwe district view the timing of the introduction of the performance appraisal system?
- What are the educators' views on the effect of the design process of the performance evaluation system in Kwekwe district?
- What are the perceptions of educators on the implementation of teacher performance evaluation in Kwekwe district?
- How does the teacher performance evaluation system in Kwekwe district contribute to accountability and professional development?

4.6.5.4 The Hybrid or Integration question

Literature argues that a mixed method study should have a hybrid or integration question that clarifies the rationale for using both the quantitative and qualitative research strands (Teddle & Tashakkori, 2009:133; Creswell, 2014:149). The hybrid question provides additional information that is generated by having a mixed method design. Since the design of this study is a parallel convergent design, the mixed method question presents the purpose that is achieved by using the design, which is to compare or determine the extent of convergence or corroboration of the two sets of results from the quantitative and qualitative strand.

The hybrid or integration question for this study is, *‘To what extent do the qualitative results on the effectiveness of teacher performance evaluation system confirm and corroborate the outcome of the quantitative survey instrument on the views and attitudes of teachers on the effectiveness of the performance evaluation system in Kwekwe district’.*

4.7 Data Collection methods

My study was a mixed method study investigating the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. The aim of the study was to find out the educators’ views, perceptions and attitudes on the effectiveness of the teacher performance evaluation system and to explore their individual experiences and views on the effectiveness of the performance evaluation system in Kwekwe district. After the district of interest had been identified, I had to establish the research population before selecting the research sample. The district was selected for convenience purpose since my home is located in that district.

4.7.1 Study population and sample

A research population is any group that is the subject of interest for a research (Goddard & Melville, 2001:34). It is the total number of people and or sites who could be included in the study. The study population constituted a total of 55 secondary schools and 985 teachers in Kwekwe district. The target study population included heads of schools, head of departments, and teachers from the secondary schools that are located in the district.

4.7.2 Selecting the research sample

Since it was impossible to collect and analyse data from the whole study population, sampling had to be done. Sampling was done to allow me to observe relatively fewer cases and yet gain an accurate picture of a much larger population (Babbie, 2014:218). Sampling involves the selection of a portion or segment of a whole which is representative of the whole (Onwuegbuzie & Collins, 2007:281). In this case, the district had 55 secondary schools and it was not feasible to go to all the 55 schools in the district and so a sample had to be selected from these schools. The district has a total of 985 teachers and again it was not possible for all the educators to take part in the study, so a sample of those who participated in the study had to be selected.

The sampling procedure is a critical component of the research process because it is the one that determines the quality of findings that are generated by the study which in turn informs the inferences that the researcher makes on the phenomenon being studied (Onwuegbuzie & Collins, 2007:281). This position is also supported by Teddlie and Yu (2007:97) who assert that sampling decisions are critical as a poor sample will inhibit clarity and adequacy of the inferences drawn from the findings that are generated by the sample. In making the sampling decisions, I was guided by the characteristic features of samples suggested by Collins et al. (2007:270) that samples should:

- i. Generate adequate data on the phenomenon being studied.
- ii. Assist the researcher to get data saturation, theoretical saturation and/ or informational redundancy.
- iii. Allow the researcher to make statistical and/or analytical generalisations.

4.7.3 The sampling design

This study employed mixed method sampling design since it used the mixed method research approach. The sampling design refers to the framework within which the sampling process occurs and this includes the types of sampling schemes and sample sizes used in the study (Collins et al., 2007:271; Onwuegbuzie & Collins, 2007:283). The first aspect of a sampling design which is the sampling scheme involves the specification of the strategies that are used in selecting the units of analysis. The possible sampling strategies that define a sampling scheme are probabilistic or random and purposive sampling techniques. The second aspect of the sampling

design is the sample size which is the specification of the number of units that are selected for a study.

The sampling schema

The sampling design is key in determining the generalisations that can be justifiable from a study (Collins et al., 2007:273). Collins and others, however, acknowledge that in the case of a mixed method study like this one, making the appropriate generalisations from the study calls for greater caution due to the complexity of the design. The mixed method sampling design can be categorised according to two factors that are the time orientation of the components (whether the qualitative and quantitative samples of the study are selected concurrently or sequentially) and the relationship between the qualitative and quantitative samples (which can be identical, parallel, nested or multilevel). For this study, the concurrent parallel sampling design was used. The samples for the two strands of the study were selected simultaneously and they had a parallel relationship meaning that the samples for the qualitative and quantitative strands were different but drawn from the same underlying population (Collins et al., 2007:276-277).

The mixed sampling design was adopted for this study due to its favourable features as outlined by Teddlie and Tashakkori (2009:181). The design generated samples that did justice in addressing the research questions since they provided the study with both breadth and depth of information from the quantitative and qualitative samples respectively. Numeric data from the quantitative sample showed the patterns and trends on the way the teachers view the effectiveness of the teacher performance evaluation system while the narrative data from the qualitative sample helped to explain the patterns and figures from the quantitative sample through the rich and detailed explanations of the key informants. The quantitative sample also provided representativeness while the qualitative strand sought out the information rich cases. Since the qualitative sample sought information rich cases, its sampling decisions were therefore guided by expert judgement while the quantitative sample was decided on the basis of mathematical formulae to ensure that it achieved representativeness.

Teddlie and Yu (2007:85), state that mixed method sampling chooses procedures that focus on generating representative samples when addressing the quantitative strand

of the study and utilises sampling techniques that yield information-rich cases to address the qualitative strand of the study. Combining these techniques assisted the study to generate complementary databases that contained information that has both depth and breadth regarding the phenomenon that was being studied (Teddlie & Tashakkori, 2009:181; Teddlie & Yu, 2007:85). This approach to sampling allowed the use of multiple samples that varied in size depending on the research strand and question. In deciding the sampling design, I had to ensure that the design is realistic, efficient, practical, and ethical which is in line with the suggestions by Onwuegbuzie and Collins (2007:305) and Teddlie and Yu (2007:97). I had to ensure that the quantitative sample was representative enough and that the qualitative sample had people who could provide in-depth information on the effectiveness of the teacher performance evaluation system.

The Sample size

The sample size refers to the actual number of units that are selected for the study. Decisions made on sample sizes are equally important as those made for the sampling scheme since the sample size determines the extent to which a researcher can make statistical and or analytical generalisations of results (Onwuegbuzie & Collins, 2007:287). It is argued that the size of the sample should neither be too large nor too small, but rather, optimum so that it is efficient, representative, reliable and flexible (Kothari & Garg, 2014:53). The decisions on sample size for the quantitative strand of the study was guided by the type of statistical analyses planned, the accuracy of the results that were required and the size of the population (Cohen et al., 2011:144-145; Maree & Pietersen, 2016:199). The quantitative sample size had to be large since inferential statistics had to be calculated. I also had to consider the error margins that I was prepared to tolerate. Error margins are expressed in two ways, confidence level and confidence interval. The confidence level is the extent of how sure we can be that the responses lie within a given variation range, for example, 95 percent of the time and the confidence interval is the degree of the variation that one wishes to ensure, for example, 5% (Cohen et al., 2011:145). The confidence level and the confidence interval levels for this study were set at 95% and +/- 5% respectively. I however extended my sample size to cater for incidences where respondents failed to return the questionnaire or returned incomplete or spoiled questionnaires.

In coming up with a sample size, I was guided by the recommendations by Onwuegbuzie and Collins (2007:288) on sample sizes. The recommendations advanced by the pair are that a sample should have a statistical power of at least .80 and 5 % level of significance. They suggest a minimum of 64 participants for correlational design with a one-tailed hypothesis, 3-5 Participants for case study and at least 12 participants for interviews. These figures were the basis for my sample sizes for the different strands of the study considering the resources that were at my disposal in terms of money and time.

Two principles guided me in deciding the sample size for the qualitative strand: saturation and information power. Saturation refers to a point when all ranges of ideas have been heard and no new information is being generated (Cohen et al., 2011:601). Information power shows that the more the information that the sample holds that is relevant for the study, then the lower the number of participants that is needed (Malterud, Siersma & Guassora, 2016:1754). In this case, the participants had high information power since they are the direct implementers of the teacher performance evaluation system.

4.7.3.1 The quantitative sample

The purpose of a quantitative sample is to allow for generalisation from the sample to a population so that inferences can be made about some perspective, attitude, and behaviour of this population (Creswell, 2014:203). The quantitative sample of the study was selected using the probability sampling technique. Probability sampling is premised on two principles, randomness and probability theory which allow for accurate generalisation to the population (Maree & Pietersen, 2016:192). This type of sampling does not make it possible to predict the sampling units that will be selected in the sample since all possible people in the population have an equal chance of being selected. The probability sampling techniques are sampling strategies that allow each sampling unit in a specified population an equal opportunity of being included in the population. There are different types of probability sampling techniques that can be used in selecting a quantitative sample. Out of the four possible techniques that could be used, which are simple random sampling, systematic sampling, stratified sampling, and cluster sampling, this study used the simple random sampling technique.

A simple random sampling technique refers to a technique which allows each member of the population of interest to have an equal chance or nonzero probability of being selected each time a unit is drawn for inclusion in the sample (Crano & Brewer, 2002:173; Creswell, 2014:204; Onwuegbuzie & Collins, 2007:283; Teddlie & Yu, 2007:79). The procedure for selecting a simple random sample on situations where the total population is not too large involves entering all names individually on slips of paper of the same size, mixing the slips of papers well, selecting an individual to pick the slips individually from a container and not returning the selected slips to the container after being selected. Violating any of the conditions that are described will result in a sample that is not a true simple random sample (Crano & Brewer, 2002:173).

In order to be able to employ this sampling technique, I had to have an up-to-date and complete sample frame. Single stage sampling procedure was conducted since I had access to the names of the teachers in the population and therefore could sample them directly (Creswell, 2014:204). In this case, I used the clocking register at each of the schools which contained a list of all the teachers at the school. From the clocking register, I was able to draw a list of the teachers at the school and code them. I then wrote the codes for all the teachers and cut these into small cards with a code for each teacher. The code representing all teachers in the school were put in a container which was shaken so as to mix well the coded cards as a way of ensuring that each teacher had an equal opportunity to be included in the sample. The required number of participants from the school was picked and the procedure was done at all the selected schools to come up with the required number of participants. The total number of teachers selected at each school was proportional to the total sample size for the sample, meaning that, the schools that had a higher staff compliment also contributed a greater portion of the total sample size.

Since I knew the population parameter, which is the true population value (Crano & Brewer, 2002:171), I was able to use the sample size determination table to determine my sample size that was commensurate with the population parameter. With a total population size of 1000, confidence level of 95% and confidence interval of 5%, the sample size was supposed to be 278 in line with the sample determination table (Cohen et al., 2011:147; Krejcie & Morgan, 1970:608). My population total was 985

and I had a sample size of 292 which is above the figure for the recommended sample size with the selected confidence level and interval of 95% and 5% respectively. I had a higher sample size because I had given out 310 questionnaires as a measure to counter the possibility of a high non-response rate. The sample size was in fact 30, 6% of the population and I believe that this sample size was representative enough to allow generalisation of the findings to the population.

4.7.3.2 Qualitative sample

Qualitative samples are focused on the purposeful selection of a sample of participants who can best help a researcher to understand the phenomenon being studied (Creswell, 2015:76) by yielding information-rich cases (Teddlie & Yu, 2007:85), rather than for generalisation, as is the case with the quantitative sample. Patton (2002:273) alludes to this by asserting that, while the intent of quantitative samples is generalisation from sample to a population, the qualitative sample focuses on the purposeful selection of information-rich cases that allow the researcher to learn a great deal about the central phenomena thereby facilitating better insights and in-depth understanding rather than making just empirical generalisations. Qualitative samples are generated using the non-probability or purposeful sampling techniques. Purposive sampling for this study involved the selection of cases based on their typicality or possession of defined characteristics done on the basis of my expert judgment (Cohen et al., 2011:156).

Since purposive selection is 'the practice of collecting detailed information from carefully chosen sources', there was a need to come up with defining criteria for the cases that were eligible to serve as data sources (Morgan, 2014:127). The sampling frame for the study constituted the secondary schools and the educators.

Defining Criteria

The eligibility criteria for the schools used was that the schools should be a government high school, located in either a rural or urban location with a staff complement of at least 20 teachers. The Heads of schools had to be substantive heads which is an appointment commensurate with experience. Heads of departments had to be substantive and the teachers had to have been in service by 2005 when the performance evaluation system that is in use was adopted. The schools had to be

government schools since the teacher performance system being evaluated is originated by the Public Service Commission which is the employer of the government workers. The rationale for opting for substantively appointed Heads of schools and Heads of departments was to ensure that I would get educators who have served for at least ten years and therefore likely to have been in the system when the teacher performance evaluation system being studied was adopted and therefore likely to be rich sources of information on the phenomenon being studied.

Stratified purposive sampling

Stratified purposeful sampling technique was used to select the schools involved in the study. Stratified sampling involves the researcher identifying subgroups or strata in the population so that each unit belongs to a single stratum (Teddlie & Tashakkori, 2009:171). I used stratified purposeful sampling as a way of ensuring that certain schools varying on predefined criteria were included in the study (Sandelowski, 2000:250). In this study, the defining criteria for the strata were geographical location of the schools. Government secondary schools in the district were therefore put into two subgroups, rural and urban schools and each of the stratum was represented in the sample. The urban schools are those located in an area under the jurisdiction of a municipality or town council while the rural schools are located in an area under the jurisdiction of a rural district council. A school that is classified as rural in this study is one that is located at least 45km from a town centre and is under the jurisdiction of a rural district council. Five secondary schools were purposively selected from each of the two strata to give a total of ten secondary schools for the survey. From these ten schools, two rural and two urban were selected for the interviews.

Educators meeting the defining criteria were selected from each of the selected secondary schools to give a total of 4 substantive Heads, 4 substantive heads of departments and 4 teachers. The total number of informants in the sample was 15 educators, and, by then, data saturation had been reached as no new codes were coming up. This is in line with the findings by Guest et al. (2006:74) that by the time they reached the twelfth interview, data saturation had been reached and no new codes were coming up. Literature acknowledges that there is really no specific figure that is agreed for qualitative samples but rather that, the principle of saturation or redundancy should be the guiding principle on this (Creswell, 2015:77). Saturation or

redundancy to a point in the data collection process is when continued data collection from new participants does not add much to new codes and themes being developed (Creswell 2015:77). The general recommendation is that qualitative sample sizes should neither be too small to achieve data saturation or informational redundancy nor too large to make it difficult to undertake deep, case-oriented analysis (Onwuegbuzie & Collins, 2007:289). Guest et al. (2006:61), however, argue that a general guideline is needed to estimate the point at which saturation is likely to occur since protocols often require stating upfront the number of participants used in the study. In the case of this study, the qualitative sample was guided by the principles of saturation and information power. The purposive selection of the actual implementers of the teacher performance evaluation system in the form of teachers, heads of departments and heads of schools ensured that the sample had high information power. Guided by the information richness of the participants and the principle of saturation as per research findings that by the twelfth interview, no new information will be coming up (Guest et al., 2006:74; Onwuegbuzie & Collins, 2007:288), I settled for a sample size of 12 for this study.

Table 4.1 Summary of the sampling procedure and sample sizes

Research Strand	Sampling technique	Participants	Sample size
Quantitative	Stratified random	Teachers	292
Quantitative and Qualitative	Stratified purposive	Schools	10 4
Qualitative	Stratified purposive	Teachers	4
Qualitative	Stratified purposive	HODs	4
Qualitative	Stratified purposive	School heads	4

4.8 Instrumentation

The study adopted the parallel convergent mixed method research design and so it had two research strands, the quantitative and qualitative strand. Different data collection strategies were used for the two strands of the study.

4.8.1 Quantitative data collection

The quantitative strand of the study generated numerical data through a questionnaire.

4.8.1.1 The questionnaire

A questionnaire is an instrument that is used to collect survey information, is structured, can be administered in the absence of the researcher and is relatively easy to analyse (Wilson & McLean, 1994:3). The questionnaire (See appendix 5.1) was appropriate and efficient as a data gathering strategy for the target population since the population did not have challenges of literacy. The use of the questionnaire made it possible to collect data from a large number of people within a short period of time.

In the process of developing the questionnaire, I instituted the following measures in line with the recommendations by Maree and Jacques (2016:178-179):

- I provided simple and clear instructions to the respondent on how to answer the questions in the questionnaire.
- I paid special attention to the appearance of the questionnaire to ensure that it was user-friendly. I did this as a way of trying to minimise the non-response rate which can occur due to the questionnaire having poor outlook and being unappealing to the eye. I had to ensure that the font size that I used was legible enough and that the printing was neat and clear. The questionnaire had sixty one close-ended questions, which is within acceptable range since a maximum of 100 items is suggested.
- In terms of questionnaire sequence, the questionnaire started with easy questions on biographical information. I made an effort to keep questions on the same topic together to make it easier for the respondent. The questions were worded in simple language and in technical language that is familiar to the respondents. This was made possible by the fact that I did documentary review of available materials on evaluation prior to the development of the questionnaire and interview schedule so this assisted me to know the specific subject terms that the population is familiar with, for example, they use the term appraisal in place of evaluation.

The questionnaire contained close-ended questions that had set response categories on a Likert Scale. The Likert scale is an ordinal measure of the respondent's attitude on how he/she feels about something (Maree & Pietersen, 2016:186). The Likert Scale

had five response categories that included a neutral option. This scale allowed me to discover the strength of feeling or attitude on a given construct of the phenomenon being studied. The questions were arranged in topic-specific sets according to the research questions.

The questionnaire was administered after the participants had signed informed consent forms to partake in the study. I reassured the respondents on the issue of anonymity and confidentiality by assuring them that the answers that they gave in the questionnaire would be aggregated with those of others to produce aggregate results and statistical information (Statistics Canada, 2010:188). The questionnaires did not have any information that could lead to the identity of the respondent as a measure to uphold anonymity.

4.8.2 Qualitative data collection

A combination of data sources, teachers, heads of departments and school heads were used under the qualitative strand of the study. This was done in order to corroborate and confirm the data from the different data sources. The main strategy that was used to collect the qualitative data was the semi-structured interviews. The reflexive and observation journal was used to complement the interview. The interview was recorded by a digital voice recorder but there were some non-verbal communication which was important for the interpretation of the data that the digital recorder could not capture and so I had to note such communication in my reflective and observation journal. A reflective journal is a place where a researcher takes note of his/her biases, values, and interest which are likely to impact the qualitative research process (Creswell, 2003:182).

It is widely acknowledged and accepted that a researcher's biases, beliefs, misconceptions, and personal experiences have an impact on the research process or product (Cutcliffe, 2003:136). It is in response to this observation that Berger (2015:220) asserts that,

...researchers need to increasingly focus on self-knowledge and sensitivity; better understand the role of the self in the creation of knowledge; carefully self-monitor the impact of their biases, beliefs, and personal experiences

on their research; and maintain the balance between the personal and the universal.

It is in view of these acknowledgments that I decided, in an effort to try and enhance trustworthiness and credibility of this research, to note down in the reflective journal any personal biases, misconceptions, attitudes, and reactions that I experienced during the course of the research process so as to monitor their impact on the research process. This was important for enhancing the credibility and accuracy of the research findings.

4.8.2.1 Interviews

I conducted individual face to face interviews with teachers, Heads of Departments and Heads of schools from the six selected schools. An interview is defined as a two-way conversation where the interviewer asks an informant questions on a phenomenon of interest so as to learn about the participant's ideas, beliefs, views and opinions (Nieuwenhuis, 2016:92). Interviews are regarded as important data gathering strategies for qualitative studies as asserted by Nieuwenhuis (2016:93) that interviews can enable a researcher to elicit information that he/she would not collect in any other way as long as he/she manages to gain the trust of the interviewees and convince them on the importance of the phenomenon being studied. To establish trust in the informant, I had to build rapport first before delving into the serious aspects of the interview. Rapport involves establishing trust and respect for the interviewee and the information that he/she shares by creating a safe and comfortable environment to allow for the sharing of the interviewee's personal experiences and attitudes as they occurred (DiCicco-Bloom & Crabtree, 2006:316).

It is argued that the critical phase of rapport building is the initial phase known as the apprehension phase and this phase is characterised by the uncertainty that arises from the strangeness and unfamiliarity of the context (DiCicco-Bloom & Crabtree, 2006:316). The goal in this phase of the interview was to set the interviewee at ease so that they could talk. I established rapport and trust by reassuring the interviewee of confidentiality and anonymity in the use of the information that they provided. The interviewee was then asked the first interview question which was broad, open-ended and non-threatening.

In this study, the interviews provided a platform for me to view the world through the eyes of the participants from the rich and deep narrative information that was generated. I was able to seek further clarification with regards the informants' responses and on aspects of the phenomenon that was not clear.

In order to increase response comparability (Teddlie & Tashakkori, 2009:230) semi-structured interviews (See appendix 5.2) were used for the study. The semi-structured interviews were organised around a set of pre-set open-ended questions with the other questions coming out from the dialogue between me and the interviewee. This type of interview allowed me as the researcher to get specific information that was compared and contrasted with information that was gained from other interviews (Dawson, 2002:28). For comparability purposes, I used an interview schedule which had some basic questions that were asked to all interviewees which were followed by further probing and clarification. I had to be very attentive to the responses given by the informants so as to be able to exploit the use of elaboration and clarification probes to gain more information on the phenomenon of teacher performance evaluation which was the subject of interest for the study. Elaboration probing allowed me to get more information that provided a fuller picture of the phenomenon and I managed to do this by asking the informants to tell me more about their responses. Clarification probing involved paraphrasing the response given so as to confirm and verify what had been said.

I used a digital voice recorder to record the interview sessions. The audio records were later transferred to my computer and hard drive for backup before being transcribed to allow the analysis processing of the data. The interviews were recorded to allow verbatim capturing of the interviews and avoid any information getting lost by taking notes. Permission to record the interview was sought from the key informant before the actual interview when the informant consented to participate in the study. At the end of each interview, I thanked the informant courteously.

4.9 Data Analysis

The Parallel mixed data analysis that I conducted constituted of separate processes of quantitative and qualitative data analyses. Bertram and Christiansen (2014:115) define analysis as a process that involves closely and systematically examining a whole by separating the whole into smaller parts to facilitate its examination. In this

study, masses of data from the two strands of the research had to be disintegrated and systematically examined to allow me to understand the effectiveness of teacher performance evaluation system in Kwekwe district. The two independent analyses both helped to provide an understanding of the phenomenon being studied. As I was analysing the data, I was aware that the two independent analyses could lead to findings that are either divergent or convergent (Teddle & Tashakkori, 2009:266). The side-by-side comparison for merged data analysis was used to merge the data from the two strands of the study. Quantitative results were presented and discussed first followed by qualitative results before commenting on how the qualitative data either confirmed or disconfirmed the quantitative results (Creswell & Plano Clark, 2011:223). Figure 4.3 illustrates the data analyses procedures.

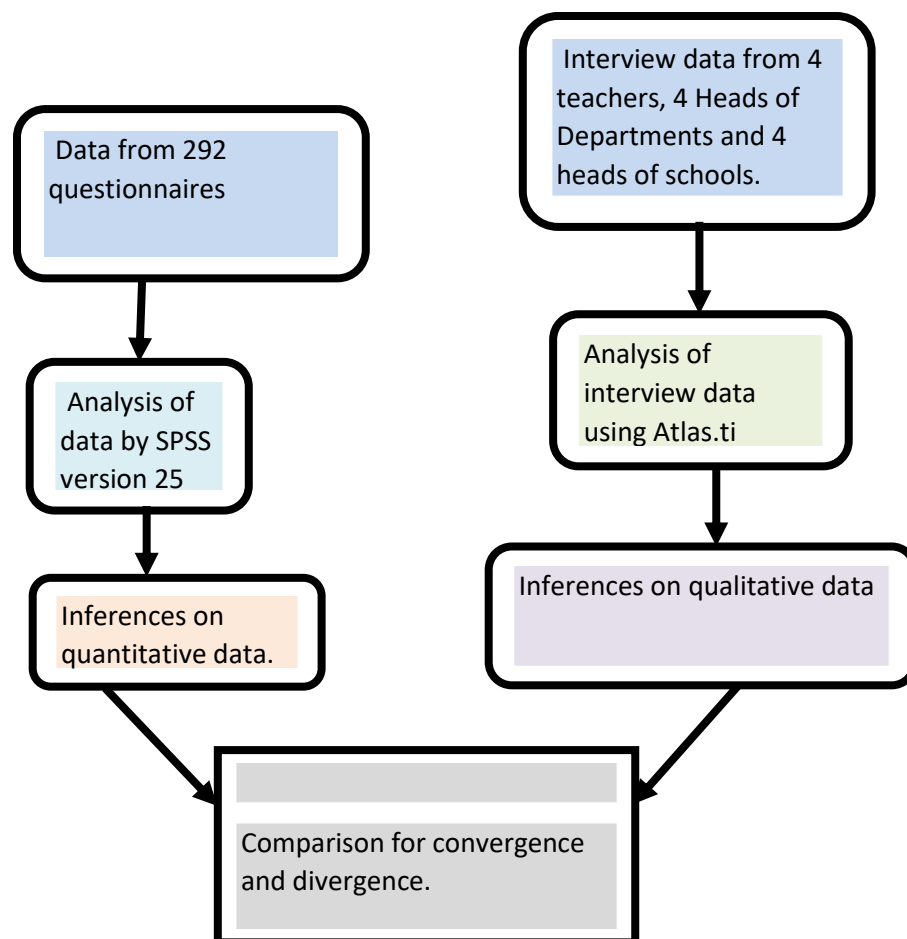


Figure 4.3: Parallel convergent/concurrent data analysis. Adapted from Driscoll et al.(2007:20)

4.9.1 Quantitative data analysis

Quantitative data was numeric data generated from the questionnaires that had been administered to the teachers. The process of analysing the quantitative data was

started by coding the questionnaires to facilitate tracking by numbering them 01 to 292. There was a need to identify the cases to facilitate easy tracking during the entering and verification of data in the statistical package. Identification numbers, therefore, had to be assigned to each questionnaire to allow for case identification as the questionnaires did not have any identifying information as a way of upholding the principle of anonymity.

After the questionnaires had been assigned identifiers, data was then entered from each of the questionnaires onto the Data Editor spreadsheet of SPSS. The Data Editor consists of the Data View and Variable View windows. Data entry was made on the Data View window, with the cases being entered on the rows and the variables on the columns. The variable specifications were entered on the Variable View window. The quantitative data for this study were analysed using descriptive and inferential statistics.

4.9.1.1 Descriptive statistics

Descriptive statistical methods were used to summarise numeric data generated from the questionnaires into tables and graphs to facilitate easy understanding of the data. Descriptive statistics make use of visual presentation of data in table or graphical form (Cohen et al., 2011:622). Frequency and percentage tables were used.

I started by conducting descriptive analysis of the demographic data of the respondents. This analysis allowed me to get an impression of the gender, age, level of education, teaching experience and geographical location of the schools of the respondents. The data was presented in a frequency and percentage table a summative description of the demographic features of the sample.

Descriptive analysis for each of the independent variables (*timing, design process, implementation process and significance of the evaluation process*) on the study was conducted. Frequency tables for each independent variable were presented to give the general attitude and opinion of the respondents on the different items for each variable. Frequency tables were also presented for the urban and rural teachers separately to get a picture of the opinions of the two groups of teachers.

4.9.1.2 Using the mean to rank the independent variables

The questionnaire used 5 point scale response categories. The mean weight for each of the independent variables was calculated so as to rank the variables in terms of their importance as exhibited by the responses of the participants. To calculate the mean weight for each independent variable, the responses were weighted with a 1 for strongly disagree up to a 5 for strongly agree. A mean weight was then calculated by adding up all the weighted responses for the participants for all the items for each variable and then the sum weight was divided by the number of items for the variable to get the mean weight which was then used to rank the variable relative to the other variables. This was done to standardise the independent variables before running regression.

4.9.1.3 Inferential statistics

Quantitative data was also analysed using inferential statistics. The inferential statistics allowed me to make inferences from the sample to the population (Cohen et al., 2011:641; Neuman, 2014:422). Regression analysis was conducted to test whether there is a relationship between the independent variables (*timing, design process, implementation and significance*) and the dependent variable (*effectiveness of teacher performance evaluation system*). The rationale behind conducting the inferential statistics was due to the nature of the research questions that required that the relationship between the independent and dependent variables of the RADPS conceptual framework be established. In establishing the effectiveness of the teacher performance evaluation system, I had to determine the relationship between the timing of introduction, design process, implementation process and significance of process (independent variables) to the effectiveness of the performance appraisal process.

Multiple regression analysis

Regression analysis is done to examine the relationships between the variables (Pietersen & Maree, 2016:269). Regression analysis models the relationship between the variables (Cohen et al., 2011:661). Regression analysis can be simple linear or multiple regression. In this study, multiple regression was used to predict and weigh the relationship between more than one independent or explanatory variables and one dependent or explained variable (Cohen et al., 2011:663; Pietersen & Maree,

2016:272). Multiple regression was thus used in this case to determine if the effectiveness of teacher performance evaluation system could be predicted based on timing, design process, implementation process, and significance of process. Different multiple regression tests namely R square, Analysis of Variance (ANOVA) and the Beta coefficient test were conducted in this analysis.

Determining how well the model fits the data

The R Square test (Goodness of fit test) was done to establish how well the regression model fits the data. The R square told me the extent of variance in the dependent variable that was explained by the independent variable in the calculation (Cohen et al., 2011:662). The adjusted R square which is more accurate was used. The adjusted R square is more accurate because it considers the number of independent variables and the fact that a sample was used rather than the whole population. To determine the goodness of fit of the regression model for my data, I was guided by Muijs (2004:165) who proffer that the value of an adjusted R square means:

< 0.1 means poor fit.

0.11-0.3 means modest fit.

0.31-0.5 means moderate fit.

> 0.5 means strong fit.

The ANOVA test

After the goodness of fit test, ANOVA test was conducted. The f-ratio in the ANOVA table was used to determine if the overall regression model was a good fit for the data. The test was done to establish the level of statistical significance of an independent variable to predict the dependent variable. A significant level of 0.000 meant that there was a statistically significant relationship between the independent variables (timing, design, process and significance) and the dependent variable (effectiveness of performance evaluation system) (Cohen et al., 2011:662).

The Beta Coefficient

Beta weight refers to the 'amount of standard deviation unit of change in the dependent variable for each standard deviation unit of change in the independent variable' (Cohen et al., 2011:663). The Beta weighting showed the degree of standard deviation units that were changed in the effectiveness of teacher performance evaluation system for each standard deviation unit of change in each of the independent variables (timing, design, process and significance).

4.9.2 Qualitative Data analysis

Data analysis involves bringing order, structure, and meaning to the mass of collected data and searches for relationships among categories of data (Marshall & Rossman, 1995:111). describe the data analysis process to constitute of mainly three activities which are data reduction, data display, and conclusion drawing and verification (Miles & Huberman, 1994:10-11; Miles, Huberman & Saldana, 2014:10). The first step of data reduction involves the sorting out of data into codes or categories and then looking for patterns between the categories. Data display is the organised, compressed assembly of information in the form of quotes and network views that permits a researcher to draw conclusions. The conclusion drawing and the verification aspect is the stage where the researcher notes the patterns in the data and develops possible explanations. It is important to note that the processes of data collection and analysis in qualitative research are interactive (Nieuwenhuis, 2016:114; Smit, 2002:66) so the provisional results of the study should be traced to the context of the data. It is with this in mind that before getting into the actual processing of the data, I started by giving a detailed description of the participants and an in-depth discussion of the context in which the study was conducted. My description included the number of participants involved, the selection criteria that was used, the job experience, and a detailed description of the context.

Qualitative data analysis can be done manually or using computer programs. Analysing data using computer software programs is referred to as Computer Assisted Qualitative Data Analysis (CAQDAS). The term was introduced by Fielding and Lee in 1991. There are different software programs that can analyse qualitative data and they include NVivo, MAXQDA, Hyper RESEARCH and ATLAS.ti. I opted to use ATLAS.ti 8 for the analysis of the qualitative data. I was motivated to use ATLAS.ti 8

rather than the manual analysis for a number of reasons. The first reason was the bulkiness of the amount of data that I had which would have been overwhelming for me to deal with effectively. The second reason was that ATLAS.ti 8 helped to speed up the coding process and made it more flexible (Hwang, 2008:521) thereby leaving me with more time to do the 'actual' analysis (Barry, 1998:2.10). The use of ATLAS.ti 8 allowed the analysis process to be more transparent and more replicable thereby increasing the credibility of the qualitative strand of my study (Hwang, 2008:521; Konopasek, 2007:294). I also opted to use ATLAS.ti 8 so as to exploit its ability to pick deviant cases with much ease as compared to the manual process (Barry, 1998:2.10).

The ATLAS.ti methodology of data analysis is based on four principles of Visualisation, Integration, Serendipity and Exploration (VISE) (Smit, 2002:73; Konopasek, 2007:288). The Visualisation principle refers to the direct supportive role of the software to the way humans think and plan while the principle of integration to the availability of the integrated whole project within the hermeneutic unit at all times. Serendipity principle involves the instinctive approach to data analysis where a researcher goes through the data and makes discoveries without a forceful approach. The exploration principle implies the use of an exploratory and discovery-oriented approach.

The analysis of data using ATLAS.ti occurs in two modes, textural and conceptual levels (Pandit, 1996:13). The textural level of analysis involves breaking primary documents into quotations, coding and memoing while the conceptual level analysis refers to the process of framework building which involves the linking of codes, concepts and categories to form theoretical networks (Pandit, 1996:13).

This analysis procedure was mainly informed by Konopasek (2007:284-293)'s guide on analysing qualitative data with ATLAS.ti. The process is presented in a step by step process although it is important to remember that qualitative analysis is iterative and non-linear (Smit, 2002:66).

i. Entering primary documents

The first step involved entering all the primary documents (PD) into the hermeneutic unit of ATLAS.ti. Twelve interview transcripts for different participant categories,

teachers, heads of departments and heads were entered. The hermeneutic unit is the data base for a particular project (Lewis, 2004:452).

ii. Defining quotations from primary documents

This step involved the breaking down of the primary documents into segments of data in the form of quotations. The cutting of primary documents into quotations assisted me to reduce the data to comprehensible levels and allowed me to focus on what I had initially found to be relevant. Quotations are basic units of analysis and their limited textural sizes and meaning make it easier to construct their meaning. Konopasek (2007:286) however also warns that the breaking down of primary documents into quotations makes the data more distant from each other as the quotations increase in numbers which therefore calls for greater effort to identify the connections. Identification of the connections between the data could require hours of careful reading and re-reading as I had to do.

iii. Reintegrating data by coding

Coding was done by assigning selected portions of the primary documents with one or more codes from a pre-set code list derived from the RADPS conceptual framework. Codes are the labels or names that indicate the kind of quotations in particular thematic group. I had to make sure that the codes were not just simple but also showed the content of the quotation clearly. The coding process gave the quotations more relevance and made them more meaningful. Codes are not fixed but are flexible since they can be moved, renamed, split or even linked to each other. Coding did not however stop with the priori codes but even extended to include the emerging codes since I was using a deductive-inductive approach because of the research method that I was using. The emerging codes were important since they were later utilised to improve the RADPS conceptual framework based on empirical data.

iv. Commenting and memoing

Comments and memos were attached to some of the analytic units like the quotations, links and networks. The comments differed depending on the analytic unit to which I had attached them and the purpose for which they were meant to serve, for example, some comments on the codes were used to describe and explain the labels that were not very descriptive while the comments on the links or quotations explained the interesting aspect of the analytic unit. Comments on the primary documents gave

details on the source of the primary document. The memos unlike the comments were more specific in the sense that they involved the integration of partial observations and informed the research report based on empirical data and the wider argument of the study. The commenting helped to preserve the ideas that came up as I was reading the data and constituted part of the interpretation process and provided room to infuse my own additions to the participants' account thereby transforming data to 'sociological data' in a gradual way.

v. Categorising data into families and seeing relevance

Comments were attached to the analytic units and links in order to justify and ground my line of argument. The codes were grouped into families and sub-families (axial coding) that were used as network nodes later in the modelling. ATLAS.ti enabled me to get a view of the nature and density of the links and connections in the data set.

vi. Reading data in a new way

The segments of data that were cut off from the primary documents told a new story when they were put together in thematically defined relationships (selective coding). I discovered that what could initially have appeared to be insignificant became noteworthy from the frequency by which it was mentioned by the different people. Space for new ideas leading to new textural additions in the form of comments, links, codings and ultimately new relevancies (serendipity principle) was developed. The gradual interpretation of the participants' voice resulted in new sociological reading that constituted the participants' voices and my voice through my construction of meaning. The analysis process led to the new accounts of the situation pertaining to the effectiveness of the teacher performance evaluation system in Kwekwe district resulting in new views, insights, and ultimately new understanding.

vii. Producing the report

Since the aim was to tell my story in a way that convinced the reader about the merit and validity of my analysis, I provided concise, logical and coherent extracts within and across the themes. I had to present sufficient evidence on each theme from the data by giving sufficient data extracts to show the prevalence of the theme. Network views were also presented. Network views are graphical presentations of quotations and codes showing relationships between the data codes (Lewis, 2004:453 & 458).

4.10 Quality Assurance Measures for the Study

Since I used the parallel convergent mixed method design, I had to aim at establishing quality assurance measures for both the quantitative and qualitative strands of the study. Morse, Barrett, Mayan, Olson, and Spiers (2002:15) proffer that the nature of knowledge in the quantitative (rationalistic) paradigm and the nature of knowledge in the qualitative (naturalistic) paradigm is different and therefore each strand of this study required its own specific criteria to address quality issues namely 'rigor' in quantitative part and 'trustworthiness' in the qualitative part. In support of this position, Golafshani (2003:601) posits that the quality of a study in each paradigm should be judged on the basis of the paradigm's own terms. It is against this background, therefore, that I present the quality assurance measures for each strand of the mixed method study separately.

This study, however, had a common quality assurance measure that was used for the two strands of the study which I present first before presenting the separate measures for the two strands. Piloting of the data collection instruments was done for both the quantitative strand and qualitative strand of the study. This is line with the recommendation by Burton, Brundrett, and Jones (2014:142) that it is usual to pre-test a research instrument with a group of people with similar characteristics to the sample prior to using it. Kimberlin and Winterstein (2008:2277), posit that pre-testing or pilot testing an instrument facilitates the identification of errors thereby giving room for the refinement of the instrument.

4.10.1 Piloting the questionnaire

The questionnaire was pre-tested on a group of teachers sampled from a school which was not among those selected for the actual study. Thirty participants were selected to answer the questionnaire. The respondents were told the objectives of the pilot exercise and were asked to comment on the length, structure, and wording of the questions as suggested by Dawson (2002:95). One purpose of piloting the questionnaire was to help me to identify issues such as ambiguities and inconsistencies in the questions. The other objective was to determine the relevance and clarity of the questions in the instrument and also to get comments on the length and outlook. The teachers were also asked to comment on the content validity of the instrument. The questionnaire was re-administered to the same teachers two weeks

after the first administration for the test-retest reliability test. The details of this test for reliability will be presented later on when I present the quality assurance measures for the quantitative strand of the study.

There were a few teachers who indicated that the questionnaire was rather too long. The indications from the pilot study indicated that on an average, a respondent would need approximately fifteen minutes to complete the questionnaire. In terms of length of a questionnaire, recommendations from literature are that a questionnaire can have up to 120 items and take up to 30 minutes if it is for adults (Maree & Pietersen, 2016:179). This questionnaire had 61 items and took an average of 15 minutes to complete which therefore made its length reasonable relative to the recommendations from literature. The respondents indicated that there were conflicting instructions for section 2. This was corrected to make sure that the instruction on this section given at the beginning of the questionnaire was similar to the one given at the beginning of the said section. The respondents however indicated that the structure of the questionnaire was appealing in terms of its presentation and legibility and that the wording was simple and clear. They also commented positively about the comprehensiveness of the instrument.

4.10.2 Piloting the interview schedule

The interview schedule was piloted with two school heads. The heads were purposively selected because of their experience and knowledge of the teacher performance evaluation being studied. The objectives of the pilot were to assess the relevance and clarity of the questions, length of the interview and flow of the questions. The piloting of the interview made it possible to plan for the actual study from the time taken at piloting. The piloting of the interview also assisted me to determine if the questions were clear and relevant. I was also able to tell from the pilot interview whether the questions were in an order that preserved the conversational and inquiry goals of the interview (Castillo-Montoya, 2016:822).

The interview schedule was reviewed after the pilot study using the comments and observations that had emanated from the pre-test. Reviews made to the interview schedule involved adding some questions in some thematic areas, changing the wording of some questions and removing certain questions that were seen as

repetition. The piloting of the data collection instruments helped to enhance the validity or appropriateness of the instruments.

4.10.3 Quantitative quality assurance measures

The research rigor of the quantitative strand of this study was ensured by establishing reliability and validity.

4.10.3.1 Reliability of quantitative research instrument

In this study, reliability is defined as the extent to which a measuring instrument is consistent and repeatable (Pietersen and Maree, 2016:238). Reliability is also defined as the degree to which a measuring instrument is free from measurement error (Scholtes et al., 2011:237). Reliability has a value between 0 and 1 with the value of 1 indicating perfect reliability and 0 indicating zero reliability. Literature proffers that a reliability of at least 0.70 indicates an acceptable reliability level (Scholtes et al., 2011:237). Cronbach (1951:297) posit that reliability of measurement should be a concern for any research that is based on measurement and since the quantitative strand of my study was concerned with the determination of the attitudes, opinions, and perspectives of the teachers on the effectiveness of teacher performance evaluation, it meant that I had to deal with reliability concerns.

Different types of reliability exist in research, but for this study, I focused on two types which are the test-retest reliability and the internal reliability.

Test-retest reliability

The test-retest reliability measure is concerned with the stability of the research instrument or its consistence over time and over similar samples (Cohen et al., 2011:200). For a research instrument to be regarded as reliable, it should yield similar data from similar respondents over time. I used the test-retest reliability measure in this study by administering the questionnaire to the same respondents on two different occasions during the piloting of the questionnaire. I compared the first set of scores to the second set of scores by calculating a correlation coefficient between the two set of scores. This test shows the stability of the score from the instrument and therefore is called a coefficient of stability (Cronbach, 1951:298). Since the reliability test evaluates reliability across different timing, the timing of the second administration was therefore critical. It is recommended that time between the two test administrations

should not be short to avoid the challenge of 'memory effect', where an artificially high reliability will occur as result of recollection. On the other hand, it should not be too long to allow external factors to distort the data (Cohen et al., 2011:200). The recommended time interval between the two tests administrations is 2-4 weeks (Scholtes et al., 2011:237). In this study, the second administration was done after 2 weeks from the first one. Reliability has a value between 0 and 1, with the value of 0 indicating no reliability while a value 1 indicates absence of measurement error or maximum reliability (Pietersen & Maree, 2016:238). The coefficient of stability for this questionnaire was 0.9 which indicates high reliability of the questionnaire.

Internal reliability

The internal reliability was used in the study to measure the degree of similarity or interrelatedness between the items that were measuring the same construct of the phenomenon (Pietersen & Maree, 2016:239; Scholtes et al., 2011:237). Items that are purported to measure the same construct should have a high degree of similarity which indicates the internal consistency or reliability of the instrument. The Cronbach's alpha coefficient measures internal consistency and it is based on inter-item correlations. Values greater than 0.70 show that the items are sufficiently correlated (Scholtes et al., 2011:237) while values lower than 0.60 are regarded as too low and therefore unacceptable (Pietersen & Maree, 2016:239). The value for Cronbach's alpha coefficient for this study was 0.95. This value of Cronbach's alpha co-efficient measure indicates high internal consistency or high degree of interrelatedness between the items measuring the same construct.

4.10.3.2 Validity of the questionnaire

Validity refers to the extent to which an instrument measures what it is supposed to measure (Pietersen & Maree, 2016:239). There are a number of conditions that potentially threaten validity. These are:

- Reliability of an instrument. If an instrument is not reliable then there is no way it can be valid.
- The tendency by some respondents to affirm to all questions, so to avert this, some questions have to be formulated in the negative like I did in my questionnaire.

- The tendency by respondents to respond in a manner that they think is socially desirable.
- Item bias on the basis of cultural differences or gender bias

(Pietersen & Maree, 2016:239).

There are a number of measures that I took in the study to ensure validity of the instrument. As is indicated in literature, for an instrument to be valid, it has to be reliable. I had to conduct the test-retest reliability test to ensure reliability of the questionnaire which in turn contributed to the enhancement of the validity of the instrument. Piloting the instrument was another measure that I took to ensure that the instrument is valid. Other people like the promoters, peers and educators were also consulted to validate the instrument. For this study, face validity, content validity, and construct validity were established.

Face Validity

Face validity refers to the extent to which an instrument appears to measure what it is supposed to measure (Pietersen & Maree, 2016:240). This type of validity is not measurable or quantifiable but is determined through expert judgment. Face validity for my questionnaire was determined by my promoters and other experts in the field that I consulted like educators. I also consulted my peers to determine face validity as they are familiar with the phenomenon as educators. Piloting the instrument was also another measure that was taken to ensure the validity of the instrument.

Content Validity

Content validity refers to the extent to which the items or content of an instrument adequately reflects the construct being measured (Scholtes et al., 2011:239). Content validity like face validity is determined by expert judgment. To ensure content validity for this study I observed the following suggestions by Scholtes et al. (2011:239):

- The items in the questionnaire were relevant and comprehensive
- The aim of the measurement was clear that it's evaluative.
- The content of the questionnaire was appropriate for the teachers.

To determine the content validity of my questionnaire I made use of my promoters as experts on teacher performance evaluation. Content validity of the instrument was also tested through the pilot study.

4.10.4 Qualitative quality assurance measures

While reliability and validity are key quality assurance measures in quantitative research, trustworthiness is considered the key quality assurance factor for qualitative studies and is measured through credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985:290). To establish data trustworthiness for my study, I employed a variety of measures that included triangulation, thick descriptions, member checking, reflective journal, debriefing sessions, and peer scrutiny (Nieuwenhuis, 2016:123-125; Onwuegbuzie & Johnson, 2006:126-129; Shenton, 2004:64-72; Teddlie & Tashakkori, 2009:295-297).

4.10.4.1 Credibility

Credibility is defined as the extent that the findings are congruent or in agreement with the reality on the ground (Nieuwenhuis, 2016:123; Shenton, 2004:64). Credibility is equated to the internal validity of quantitative research (Teddlie & Tashakkori, 2009:296). There are a number of measures that I used to ensure credibility for this study. In line with the suggestions by Nieuwenhuis (2016:123) and Shenton (2004:64), I used well-established research methods for my study. I did this by selecting a research design that would allow me to adequately address the research questions of my study. Well-defined purposive sampling was done taking care that rich-information cases were selected and that the necessary stakeholders were involved as key informants which led me to include teachers, heads of departments and heads of schools as the key informants for the study. This was meant to allow me to get a wide array of perspectives and views on the phenomenon from all key constituencies for credible findings to be made. The documentary review I did also assisted me to develop early familiarity with the organisational policies pertaining to the phenomenon in line with the suggestions by Shenton (2004:64). This, in turn, helped me to develop relevant instruments which were able to elicit the required data. The frequent debriefing sessions that I held with my promoters widened my vision and drew attention to flaws in the proposed procedures while opening up discussions for alternative approaches. The use of peers to scrutinise my work helped me to refine

my methods and strengthen my arguments for the different procedures of the research.

I also made use of triangulation by having a variety of sources to corroborate my data and findings. Data sources were triangulated by having informants representing different sectors of the educators in the form of teachers, heads of departments and heads of schools. These were expected to give comprehensive views thereby creating a rich picture of the perspectives, attitudes, and opinions on the teacher evaluation system being used as I had representatives of evaluatees and at the same time representatives of evaluators which allowed me to get experiences of people from different standpoints. Site triangulation was done by having schools from different geographical locations, rural and urban sites. This assisted me to determine local factor effect on the findings thus enhancing the credibility of the findings (Shenton, 2004:66).

To further enhance the credibility of this study, I made rich and thick descriptions of the research procedures and processes by providing rich and detailed descriptions of the setting, the participants, the methods and the analysis procedures (Nieuwenhuis, 2016:123; Onwuegbuzie & Johnson, 2006:127; Shenton, 2004:69). I also conducted member checks which Lincoln and Guba (1985:314) describe as the most crucial technique for establishing the credibility of a study. I gave the informants my observation notes made during the interviews where I noted the non-verbal communication of the informant and the transcription of the interview to verify and comment on their accuracy before analysing the data. In this study, member checks did not go beyond the observational notes and the transcriptions because, after results synthesis, informants could not be involved in member checking as this would have forced me to keep the results to be more descriptive in order to address the individual concerns which would have compromised my work. This is in line with the assertion by Morse et al (2002:16) that member checks may actually invalidate the work of the researcher and keep the level of the analysis inappropriately close to the data if taken thus far into the research process.

4.10.4.2 Transferability

Transferability refers to the extent to which the study can be applied to other situations (Shenton, 2004:70). Teddlie and Tashakkori (2009:296) equate transferability to the

external validity of quantitative studies. In order to enhance the transferability of my study, I provided thick descriptions of the contextual and procedural information to allow the reader to make connections between aspects and elements of the study that they could connect to their own experience or research (Nieuwenhuis, 2016:124). Providing vivid detail of the processes was meant to facilitate the reader to be in a position of making an informed decision on the applicability of the findings to other settings or similar context (Onwuegbuzie & Johnson, 2006:129). I provided sufficient detail on the phenomenon to allow the reader to properly understand it. There is generally controversy on the nature and extent of background information that has to be provided in the descriptions. In line with the suggestions by Shenton (2004:70), I provided the following detail in my descriptions:

- A number of the institutions taking part in the study and their geographical locations.
- The number of participants involved in the field work.
- The data collection methods used.
- The number and length of data collection sessions conducted.
- The total time period over which the data was collected.

It is important that the results of the qualitative strand of my study be understood within the context of the characteristic features of the sites and participants. Nieuwenhuis (2016:124) also adds that purposeful sampling enhances transferability if care is taken to ensure that participants represent the entire population relative to the phenomenon or context being studied. I did this by ensuring that I had representatives from key stakeholders such as teachers, heads of departments and heads of schools.

4.10.4.3 Dependability

Bertram and Christiansen (2014:190) define dependability as the ability of the researcher to account for the variations in the study, for example, between cases and comparing the study to previous studies in the same field and explaining the main differences. Teddlie and Tashakkori (2009:296) equate the dependability criterion to reliability in quantitative studies. Shenton (2004:71) proffers that dependability can be

enhanced by adequately reporting on the process within which the study was conducted in such a way as to allow another researcher to be able to repeat the work. Shenton (2004:71) adds that to allow the reader to understand the methods used, the report should have details on the following:

- The research design and its implementation, detailing what was planned and what was actually executed.
- The operational detail of data gathering including what was done in the field.
- Reflective appraisal of the study.

For this study, to implement the suggestions by Shenton, I kept a journal where I memoed my decisions through the research process. This was meant to allow others to follow my reasoning. Most of the memoing was done for the data collection and analysis processes. Birks, Chapman and Francis (2008:69) proffer that memoing helps to ensure that ideas that could be lost are retained, record a decision making trail of how decisions were made and conclusions reached throughout the research process, and allow the researcher to articulate, explore, contemplate and challenge their interpretation thereby engaging with the data more deeply. Memoing helped to enhance my research experience and ultimately led to my producing a product that possessed greater depth, quality, and elements of self as well as shared meaning that is characteristic of qualitative research (Birks et al., 2008:74).

4.10.4.4 Confirmability

Confirmability is described as the degree of neutrality or the extent to which findings of a study are shaped by the participants and not by researcher bias or interests (Nieuwenhuis, 2016:125). Teddlie and Tashakkori equate confirmability to objectivity. Shenton (2004:72) recommends that effort should be taken to ensure that findings are the result of the experiences and ideas of informants rather than the preferences of the researcher. To enhance confirmability and minimise researcher bias, I utilised the technique of member checks. This helped to ensure that I captured the informants' experiences and ideas accurately thereby eliminating the representational challenge which is a threat to validity (Ownwuegbuzie & Johnson, 2006:49). Triangulation also helped to minimise the effects of researcher bias which could have occurred as a result of my having developed a relationship with the participants and therefore likely to cloud

my judgment resulting in me missing certain things. The memoing technique further enhanced confirmability of my findings as I was able to acknowledge and admit my predispositions and beliefs underpinning decisions that I made in the research process thereby making me more self-aware which minimised researcher bias and enhanced confirmability.

4.11 Ethical considerations

The goal of any study is to get answers that are credible, but, these can only be accepted if the welfare of the participants has been preserved (Creswell, 2014:99). The protection of human participants in any study is critical and harm to research participants can be minimised or eliminated by observing suitable or fitting ethical principles (Orb, Eisenhauer & Wynaden, 2001:93). Ethics involve the moral choices that impact on the decisions, standards, and behaviour of the researcher (Greener, 2008:40). Although this study is a mixed method study, the ethical considerations that I observed applied across the two different strands of the research. I observed ethical considerations in three broad areas for this study, in seeking approval to conduct research, protecting participants' rights and in maintaining my integrity as a researcher.

4.11.1 Approval for conducting research

As a student of the University of Venda, my research proposal had to go through the School Higher Degrees Committee (SHDC) before proceeding to the University Higher Degrees Committee (UHDC) for approval. After this approval was granted (See appendix 1), it is the ethical procedure of the university that those research projects that deal with human and animal subjects have to get an ethical clearance from the University Ethics Committee to be allowed to conduct the study. To fulfil this institutional requirement, I thus had to apply for ethical clearance determining the level of participant risk posed by my study and outlining the measures that I would employ to establish the principle of beneficence (avoiding harm) to the participants. I was awarded an ethical clearance certificate (See appendix 2.2) after I had effected recommendations from the ethics committee (See appendix 2.1) on my application. After getting the ethical clearance from the University Ethics Committee, I then applied for permission to conduct the study from the Permanent Secretary of the Ministry of Primary and Secondary Education in Zimbabwe and permission was granted (See Appendix 3.1). This submission detailed the aims and objectives of the study, the

significance of the study, proposed study sites, sample frame, and sample size. The following documentation was also attached as supportive documents:

- Ethical certificate
- Consent forms
- Data collection instruments
- A supportive letter from the office of the registrar of my university.

The Midlands province then issued a letter which I took to Kwekwe district (See Appendix 3.2). At the district, the District Schools Inspector simply stamped the letter from the province with the district stamp to show that I had passed through the office (See Appendix 3.3). At school level, permission to conduct the study was sought from the Head of School before seeking for individual informed consent from the educators.

4.11.2 Participants' rights

Various ethical considerations were observed in an effort to uphold the rights of the participants.

4.11.2.1 Informed and voluntary consent

A basic ethical principle is that no one should be coerced to participate in a study, but rather, should participate voluntarily after making an informed decision (Neuman, 2014:151). The informed consent was a statement where I wrote and explained aspects of the study to the participants and asked for their voluntary agreement to participate before the onset of the data collection process. My consent form contained the following:

- i. Brief description and purpose of the study.
- ii. Statement of any discomfort associated with the study.
- iii. A guarantee of anonymity and confidentiality.
- iv. A statement that participation is completely voluntary and that it can be terminated anytime without explanation

The contents of the consent form were informed by Greener (2008:44); Neuman (2014:151) and Baker (1988:76). I had two consent forms, one for the questionnaire (See appendix 4.1) and another for the interview (See appendix 4.2).

4.11.2.2 Privacy, confidentiality and anonymity

As an ethical researcher who had considerable respect for the participants, I had to assure them that their privacy, confidentiality, and anonymity would be upheld throughout the course of the research. In my contact with the research participants, I had to ensure that I did not invade their privacy by sticking to the agreed timings of contact. I also made sure that I minimised violating participants' privacy by limiting my contact to the minimum possible and to what was only for legitimate research purposes (Neuman, 2014:154). I assured participants of confidentiality by holding the research data in confidence and keeping it out of the public domain. I released information in such a way that did not allow it to be linked to specific individuals and therefore, for the quantitative data, where individual-specific information is not the goal (Baker, 1988:75), information was only released publicly in collated form through averages and percentages. Another measure that was taken to ensure anonymity was to ensure that the questionnaires did not have any identifying information on them and they were only coded at data capturing using numbers. There was no need for me to trace the questionnaire to the respondent since I delivered and collected the completed questionnaire. The interview transcriptions were coded using the geographical code letter (R for Rural and U for Urban), and informant code (T for teacher, HOD for Head of Department and H for Head of School) and a number for the interview (1 to 4) for example, UT 1 means urban teacher 1; RHOD 1 means rural head of department 2; RH 1 means rural head 1. This form of coding was done for all the different categories of participants in the study. These measures were taken so as to ensure that the information was released in a way that did not allow it to be linked to specific individuals or sites. For the purpose of the analysis and the research report, participants were coded from 1 to 12. Participant 1 to 4 were teachers, participant 5 to 8 were heads of departments and participants 9 to 12 were school heads.

4.11.3 Researcher's conduct

A researcher's conduct during the research process is critical in determining the credibility of the study. A number of measures were implemented to ensure that my conduct upheld the integrity of the research profession.

4.11.3.1 Use of high technical standards

I maintained high technical standards during the research process by first familiarising myself with the range of types of research that have already been done on the topic so that I could understand the procedures that have been used and therefore be better positioned to decide on the appropriate design and methodology for this study. I was also critical in my approach by considering alternative ways of conducting the study (Baker, 1988:76) which helped to refine my methods and procedures. I desisted from bringing the research profession into disrepute by not falsifying, distorting or sensationalising the findings of the study.

4.11.3.2 Full disclosure

In an effort to establish transparency and dependability, I made it a point to disclose all aspects of the research for inspection and understanding by the research community. I achieved this by openly and fully disclosing the methods and all the findings including those that appeared insignificant. The disclosure was detailed enough to allow replication and this was made possible by keeping careful records of all steps of the research project.

4.12 Chapter Summary

Any research process is underpinned by the ontological, epistemological, methodological and axiological assumptions that the researcher holds. These assumptions are however driven by the research questions for the study. In this study, the pragmatist paradigm which is informed by both the positivist and constructivist assumptions was adopted for the study.

The convergent parallel design which involved conducting both the quantitative and qualitative strands of the research concurrently was used. Parallel sampling design was also used in the study. A quantitative sample comprising of 292 respondents generated by the stratified random sampling technique and a qualitative sample comprising of 12 respondents generated through stratified purposive sampling technique were used for the two strands of the study. Data was collected using a close-ended questionnaire for the quantitative strand and semi-structured individual face-to-face interviews for the qualitative strand. The two data sets, quantitative and qualitative were analysed separately but concurrently and then merged at interpretation and discussion phases. To ensure research rigor, quality control

measures to ensure reliability and validity in the quantitative strand and trustworthiness in the qualitative strand were instituted. For the research study to be respected and accepted, ethical considerations on permission to conduct and gain field entry, upholding the rights of the participants and on upholding researcher's ethical conduct were observed.

CHAPTER FIVE

Data presentation, Analysis and Discussion

5.1 Chapter Overview

This chapter presents and discusses the results from the quantitative and qualitative research strands of the study. The chapter starts by presenting descriptions of quantitative and qualitative samples before presenting and explaining the results of the regression analysis that was run. Presentation and discussion of results for each of the four objectives of the study was then done.

This study aimed to assess the effectiveness of the teacher performance evaluation system of Kwekwe district. The aim of the study was achieved through the following specific objectives that intended to:

- Investigate the relationship between the timing of the introduction of the teacher performance evaluation system and its effectiveness.
- Examine the design process of the teacher performance evaluation system in Kwekwe District.
- Appraise the implementation process of teacher performance evaluation system in Kwekwe district.
- Assess the significance of the teacher performance evaluation process in Kwekwe District in promoting professional development and accountability.

In presenting results for each of the objectives, the following approach was adopted. The quantitative results were first presented by giving an overview of the perceptions of the whole sample on the different items on the objective. The quantitative results on the same objective were then presented showing the perceptions of the urban and rural teachers separately. Thereafter, the qualitative results of the objective were presented. Participants 1 to participant 4 were teachers, participant 5 to participant 8 were heads of departments, and participant 9 to participant 12 were school heads. Quantitative and qualitative results were then merged before they were discussed. This was done for each of the four objectives. Participants also gave suggestions on what could be done to: improve appropriateness of timing of introduction, design

process, implementation process, and significance of the evaluation process and these recommendations are presented before giving a summary of the chapter.

5.2 Introduction

The quantitative data were collected using questionnaires completed by randomly selected rural and urban secondary school teachers while the qualitative data were collected from individual face to face interviews conducted with purposively selected key informant teachers, heads of departments and school heads from Kwekwe district in Zimbabwe. Data on multiple regression analysis tests namely, R Square, Analysis of Variance (ANOVA), and Beta Coefficient are presented before giving the results for each of the four subsidiary objectives derived from the RADPS conceptual framework on performance evaluation effectiveness. Data from the quantitative strand of the study was analysed using SPSS version 25 and is presented in frequency tables while the data from the qualitative strand was analysed using ATLAS. ti 8 (See appendix 6) to produce codes, themes, quotations, and networks views. Quotations and networks generated by ATLAS. ti 8 were used to illustrate the views and attitudes of the participants as revealed by the interviews.

5.3 Demographic information

This research adopted a mixed method study, therefore two samples of participants were used, the quantitative and qualitative samples.

5.3.1 Quantitative sample

The details of the quantitative sample are presented in Table 5.1.

Table 5.1 Demographic information on Quantitative sample

Variable		Frequency	Percentage
	Less than 25 years	2	0.7
	26 to 35 years	61	20.9
	36 to 45 years	120	41.1
	Over 45 years	109	37.3
	Total	292	100.0
Gender	Female	209	71.6
	Male	83	28.4
	Total	292	100.0
Highest Qualification	Certificate/Diploma in Education	161	55.1
	Bachelor of Education	79	27.1
	Master's in education	11	3.8
	Other (Please specify)	41	14.0
	Total	292	100.0
Teaching Experience	Below 5 years	22	7.5
	6 to 10 years	87	29.8
	11 to 20 years	100	34.2
	Over 20 years	83	28.4
	Total	292	100.0
School classification	Urban	189	64.7
	Rural	103	35.3
	Total	292	100.0

The quantitative sample was composed of a total of 292 respondents. A descriptive summary of the quantitative data is presented in Table 5:1. Of these 292 respondents, 41.1% were in the 36-45 years age group, 37.3% in the over 45 age group; 20.9% in the 26-35 years age group, and only 0.7% were below 25 years. A total of 78.4% of the sample was thus over 35 years old, an indication that there were very mature individuals dominating the sample. 71.6% and 28.4% of the respondents were female and male, respectively. In terms of the highest qualification, 55.1% have a Certificate or Diploma in Education; 27.1% hold a Bachelor of Education degree; 3.8% hold a Masters of Education degree and 14% hold other qualifications which are not specifically qualifications in education. The majority of the teachers in the sample, 62.4% have over 10 years of teaching experience and of these, 28.4 % have over 20 years of teaching experience; 29.8% have 6 to 10 years and only 7.5% have less than 5 years of teaching experience. The sample also constituted 64.7% teachers at urban schools and 35.3% teachers at rural schools.

Qualitative Sample

The details of the qualitative sample are presented in Table 5:2.

Table 5.2 Demographic information on qualitative sample

Variable		Frequency
Location	Urban	6
	Rural	6
	Total	12
Occupation	Teacher	4
	HOD	4
	Head	4
	Total	12

The qualitative sample was composed of a total of twelve participants and six of these were from urban schools while the other six were from rural schools. In terms of the appointments of the participants, the sample was composed of 4 participants in each of the 3 categories namely teachers, Heads of Departments and Heads of Schools.

5.4 Results and discussions

This section presents and discusses educators' perceptions and attitudes pertaining to the four objectives on effectiveness of the teacher performance evaluation system.

In presenting and discussing the results for each of the four objectives, the quantitative results of the whole quantitative sample are presented first. Quantitative results showing the comparative views of teachers based on their location (urban and rural) are then presented next before following a similar procedure with the qualitative data. The two data sets are then merged by comparing and contrasting the results from each of the two strands to either confirm or disconfirm each other. The presentation on each objective is wrapped up by discussing the results relative to existing literature and the RADPS conceptual framework on performance evaluation system effectiveness.

In presenting data for the different questionnaire items under each objective, I collated the strongly agree and agree together and presented them as respondents who agreed to a given statement and did the same for those who disagreed although the frequency table shows the different levels of agreement and disagreement. Percentages in the frequency tables were rounded up or down to one decimal point. The qualitative data was presented in quotations, and network views that were

generated using ATLAS. ti 8. The network views are the graphical presentations of quotations and codes shown as network relationships between the data codes (Lewis, 2004:453 & 458).

5.4.1 Presentation of Multiple regression analysis results

As part of the quantitative data analysis, multiple regression analysis was done and different tests that include R square, Analysis of variance (ANOVA) and Beta coefficient tests were conducted and the results of these are presented first before presenting the results for the different objectives. The purpose of the R square was to determine how well the regression model fits the data. The ANOVA test was done to establish the level of statistical significance of an independent variable (*timing, design, process* and *significance*) to predict the dependent variable (effectiveness of the evaluation system). The Beta co-efficient test was conducted to show the strength of the relationship between the dependent and independent variables.

5.4.1.1 Model summary

The model summary reports the strength of the relationship between the model and the dependent variable. The model summary table results are presented in Table 5:3.

Table 5:3 Model Summary table

Model Summary ^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.999 ^a	.998	.998	.02708
a. Predictors: (Constant), Significance, Timing, Design, Implementation				
b. Dependent Variable: Effectiveness				

The first column of the model summary has R which shows the multiple correlation coefficient (linear correlation between the observed and the model predicted values of the dependent variable). R square gives us the co-efficient of determination which is the squared value of the multiple correlation coefficient. It indicates the degree of the variation explained by the model. The adjusted R square is the corrected R square statistics, and it is the one that is considered for this study. In this case, the adjusted R square has a value of 0.998 to the co-efficient of determination which means that 99.8% of the variation in time is explained by the model. An R square value of >0.5 or

50% means strong fit (Muijs, 2004:165). This therefore means that the model was a strong fit for the data.

5.4.1.2 ANOVA Test

The ANOVA tests shows acceptability of the model from a statistical perspective and reports how well the regression equation fits the data. The ANOVA results for this study are presented in Table 5:4

Table 5:4 ANOVA table

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	103.249	4	25.812	35187.882	.000 ^b
	Residual	.211	287	.001		
	Total	103.460	291			
a. Dependent Variable: Effectiveness						
b. Predictors: (Constant) Timing, Design, Implementation and significance.						

The regression row shows information about the variation accounted for by the model. The residual row shows information about the variation that is not accounted for by the model. The regression sum of squares (103.249) is considerably larger than the residual sum of squares (.211) which shows that most of the variation is explained by the model. In this case the significant value is 0.000, and this means that there is a statistically significant relationship between the independent variables and the dependent variable (Cohen et al., 2011:662), thus, the independent variables of *timing*, *design*, *implementation process* and *significance of process* are statistically significant in predicting the effectiveness of teacher performance evaluation systems. In other words, this means that the variation explained by the model is not due to chance.

5.4.1.3 The Beta coefficient table

The Beta coefficient value explains the strength of the relationship between the independent variable and the dependent variable. In this study it explains the extent of the influence of timing, design process, implementation process and significance of

process on the effectiveness of the teacher performance evaluation system. The results of the Beta coefficient test are presented in Table 5:5.

Table 5:5 Beta Coefficient table

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.016	.009		-1.709	.089
	Timing	.186	.003	.213	60.308	.000
	Design	.204	.004	.204	50.944	.000
	Implementation	.298	.004	.368	81.983	.000
	Significance	.313	.004	.368	81.801	.000
a. Dependent Variable: Effectiveness						

In this study I used standardised Beta coefficient as I had standardised the variables before running the regression analysis. Standardisation of the variables was done by transforming the data from categorical scale to continuous scale. According to the results shown in this table, both implementation and significance have a Beta value of 0.368, which is the highest, meaning that the two have the strongest influence on effectiveness, followed by timing which has a .213. The Beta coefficient test shows that the design independent variable has the least influence on effectiveness relative to timing, implementation and significance. The standardised beta value of design is 0.204 and is the least value compared to the values of the other three independent variables.

5.5 Objective 1: To investigate the relationship between the timing of introduction and effectiveness of evaluation system

This objective sought to establish the effect of the timing of introduction of the performance evaluation system on its effectiveness.

5.5.1 Quantitative Results on objective 1

Questions in Table 5:6 were used to find out the teachers' perceptions on the appropriateness of the timing of introduction of the evaluation system.

Table 5.6 Perceptions of teachers on the appropriateness of the timing of introduction of the teacher evaluation system

	Frequency /Percentage				
	S D	D	NS	A	S A
Timing of inception of the performance evaluation system					
There was adequate consultation before inception of Results Based Management	86 29.5%	73 25%	74 25.3%	55 18.8%	4 1.4%
I was trained on Results Based Management System (RBM) before it was introduced.	76 26%	87 29.8%	22 7.5%	99 33.9%	8 2.7%
I was ready for RBM when it was introduced.	84 28.8%	119 40.8%	35 12%	48 16.4%	6 2.1%
I knew all the requirements for RBM.	78 26.7%	113 38.7%	51 17.1%	41 14%	9 3.1%
My school had all the necessary requirements for RBM when it was introduced.	58 19.9%	87 29.8%	84 28.8%	48 16.4%	9 3.1%
When RBM was introduced, I was not sure of what was expected of me	46 15.8%	37 12.7%	30 10.3%	45 15.4%	134 45.9%
RBM was needed in the education system to appraise teacher performance.	61 20.9%	49 16.8%	63 21.6%	92 31.5%	27 9.2%
My school head as the manager supported the Introduction of RBM as an appraisal system.	18 6.2%	18 6.2%	69 23.6%	143 49%	44 15.1%
For me, the introduction of RBM was at the right time.	69 23.6%	98 33.6%	69 23.6%	48 16.4%	8 2.7%
My opinion is that the timing of the inception of the teacher evaluation system has impacted positively on the teacher performance evaluation effectiveness.	73 25%	77 26.4%	73 25%	63 21.6%	6 2.1%

Questions in Table 5.6 of the questionnaire sought to find out the perceptions of the teachers on the appropriateness of the timing of the introduction of Results Based Management as a teacher performance evaluation system. These perceptions can be examined in greater detail by studying the frequency distribution of the individual questionnaire statements as reflected in table 5.6. Data indicate that out of the total of 292 respondents that participated in the study, 159 respondents (54.5%) disagreed, 71 (25%) were non-committal and only 59 (20.2%) agreed that adequate consultations had been done before the introduction of RBM as a performance evaluation system. There is an indication that training provided was not adequate as 163 respondents (55.8%) disagreed that they had been trained while 107 (36.6%) respondents agreed that they had been trained before the introduction of RBM. In terms of respondents' readiness for RBM at inception, 203 respondents (69.5%) did not agree that they were prepared for the system whereas 54 respondents (18.5%) respondents agreed that they had been prepared for RBM at its inception. Data also show that 167 respondents (57.2%) did not agree that the performance evaluation system had been introduced at the appropriate time but on the other hand, 56 respondents (19.2%) agree that the introduction of RBM was at the right time. In terms of the impact of the timing of the introduction of the performance evaluation system on the system effectiveness, 150 respondents (51.4%) disagree with the statement that the timing of the introduction

positively impacted the system effectiveness while 69 respondents (23.6%) think the timing of introduction impacted positively on the effectiveness of the system.

The data overall indicate that:

- Most teachers (54.5 %) indicated that they were not adequately consulted.
- Added to inadequate consultation, most teachers (55.8%) echoed that they were not adequately trained.
- One of the common perceptions by most teachers (69.5%) was that implementers were not ready when the system was introduced.
- The majority of the teachers (65.4%) said they did not know the requirements of the system at its introduction.
- A popular perception from the quantitative sample (57.2%) was that the system was not introduced at the right time.
- Most of the quantitative respondents (51.4%) were also of the opinion that the timing of introduction impacted negatively on the effectiveness of the performance evaluation system.

5.5.2 Comparison of quantitative results on objective 1 according to location

The perceptions of the urban and rural teachers on the appropriateness of the timing of introducing RBM were compared and the results are presented in Table 5.7

Table 5.7 Perceptions of teachers on timing based on the location of their school

Timing of inception of the performance evaluation system		SD	D	NS	A	SA	Total
There was adequate consultation before inception of Results Based Management.	Urban	59 31.2%	48 25.4%	46 24.3%	34 18%	2 1.1%	189
	Rural	27 26.2%	25 24.3%	28 27.2%	21 20.4%	2 1.9%	103
I was trained on Results Based Management System (RBM) before it was introduced.	Urban	50 26.5%	54 28.6%	13 6.9%	68 36%	4 2.1%	189
	Rural	26 25.2%	33 32%	9 8.7%	31 30.1%	4 3.9%	103
I was ready for RBM when it was introduced.	Urban	63 33.3%	73 38.6%	22 11.6%	28 14.8%	3 1.6%	189
	Rural	21 20.4%	45 44.7%	13 12.6%	20 19.4%	3 2.9%	103
I knew all the requirements for RBM.	Urban	57 30.2%	69 36.5%	36 19%	22 11.6%	5 2.6%	189
	Rural	21 20.4%	44 42.7%	15 14.6%	19 18.4%	4 3.9%	103
My school had all the necessary requirements for RBM when it was introduced.	Urban	42 22.2%	56 29.6%	58 30.7%	28 14.8%	5 2.6%	189
	Rural	18 17.5%	33 32%	28 27.2%	20 19.4%	4 3.9%	103
When RBM was introduced, I was not sure of what was expected of me	Urban	37 19.6%	23 12.2%	17 9%	32 16.9%	80 42.3%	189
	Rural	9 8.7%	14 13.6%	13 12.6%	13 12.6%	54 52.4%	103
RBM was needed in the education system to appraise teacher performance	Urban	46 24.3%	36 19%	35 18.5%	57 30.2%	15 7.9%	189
	Rural	15 14.6%	13 12.6%	28 27.2%	35 34%	12 11.7%	103
My school head as the manager supported the introduction of RBM as an appraisal system.	Urban	14 7.4%	12 6.3%	49 25.9%	83 43.9%	31 16.4%	189
	Rural	4 3.9%	6 5.8%	20 19.4%	60 58.3%	13 12.6%	103
For me, the introduction of RBM was at the right time.	Urban	52 27.5%	64 33.9%	44 23.3%	26 13.8%	3 1.6%	189
	Rural	17 16.5%	34 33%	25 24.3%	22 21.4%	5 4.9%	103
My opinion is that the timing of the inception of the teacher performance evaluation system has impacted positively on the teacher performance evaluation effectiveness	Urban	53 28%	50 26.5%	48 25.4%	36 19.1%	2 1.1%	189
	Rural	20 19.4%	27 26.2%	25 24.3%	27 26.2%	4 3.9%	103

The results shown in Table 5.7 show the comparative opinions of the teachers on the appropriateness of the timing of the introduction of RBM as a teacher evaluation system based on the geographical location of their school, that is, urban and rural. A total of 189 teachers in urban secondary schools took part in the study against 103 teachers from rural secondary schools. In terms of adequacy of consultation, 56.6% of the urban respondents against 50.5% of the rural respondents disagreed that adequate consultations had been done before the introduction of RBM. Of the 189 urban respondents, 55.1% compared to 57.2 % of the 103 rural respondents disputed

that they had been trained adequately before the introduction of RBM. With reference to readiness, 71.9% of urban respondents versus 65.1% of the rural respondents argued the fact that they had been ready when RBM was introduced. These results generally show that both the urban and rural teachers have similar perceptions on the appropriateness of the timing of introduction of RBM. There were however a bit of variation in terms of the perceptions of the respondents from the two geographical locations. With regard to the system being introduced at the right time, 61.4% of the urban respondents compared to 49.5% of the rural respondents opposed the fact and in terms of the timing of inception having positively impacted on the effectiveness of the system, and 54.5% of the urban respondents challenged the notion in contrast to 45.6% of the rural respondents. Although the percentage proportions of the respondents of the two geographical areas were quite different on these two aspects, what can be noted however is that in both cases there was similarity on the views of the majority of the respondents of the two locations. This therefore effectively means that the views were generally congruent.

5.5.3 Qualitative results on objective 1

The interview participants raised a variety of concerns pertaining to the timing of the introduction of RBM as a teacher performance evaluation system. The major concerns that were raised included provision of inadequate training on the RBM system, imposition of RBM system, lack of buy-in of RBM by the educators and the issue of having inadequate resources to support a new system. Figure 5.1 presents a summative picture of the different concerns raised by the 12 participants.

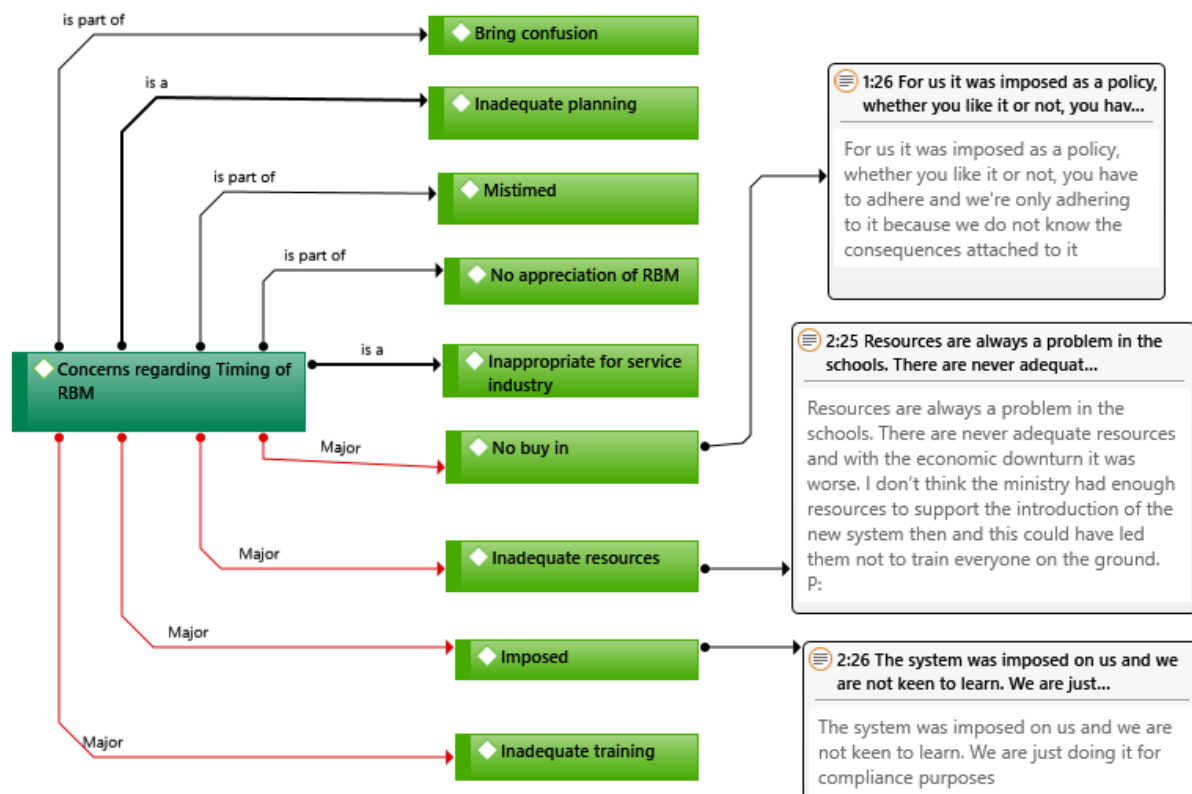


Figure 5.1 Concerns regarding the timing of the programme

Theme 1: Inadequate training

The majority of the participants indicated that one concern they had was that adequate training was not provided prior to the introduction of the performance evaluation system. The lack of adequate training meant that the majority of the teachers did not really know and understand what was expected of them in the implementation process of the system. This is reflected in the quotations by some of the participants:

We didn't actually know what was expected of us... When the system came, only a few people were coached and when they came back to the schools, they didn't actually understand what they had learnt. We were all confused about what we were supposed to do, so we were simply copying and pasting from each other. The ones that knew a little would just give us some guidelines and then the rest of the school would copy from each other (Participant 2: teacher).

Similar to participant 2 (teacher), Participant 5 (head of department) also indicated that, “We did not know exactly what it was all about and sometimes we get confused. We simply complete the forms because they have to be completed”. Participant 5 (head of department) further explained that, “Yes they did not have much information,

although workshops were held but then people were still confused. They did not get adequate information on what it was all about”.

These sentiments by the participants indicate that the trainings that were conducted did not help the educators to know and understand how they were supposed to implement the RBM system. One of the participants also highlights that not all the educators were trained. He stated that, *“HODs were called for training and when they came back, we asked them exactly what we were supposed to do?”* (Participant 3). Indications were that a few were nominated for training and these were then expected to train their colleagues on returning to their station. The challenge that arose was that those who had been trained had not understood the implementation procedures well thereby leading to them providing sub-standard training which failed to demystify the new system, but rather, only brought confusion to the educators. Due to the inadequate understanding and knowledge, educators ended up simply copying from the few who had limited knowhow in order to comply with the policy requirements.

Theme 2: Imposition of RBM system

Participants also felt that the system had just been imposed on them without any consultation or involvement. Eight of the participants raised the issue of imposition as a concern as reflected in some of the participants comments that, *“I personally feel that this was imposed on us. There wasn’t any dialogue. We were just given the forms and asked to fill them in”* (Participant 7: head of department). Another participant also reiterated this point by highlighting that, *“The system was imposed on us and we are not keen to learn. We are just doing it for compliance purposes”* (Participant 2: teacher). The sentiments of the participants are that there was no dialogue between the RBM drivers and the supposed implementers. The interviews revealed that because of the imposition of the RBM, educators were thus not keen to learn. The imposition created the resistant attitude among the educators as explained by participant 2. Participant 8 (head of department) weighed in and said, *“It was just mandatory for people to do it but most of us were not aware why we were doing that and we didn’t even see the importance of the appraisal process or understand what had to be done”.*

The views of the participants illustrate inadequate ownership of the system as it was imposed on them which brought some resistance in the implementers as evidenced by the statement by one of the participants that, *'... we are not keen to learn'*. Educators simply do implementation for compliance with policy as the implementers claim they do not appreciate the appraisal process since they do not see any benefits or understand how they are supposed to implement the system because the system had simply been imposed on them. There is also an implication that the drivers of RBM did not attempt to market the system enough as revealed by Participant 8 (head of department) above that, *"... we didn't see the importance of the appraisal process..."*. The educators did not appreciate the system because they were not aware of its benefits and advantages.

Theme 3: Lack of buy-in of RBM

During the interviews, it also emerged from the interviews that there was no buy-in because the system had just been imposed on the teachers. The ultimate effect of the top-down approach was that there was no buy-in or ownership of the system by the teachers and administrators who were supposed to implement it. Participant 7 (head of department) explained that, *"... as a result embracing that system might be very difficult especially if you do not know what you are supposed to be doing"*. The implication that can be deduced here is that educators failed to embrace the system partly because of the inadequate training that they had. The participants accordingly did not even understand or know what they were supposed to do in implementing it. Participant 4 (teacher) on the other hand argued that, *"There was a need for the ministry to assess if the stakeholders embraced the system before introducing it"*.

It also emerged from the interview that educators feel deceived because the impression given is not what is happening, *"... while this system was brought on the pretext of upholding and improving the performance of teachers, it is in fact just meant to burden us"* (Participant 7: head of department). They regard it just as a burden rather than anything of benefit hence the refusal to embrace it.

Theme 4: Inadequate resources to support new system

It also arose during the interviews that resources were inadequate and this could have affected the appropriateness of the timing of the performance evaluation system.

Fiscal resources were needed in many areas to ensure effective implementation of the system. Fiscal resources were needed to fund the requisite training for the different implementers, buy material resources, fund the targeted professional development programs and also for the provision of incentives. In support of this assertion, participant 2 (teacher) said, “... *I don’t think the ministry had enough resources to support the introduction of the new system then and this could have led them not to train everyone on the ground*”. The lack of resources is therefore blamed for the inadequate training that was provided to the implementers. To emphasise the inappropriateness of the timing of introduction of the system, the same participant explains that, “*Resources are always a problem in the schools. There are never adequate resources and with the economic downturn it was worse*”. Indications were that resources have always been a problem in the schools and the economic downturn worsened the situation which would thus hinder teachers from effectively discharging their duties. The implication is that introducing an appraisal system then would therefore be inappropriate since it would be very challenging for teachers to meet the new policy requirements.

Educators also revealed that it was unwise for the driving ministry to introduce the RBM system without budgetary support to cater for necessities like incentives. This illustrated in the quotation by one of the participants that, “*There was a need for the system to have monetary incentives because without such incentives, the teachers cannot be motivated*” (Participant 4: teacher). The argument that is advanced is that teachers cannot be motivated enough to assume the additional load without the corresponding associated benefits.

The general findings that emerged from the individual interviews were that:

- Initial training was inadequate leading to confusion and lack of knowledge of implementation as illustrated by one of the teachers that, “*There are a lot of grey areas that had not been cleared in terms of training and equipping the teachers to understand RBM as the new system of assessment,*” (Participant 7: head of department).
- RBM was imposed and there was no stakeholder involvement. This general view is illustrated by a statement by one of the participants that, “... *I agree with*

the sentiments that it was just something that was just imposed on the teachers without consultation..." (Participant 10: school head).

- The non-participatory, top-down approach resulted in lack of buy-in and acceptance of the programme. Evidence in support of this claim can be illustrated by Participants 12 (school head) who said, *"... It was information coming from the top. There was no consensus before that instrument was used. The concern was how can we implement an instrument that we did not take part in crafting,"*.
- Resources were inadequate to support the introduction of a new programme then. This is revealed in the statement by one of the participants that, *"I don't think the ministry had enough resources to support the introduction of the new system then and this could have led them not to train everyone on the ground"* (Participant 2: teacher).

5.5.4 Comparing and contrasting quantitative and qualitative results on the appropriateness of timing of introduction of the evaluation system

Results from both the quantitative and qualitative strands point towards a similar opinion on the appropriateness of the timing of the introduction of the evaluation system. Quantitative findings indicate that there was no consultation (54.5%), meaning that the system was just imposed, training was inadequate so implementers did not know the requirements of the system (65.4%) and therefore they do not believe that the timing of introduction of the system was appropriate (57.2%). Quantitative results also indicate that respondents feel that the inappropriate timing of the introduction stage negatively affected the effectiveness of the performance evaluation system (51.4%). Similar sentiments were also raised from the interviews that the system had simply been imposed leading to a lack of buy-in as illustrated in the quotation from participant 12 (school head) that, *"There was no consensus before that instrument was used. The concern was how can we implement an instrument that we did not take part in crafting"*. This implies that the lack of buy-in or failure by the implementers to embrace the system emanated from the negative feelings evoked by the imposition. The aspect of training also emerged from the interviews in support of what was indicated by the respondents in the questionnaire. It emerged that training had been inadequate resulting in implementers failing to both understand and know how the

system was to be implemented. A participant explained that, “... *A lot of things remained hazy for the teachers.... We just do it because it is policy but to be very honest with you, many people, even today, are still not conversant with the RBM*” (Participant 9: school head). Respondents to the questionnaire indicated that timing of the introduction of RBM negatively affected its effectiveness. This was explained in the interviews by one of the participants that, “*The ministry should not have rushed to introduce the system but should have waited until it had adequate financial resources*” (Participant 4: teacher). The implication is that the resources had been inadequate then to support the demands of a newly introduced system which would accordingly negatively impact on the effectiveness of the appraisal system. Results from the quantitative and qualitative research strands therefore corroborate each other in that they both point to the fact that the timing of the introduction of the performance evaluation system was inappropriate because:

- There was no consultation before the introduction of the system thereby implying that the system was imposed on the stakeholders. This was confirmed by participants such as Participant 7 (head of department) who declared that, “*I personally feel that this was imposed on us. There wasn’t any dialogue*”.
- Training was inadequate, thus the stakeholders were not even sure of how they were supposed to implement the system and a number of participants described the situation as so dire that some teachers just copied from their colleagues to comply with the requirement. Participant 2 (teacher) stated that, “... *We were all confused about what we were supposed to do so we were simply copying and pasting from each other...*”
- Participants felt that the system was introduced at an inappropriate time because there was inadequate training as remarked by Participant 1 (teacher) that, “... *the second concern was the timing of its introduction, there was no proper workshop held to train and equip all the educators...*”. It also emerged that the driving ministry had inadequate resources to support the new system which could have compromised the training that was provided. This was highlighted in comments like, “*I don’t think the ministry had enough resources to support the introduction of the new system then and this could have led them not to train everyone on the ground*” (Participant 2: teacher). It was also revealed that lack of buy-in ensued due to lack of marketing as indicated, “...

very unfortunate to say it was not sold but it was imposed..." (Participant 1: teacher).

- The teachers therefore felt that the timing of the introduction of the system was inappropriate and thus impacted negatively on the effectiveness of the performance evaluation system. In explaining this fact, Participant 4 (teacher) said, *"The ministry should not have rushed to introduce the system but should have waited until it had adequate financial resources"*.

5.5.5 Discussion of quantitative and qualitative results on timing of introduction

A number of issues emerged from the study pertaining to the timing of the introduction of the RBM as a teacher performance appraisal system. Key issues that were raised included the imposition of the system, inadequate resources, and inadequate training.

Imposition of the performance evaluation system

Results from the study show that the performance evaluation system was imposed on the teachers without them being consulted in any way or making any input. This lack of involvement resultantly led to lack of ownership and buy-in by the teachers and may be indicative of a system that was introduced prematurely with inadequate marketing having been done to the stakeholders. Pretorius and Ngwenya (2008:156) emphasize the point that a program has to be marketed to the implementers so that they are made conscious of its benefits if the stakeholders are to embrace and accept it. The teachers in this study have not embraced the system 12 years after its adoption and are simply implementing it for compliance purposes without much appreciation of its objectives and benefits. Machingambi (2013:223) and Gutuza (2016:394) also note that there has been very little stakeholder involvement in the Zimbabwean case and this has ultimately resulted in lack of ownership of the system. Gutuza (2016:395) asserts that lack of involvement of the stakeholders breeds opposition and negative regard of the system. Santiago and Benavides (2009:27) on the other hand proffer that involvement of the teachers or their unions minimises teacher resistance and promote buy-in as occurred in the case of Chile.

In conducting a readiness assessment as proposed by the RADPS conceptual framework, ascertaining the demand for the intended programme is also emphasised. The study has confirmed the importance of readiness assessment when the issue of

imposition and lack of dialogue was raised. Demand for a product can be created by marketing the product so that people's awareness of its benefits is raised which in turn promotes appreciation and buy-in. As marketing is done, a dialogue is created which helps to demystify the new product that is aimed for introduction hence promoting acceptance. It can therefore be concluded that the findings from the study echo the propositions of the RADPS conceptual framework.

Inadequate resources to support new system

Resources are key in the introduction of any new programme. This study established that resources were inadequate to support the introduction of RBM at the time. Kusek and Rist (2004:43) assert that it is critical to identify existing gaps that have to be built before introducing a new evaluation system if it is to be sustainable. Institutional capacity assessment has to be done to determine institutional capacity in terms of the available skills and expertise, fiscal and material resources. This is deemed critical since it facilitates the determination of the feasibility of the intended programme. The findings from this study on resources resonate well with the suggestions from the RADPS conceptual framework for performance evaluation system effectiveness. The conceptual framework submits that in conducting a readiness assessment, which ascertains the appropriateness of timing, it is necessary to determine the institutional capacity to implement and sustain the new system. This involves establishing if the resources are adequate to support the introduction of a new system which this study has also confirmed.

Findings from the individual face to face interviews also shed a light that the inadequacy of resources extended beyond the fiscal resources to include the material resources in the form of teaching and learning material which further complicated and compromised the implementation of the teacher evaluation system. Machingambi (2013:224) supports this by asserting that availability of resources enables schools to purchase materials and resources that support the teaching process and professional development programmes that capacitate the teachers with better instructional skills.

Inadequate training

Interview participants revealed that the teachers were confused and did not know or understand much on how they were supposed to implement RBM. This finding resonates with those established by Machingambi (2013:223) and Gutuza (2016:395) that the teachers were incapacitated and lacked skills to implement the evaluation system properly. This was attributed to the inadequate training that had been provided to them probably due to the limited availability of resources to fund the training programme. Literature substantiates that resource constraints led to most of the employees not getting the requisite training to allow them to be conversant with the new system (Madhekeni, 2012:125; Machingambi, 2013:223) which resultantly led to their ability to implement the system being compromised. Pre-implementation training as confirmed in this study and extant literature, and similarly proposed by the RADPS conceptual framework, helps to ensure that the requisite expertise and knowledge is inculcated in the implementers to capacitate and ready them for the process.

5.6 Objective 2: To examine the design process of the teacher performance evaluation system

The question of this study was to find if the design process of the teacher performance evaluation system in Kwekwe influences its effectiveness.

5.6.1 Quantitative results on objective 2

The perceptions of teachers on the design process were established through the questions in Table 5:8

Table 5:8 Perceptions of teachers on the design process

Design process of teacher performance evaluation system	Frequency /Percentage				
	S D	D	NS	A	S A
There was adequate consultation when the teacher performance evaluation system was designed.	95 32.5%	80 27.4%	69 23.6	44 15.1%	4 1.4%
The appraisal process at my school is guided by regulations on appraisal Statutory 1 of 2000.	14 4.8%	11 3.8%	90 39%	148 50.7%	29 9.9%
Regulations on appraisal in Statutory 1 of 2000 are not used at my school.	10 3.4%	35 12%	114 38.4%	80 27.4%	53 18.2%
The appraisal process is based on clearly stated job descriptions	27 9.2%	41 14%	54 18.5%	140 47.9%	29 9.9%
RBM is based on clear standards of performance	28 9.6%	41 14%	53 18.2%	142 48.6%	28 9.6%
I know enough about what constitute good teaching as a practitioner.	16 5.5%	11 3.8%	28 8.9%	151 51.7%	86 29.6%
I know the methods that can be used to get accurate performance information.	18 6.2%	38 13%	74 25.3%	126 43.2%	36 12.3%
The RBM appraisal system at my school has rating levels that differentiate well-performing teachers from poor performing teachers.	32 11%	45 15.4%	76 26%	121 41.4%	18 6.2%
Teacher performance appraisal at my school is based on classroom observation only.	39 13.4%	100 34.2%	46 15.8%	91 31.2%	16 5.5%
Learner achievement is used to appraise teacher performance.	148 50.7%	16 5.5%	46 15.8%	52 17.8%	30 10.3%
The appraiser uses information from different sources to Appraise me.	26 8.9%	34 11.6%	68 23.3%	131 44.9%	33 11.3%
The teacher appraisal process at my school has clear guidelines.	34 11.6%	150 51.4%	51 17.5%	37 12.7%	20 6.8%
The appraisal process at my school is well planned	148 50.7%	44 15.1%	55 18.8%	23 7.9%	22 7.5%
My opinion is that the design of the teacher evaluation system has impacted positively on teacher performance evaluation effectiveness.	60 20.5%	105 36%	61 20.9%	19 6.5%	47 16.1%

The second part of the questionnaire sought to establish the perceptions of the teachers on the design process of the evaluation system and its possible impact on the effectiveness of the evaluation system and the design itself by getting the views of the teachers on the aspects of the system like clarity of evaluation guidelines, multiplicity of sources of evaluation information, and ability of evaluation process to distinguish teachers performing at different levels. Table 5.8 presents detailed results of the teachers' perceptions on the various items on the design aspect of the evaluation process. The majority of the respondents, 175 (59.9%) indicated that there had been no adequate consultation during the evaluation system design process. This lack of active teacher participation was there in spite of the majority teachers 237 (81.2%) indicating that as practitioners they know about what constitute good teaching and 162 (55.5%) teachers also pointing out that they know the methods that can be used to get the accurate teacher performance information. Teachers also pointed out that the appraisal process does not have clear guidelines which could affect their

ability to implement the system. On the other hand, 170 (58.2%) teachers acknowledged that the evaluation process is based on clear standards of performance and multiple sources of information as indicated by 164 (56.2%) teachers. On the ability of the rating levels to distinguish the well-performing teachers from the poor performing teachers, 139 (47.6%) teachers agree that the system is differentiating although 77 (26.4%) teachers feel otherwise. In spite of learner achievement being the key outcome of an educational process, 164 (56.2%) teachers admit that learner achievement is not used as a measure of teacher performance in the evaluation process. Some teachers, 165 (56.5%) feel that the design process of the teacher evaluation system generally negatively impacted on the effectiveness of RBM.

The results on the design process indicate that:

- The majority of teachers (59.9%) indicated there was no adequate consultation when teacher evaluation system was designed.
- A higher proportion of teachers (58.2%) acknowledged that the evaluation system is based on clear standards of performance and 47.6% of the teachers pointed out that the evaluation system rating levels are quite differentiating. It should be noted however that 26% of the teachers indicated that they were not sure if the rating levels were differentiating.
- Indications by most of the teachers (56.2%) were that the evaluation process is based on multiple sources of evaluation information.
- Although most teachers (81.3%) admitted that as practitioners, they know a lot about what constitutes good teaching and 55.5% of the teachers added that they know the methods that can be used to get accurate performance information, 59.9% of the teachers however pointed out that they were excluded from the design process.
- Most teachers (63%) attested that there are no clear guidelines to the implementation process of the evaluation system and (56.2%) of the teachers highlighted that learner achievement is not used as a measure of teacher performance.

- The majority of teachers (56.5 %) concluded that the design process negatively impacted on the performance evaluation system effectiveness.

5.6.2 Comparison of perception of teachers on design process of the evaluation system

The perceptions and opinions of the urban and rural teachers on the design process were compared and the results presented in Table 5:9.

Table 5:9 Perceptions of teachers on the design process based on the geographical location of the school

Design process of teacher performance evaluation system		SD	D	NS	A	SA	Total
There was adequate consultation when the teacher performance evaluation system was designed.	Urban	63 33.3%	55 29.1%	46 24.3%	22 11.6%	3 1.6%	189
	Rural	32 31.1%	25 24.3%	23 22.3%	22 21.4%	1 1%	103
RBM is based on clear standards of performance	Urban	21 11.1%	24 12.7%	36 19%	92 48.7%	16 8.5%	189
	Rural	7 6.8%	17 16.5%	17 16.5%	50 48.5%	12 11.7%	103
I know enough about what constitute good teaching as a practitioner.	Urban	12 6.3%	9 4.8%	18 9.5%	94 49.7%	56 29.6%	189
	Rural	4 3.9%	2 1.9%	10 9.7%	57 55.3%	30 29.1%	103
I know the methods that can be used to get accurate performance information.	Urban	14 7.4%	28 14.8%	46 24.3%	79 41.8%	22 11.6%	189
	Rural	4 3.9%	10 9.7%	28 27.2%	47 45.6%	14 13.6%	103
The RBM appraisal system at my school has rating levels that differentiate well-performing teachers from poor performing teachers.	Urban	23 12.2%	29 15.3%	50 26.5%	75 39.7%	12 6.3%	189
	Rural	9 8.7%	16 15.5%	26 25.2%	46 44.7%	6 5.8%	103
Teacher performance appraisal at my school is based on classroom observation only.	Urban	27 14.3%	59 31.2%	30 15.9%	63 33.3%	10 5.3%	189
	Rural	12 11.7%	41 39.8%	16 15.5%	28 27.2%	6 5.8%	103
Learner achievement is used to appraise teacher performance.	Urban	100 52.9%	29 15.3%	29 15.3%	21 11.1%	9 4.8%	189
	Rural	48 46.6%	23 22.3%	16 15.5%	9 8.7%	7 6.8%	103
The appraiser uses information from different sources to appraise me.	Urban	18 9.5%	23 12.2%	43 22.8%	81 42.9%	24 12.7%	189
	Rural	8 7.8%	11 10.7%	25 24.3%	50 48.5%	9 8.7%	103
The teacher appraisal process at my school has clear guidelines.	Urban	21 11.1%	96 50.8%	31 16.4%	25 13.2%	16 8.5%	189
	Rural	13 12.6%	54 52.4%	20 19.4%	12 11.7%	4 3.9%	103
The appraisal process at my school is well planned	Urban	97 51.3%	26 13.8%	37 19.6%	15 7.9%	14 7.4%	189
	Rural	51 49.5%	18 17.5%	18 17.5%	8 7.8%	8 7.8%	103
My opinion is that the design of the teacher evaluation system impacted positively on teacher performance evaluation effectiveness.	Urban	46 24.3%	54 28.6%	41 21.7%	14 7.4%	34 18%	189
	Rural	14 13.6%	51 49.5%	20 19.4%	5 4.8%	13 12.6%	103

The results shown in Table 5:9 indicate the comparative perceptions of the teachers in urban and rural secondary schools on the design process of RBM. Regarding consultation during the design process, 62.4% of the urban respondents compared to 55.4% of the rural respondents disputed that adequate consultation had been done. In regard to the standards of evaluation, 57.2% of the urban respondents against

60.2% of the rural population agreed that the evaluation process is based on clear evaluation standards. In respect of the multiplicity of the sources of evaluation information, 55.6% of the urban respondents contrary to 57.2% agreed that multiple sources of information are used in the evaluation process. While 46% of the urban respondents agreed that the ratings levels used are differentiating, 50.5% of the rural respondents affirmed to it too. Regarding teachers' knowledge on what constitutes good teaching, 79.3% of the urban respondents compared to 84.4% of the rural respondents accepted that teachers know enough of what constitutes good teaching. Regarding the methods that can be used to get an accurate teacher performance, 53.4% of the urban teachers against 59.2% of the rural teachers admitted that they are cognisant with the procedures that can be used to get an accurate teacher performance. Concerning evaluation guidelines, 61.9% of the urban respondents compared to 65% of the rural respondents disputed that there are clear evaluation guidelines.

It was also noted that 68.2% of the urban teachers versus 68.9% of the rural teachers disagreed that learners' achievement is used to measure teacher performance. With reference to the effect of the design process on the evaluation system effectiveness, 52.9% of the urban teachers against 63.1% of the rural teachers disputed that design procedure had a positive impact. The results as indicated in the table generally show that both the urban and rural teachers have similar perceptions pertaining to the design aspect of RBM. This claim is based on the comparative proportion of the opinions of urban and rural teachers on given aspects of the design process.

5.6.3 Qualitative results on the design process of RBM

A number of issues were revealed from the individual face-to-face interviews. The main issues that were raised included lack of stakeholder involvement, usefulness of evaluation guidelines, ability of evaluation form to capture teacher performance accurately, ability of evaluation form to measure performance comprehensively and capacity of the instrument to be informative. Detailed explanations of the participants' views on each of the raised issues are presented in the following section of the report.

Theme 1: Inadequate stakeholder participation in the design process.

The majority of the participants, in the study indicated that there was no stakeholder involvement in the design process as the system was simply imposed on the implementers. The following sentiments raised by some participants in the individual face-to-face interviews help to illustrate the point. Participant 2 (teacher) posits that, *“We were never involved in anything. It was just imposed on us. We were simply told that there is a new system and we should adhere to the new system. We were never involved as stakeholders”*. The implication of this statement is that besides the lack of involvement in the design process, the system was simply imposed without any dialogue or consultation with the educators. In support of this assertion, another participant pointed out that, *“As far as I know, we were not involved in the design process. Like I said, we had the KRA system before RBM and out of the blue, we were told that we are now using RBM system, so we were not involved in any way”* (Participant 7: head of department). It was also revealed from the interviews that in addition to being side-lined in the design process, the educators were not even aware of how the system had been developed. Indications are that no marketing was attempted as people were simply told that a new system was replacing the one that had been in place as revealed in this quotation, *“It was just imposed on us. We were just told that we have changed from key result areas to RBM. How it was developed, I am not aware”* (Participant 9: school head). Another participant weighed in saying, *“I don't know (laughs) I simply have to use that instrument but I don't know how it was developed”* (Participant 5: head of department). This accentuates the point that there was neither involvement nor dialogue with the educators during the development process of RBM.

Theme 2: Unavailability performance evaluation guidelines

The Public Service Commission has produced guidelines that are meant to facilitate the proper implementation of the evaluation process. This study revealed that the guidelines are quite useful as indicated by the comments by some participants that, *“Yes the guidelines are quite helpful in completing the appraisal instrument”* (Participant 12: school head). Participant 8 (head of department) also reiterated the usefulness of the guidelines by elucidating that, *“Yaa, I think if you do it without the guidelines it's even tough. Some time we go back to the guidelines when you are rating to check what the different ratings mean”*. These sentiments imply that the guidelines

are quite helpful in assisting the educators in the evaluation process, hence, they are useful.

There was however an argument which was advanced that, *“There is a need for more education on the RBM, more trainings on the system and how to use the guidelines so that we are clear. At the moment we are still confused and not clear on what has to be done”* (Participant 4: teacher) .This remark reiterates the importance of the issue of training in the success equation for the teacher performance appraisal process. The statement implies that availing the evaluation guidelines without providing the requisite initial training may render the effort futile as this is necessary for the implementers to be able to understand the process clearly.

Although the Public Service Commission has guidelines on the completion of the appraisal form, the study revealed that these are not available to most of the educators. Indications by the majority of the participants were that the guidelines have never been availed to them and hence, they were even unaware of their existence. The following remarks by some participants show that evaluation guidelines have not been readily availed to most of the participants thereby limiting their intended usefulness in the process. Participant 2 (teacher) declared that, *“The guidelines? We do not even know the guidelines. We just copy previous forms. Probably the previous forms are the guidelines since they are the ones that we just copy (laughs). We have never seen it. It has never been given to us at this station”*. Participant 1 (teacher) also weighed in affirming that, *“...There are no other guidelines besides the appraisal form itself. You simply read the form, the subheadings on the form, those should give you the guidelines on how to complete the form”*. These remarks from the participants show that the evaluation guidelines are not readily available to the participants. These participants are not even aware of their existence and yet these very guidelines are meant to help the educators in completing the appraisal form. It emerged that the participants resort to copying from each other because they have challenges in completing the evaluation form and yet there are guidelines that are meant to assist in this regard but unfortunately are not readily availed to the supposed-users.

It was also revealed from the interviews that even in some cases where the guidelines are available, the participants do not really use them because of the negative attitude

or refusal to embrace the performance evaluation system as illustrated by some participants' responses:

I think I have it but I have never read it (laughs). I am being honest (laughs again). I can't comment on their usefulness (laughs) because I have never read it (laughs again). The actual thing madam is that this thing is being imposed so we lose interest. It is supposed to be something that you embrace, something that you would want to participate in but if it is just imposed, you tend to ignore some things that may take your time and you just concentrate on producing what has to be done because at the end of the day what they want is a completed form and I will produce it and give it to them so the guidelines might not be necessary to me (laughs)
(Participant 7: head of department).

Another participant reiterated that, *"No, I am not sure if I use the guidelines because I am still using the work plan from 2005 just dubbing because I don't know anything about this RBM?... More work and less money, so you just dub and leave it"*
(Participant 6: head of department).

This negativism shows a refusal to embrace the performance evaluation system. The fact that some teachers acknowledge having the evaluation guidelines that are supposed to guide and assist them in the implementation process, but openly admit that they feel it is a further waste of their time to refer to the guidelines is rather unfortunate and a cause for concern. It just emphasises the lack of buy-in by the educators which in turn compromises the implementation process.

Theme 3: Inaccurate performance evaluation form

The interviews also sought to establish the perception of the participants on the ability of the evaluation instrument to capture performance accurately. The majority of the participants, felt that the evaluation instrument has no capacity to capture the performance accurately. A remark by one of the participants to illustrate the point was that, *"There is a contradiction... a teacher who gets a high rating in the appraisal process can get 0% pass rate for national examinations. This therefore makes the ability of the form to capture the teacher's actual performance level doubtful"*
(Participant 1: teacher). This statement insinuates that the capacity of the evaluation form to capture teacher performance is questionable and this is supported by Participant 8 (head of department) who weighed in saying, *"Aaah, I don't think it captures the performance level accurately because the rating is done when the results are not yet out and people simply estimate. There is no way therefore that the form*

can really capture the actual performance unless the rating was done after the results are out". The observation from these remarks is that for the evaluation instrument to capture the actual performance of a teacher, then, it has to have a way of including the students' results as part of the measures of teacher performance. The suggestion therefore is that the instrument was wrongly designed as it fails to capture an important indicator of a teacher's performance thereby pointing to a faulty design process that could be attributable to the exclusion of the teachers as experts in the teaching science.

Observations by some participants are that there is also a tendency to bunch teachers in certain ratings in spite of their actual performance probably as a way of ensuring protection of each other's careers. A remark by a participant to illustrate this point said,

Most teachers are bunched between 3.5 and 4. A 5 is avoided because they claim that it will attract avoidable scrutiny from the authorities. It is therefore just an exercise done to fulfil the requirement of the ministry and not one done to capture the actual performance of the teacher (Participant 1: teacher).

This remark implies that 'safe ratings' are awarded to teachers to comply with the policy requirements while at the same time avoiding attracting external authorities to validate either very low or high ratings. This remark suggests that the ratings are unsupported by empirical evidence, and hence, do not necessarily reflect the actual performance of the teacher. Participant 10 (school head) also weighed in saying,

Three people with everything the same, individuals are different but when we come to the ratings you will find that people in the same department will have one final rating. If it is 4, then it is 4 for everyone in the department but because we are different individuals and likely to perform differently so it makes sense that the marks should be different but because of a problem, manpower shortage, there are no people who are following to make sure those variations are evident.

The argument presented is that it is not practical and real that people in a department can all deserve a similar rating. The usual and real scenario is that people are bound to perform differently and thus deserving different ratings. The idea of having all people in a department getting a similar rating is suspicious and hence the ability of the evaluation form to capture the actual performance is doubted. It shows that it is open to manipulation that is why Participant 10 (school head) decried the lack of manpower to validate the summative ratings awarded to the educators.

The performance evaluation system is purported to be results based. The remarks by the majority of participants however show that they query the capacity of the system to measure performance based on the argument that the system simply measures activities in spite of it supposedly being results-based. According to the teachers, the results of any teaching process cannot be divorced from the learning outcomes which are partly seen through the results of the learners. Ideally, the system is therefore supposed to go beyond the outputs but rather extend to the outcomes. The fact that the evaluation instrument is limited to the outputs is blamed on a faulty design process that occurred due to the side-lining of the educators as experts in their own area. Participant 1 (teacher) explained that,

Yes, it measures activities, did you go for your lesson, yes I did, then you are a good teacher. Did you mark the exercise books, yes I marked the exercises, then you are a good teacher and so we're saying only laid out activities are the ones that are measured but not the final impact of the activities on the learner. The final mark and everything else are not related.

In support, another participants added that,

We are looking at how many lessons have you conducted, how many co-curricular sessions did you conduct but there is nowhere where we will talk about whether the pupils won any competition or what did they achieve, we don't write that. We just look at how many sessions did I attend, 12 sessions, and we assess that a teacher attended 9/12 possible sessions, that is that. (Participant 4: teacher)

The opinion of the educators is therefore that the instrument does not manage to capture the actual performance of the teacher and attribute this shortfall to a defective design process.

Theme 4: Ability of the evaluation instrument to capture performance comprehensively

The opinion of most of the interview participants is that the evaluation instrument captures the performance comprehensively. The following remark by one of the participants indicated this:

It tries to do that because it starts from well before delivery of a lesson from lesson preparation, the actual process of delivering a lesson, monitoring and assessment of learners work to extending beyond the classroom to co-

curricular activities. ... it therefore means that covers it all aspects of teaching (Participant 7: head of department).

In support of this assertion, Participant 11 (school head) declared, *“In a way we can say it does because it looks at the academic and co-curricular aspects. It therefore tends to cover all the aspects of teaching”*.

The study thus reveals that the evaluation instrument is comprehensive. Although the participants acknowledged that the system is comprehensive, it was however revealed that the implementation process is not done as it should. Participant 8 (head of department) explained, *“I would say that if it was done properly, it is quite comprehensive but then in fact it is not done as it should”*.

The implication is that although the instrument is comprehensive, the challenge however, is that it is not implemented correctly as people simply complete the form for compliance purposes and thus do not follow the correct procedure. This in turn tends to defeat the purpose of the process.

Theme 5: Weakness of the evaluation instrument to be informative

This study also explored the opinions of the participants on the capacity of the instrument to be informative. Figure 5.2 gives an overview of the opinions of the interview participants on the informativeness of the evaluation instrument.

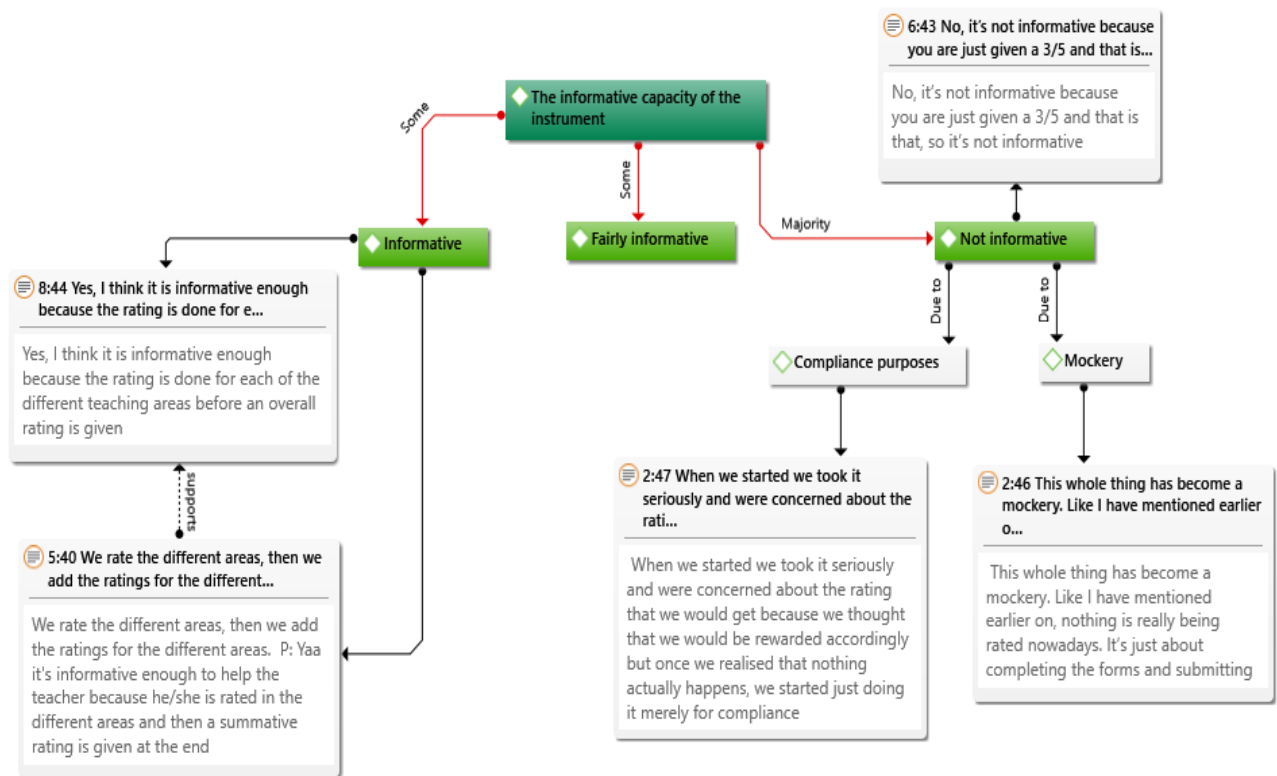


Figure 5:2 Capacity of the evaluation instrument to be informative

This study established that the evaluation instrument is not adequately informative in spite of it being comprehensive in its coverage. The majority of the participants argued that the instrument is no longer used appropriately thereby rendering the whole exercise a mockery and a mere compliance exercise that is meaningless. The majority of the participants indicated that the implementers simply put ratings which are not evidence-backed as there is no evaluation activity on the ground besides simply completing the form for compliance purposes. As illustrated by a remark presented in Figure 5.7, Participant 2 (teacher) asserted that, “*This whole thing has become a mockery. Like I have mentioned earlier on, nothing is really being rated nowadays. It’s just about completing the forms and submitting*”. The whole process is compromised because it has been turned into a mere facade. This is further supported by observations made by Participant 11 (school head) that,

I would not describe the rating process as informative enough to allow improvement because this is not done the way that it should. It is being done just for compliance purposes so it’s not helpful. The rating is usually an empty number that is not a reflection of the situation on the ground.

The implication here is that there was inadequate performance monitoring that is done. The exercise is merely for compliance purposes and not backed by practice thereby rendering the summative rating that is awarded hollow and meaningless.

There are however some participants who are of the view that the instrument is informative since it gives ratings on the different aspects of teaching thereby making it instructive and communicative. This is supported by remarks like the one by Participant 8 (head of department) in Figure 5.7 that says, *“Yes, I think it is informative enough because the rating is done for each of the different teaching areas before an overall rating is given”*. Participant 12 (school head) further added that, *“Yes, rating process is informative because the different teaching areas are rated separately before an overall rating is given”*. The implication of what was revealed from the study is that the evaluation instrument is designed well to be informative as the ratings are done for the different teaching areas before they are collated to give an overall rating. The rating of the different aspects therefore can assist the evaluatee to identify their areas of strength and weakness thereby making the instrument informative. The complication however is that although the instrument has been designed well, the inadequacy to implement compromises the capacity of the instrument. This is backed by the assertion that the design of an evaluation system may be commendable but its success is to a large extent dependent on how well it is implemented (TNTP, 2010:9).

In brief, the following findings emerged from the individual face-to-face interviews:

- Participants feel that the design process was not all-inclusive but rather excluded the implementers which in turn negatively affected buy-in and probably resulted in the development of inappropriate evaluation tools. This is illustrated by remarks such as, *“It was just an instruction given to school heads to carry out.... There was no consensus before that instrument was used.... We were not given time to try the instrument and give an input back to the superiors so it was do this now...”* (Participant 12: school head)
- The performance evaluation guidelines are useful. This was revealed by some sentiments like, *“Yaa, I think if you do it without the guidelines it's even tough. Some time we go back to the guidelines when you are rating to check what the different ratings mean,”* (Participant 8: head of department).

- Guidelines on completion of evaluation form are not available in most of the schools. To explain the situation, Participant 2 (teacher) said, *“The guidelines? We do not even know the guidelines”*.
- Most participants felt that the instrument has no capacity to capture the performance accurately. They argued that the system is purported to be results-based and yet it focuses on the outputs or activities and disregards the outcomes which are the results. Participant 8 (head of department) argued that, *“... I don't think it captures the performance level accurately because the rating is done when the results are not yet out and people simply estimate. There is no way therefore that the form can really capture the actual performance unless the rating was done after the results are out”*. Participants also argued that the evaluators want to avoid attracting the external evaluators so there is a tendency to avoid the extreme ratings (both the outstanding and dismal ratings) and rather bunch the teachers between 3.5 and 5 which is not usually the true reflection of the reality of the teacher performance on the ground. Participant 1 (teacher) explained that, *“Most teachers are bunched between 3.5 and 4. A 5 is avoided because.. it will attract avoidable scrutiny from the authorities. It is... just an exercise done to fulfil the requirement... not one done to capture the actual performance of the teacher”*.
- Most participants felt that the evaluation instrument captures performance comprehensively although they acknowledge that the challenge is that the whole thing is a fallacy since there is really nothing on the ground. This is revealed in the following statement, *“I would say that if it was done properly, it is quite comprehensive but then in fact it is not done as it should”* (Participant 8: head of department).
- Some participants felt that the instrument is informative since it gives the ratings on the different aspects of teaching which makes it instructive. This is confirmed by the following remark, *“Yaa it's informative enough to help the teacher because he/she is rated in the different areas and then a summative rating is given at the end”* (Participant 5: head of department). The majority of the participants however argued that the evaluation process is a mere illusion since nothing is really taking place on the ground in most cases except filling the

evaluation reporting form for compliance when it is due, participants therefore felt that the whole exercise is meaningless and uninformative. In support of this claim, Participant 11 (school head) said, *“I would not describe the rating process as informative enough to allow improvement because this is not done the way that it should. It is... just for compliance purposes so it’s not helpful. The rating is usually an empty number...”* This claim by the participants resonates with the assertion by TNTP (2010:9) that although the design of the evaluation system may be commendable, its success however largely depends on how well it is implemented.

5.6.4 Comparing and contrasting of quantitative and qualitative results on the design process

It has emerged from both the quantitative and qualitative strands of the study that the design process was not participatory as it excluded the key stakeholders thereby leading to a majority of the implementers failing to embrace and accept it. A total of 54.5% of the questionnaire respondents affirmed that no consultation had been done before introducing RBM and this was substantiated in the interviews through remarks such as, *“The system was imposed on us and we are not keen to learn. We are just doing it for compliance purposes”* (Participant 2: teacher).

The majority of the teachers (63%) disputed that there are clear evaluation guidelines in the school. This was corroborated by the participants in the interviews that evaluation guidelines are not available in most of the schools that participated in the study. Participant 2 (teacher) explained that, *“... We have never seen it. It has never been given to us at this station”*. Another participant reiterated that, *“Aa, I don’t remember us getting one. I am sure that we do not have it... But I do not think we got guidelines that you are talking about”* (Participant 10: school head). These indications by the participants suggest that they were not even aware of the existence of the guidelines that are supposed to guide them in completing the evaluation form. It was revealed by the participants that most simply copy from peers and old evaluation work plans. The challenge by the Public Service Commission to avail the evaluation guidelines to all the implementers further compromised the implementation process that had already been compromised due to the failure to provide adequate pre-implementation training. There were however a few participants who revealed from the interviews that they had the guidelines and acknowledged their usefulness in

facilitating successful implementation. Participant 12 (school head) confirmed that, *“Yes the guidelines are quite helpful in completing the appraisal instrument”*.

It was also revealed that the evaluation instrument has inadequate capacity to capture teacher performance accurately. Educators argued that the instrument is open to manipulation as it allows the bunching of teachers in certain ‘safe ratings’, allows rating to be done even in the absence of anything taking place on the ground and disregards learner achievements in spite of them being the major outcome of the learning process. Participant 1 (teacher) explained that, *“Most teachers are bunched between 3.5 and 4. A 5 is avoided because they claim that it will attract avoidable scrutiny from the authorities. It is.... done to fulfil the requirement of the ministry and not done to capture the actual performance of the teacher.* The quantitative strand also established that learner achievement is not considered as a measure of teacher performance as indicated by 56.2% of the questionnaire respondents and this was corroborated during the interviews. Most of the interview participants argued that disregarding the outcomes of the teaching process by overlooking learner achievement as a measure of teacher performance renders the instrument irrelevant as student learning is the major outcome of the learning process. Participant 2 (teacher) explained that, *“... It ignores the learner achievements and therefore we cannot actual say that it measures results”*.

Findings from the two strands therefore corroborate each other as they both give the impression that the non-participatory approach to the design process was faulty and led to lack of buy-in and acceptance by the implementers thereby frustrating the effort of the practice as the whole process turned into a mere perfunctory, compliance exercise.

On a more positive note, the study established that the evaluation process is based on clear standards as pointed out by 58.2% of the questionnaire respondents. The performance targets are discussed and agreed to by the evaluator and evaluatee as clarified by Participant 9 (school head) in the statement that, *“... The appraiser who is the HOD discusses the work plan with the appraisee and they agree on the targets and the review dates...”*. The discussion of the performance targets by both the evaluator and evaluatee induces critical reflection by the concerned parties which is positive for the evaluation process. It also emerged from both strands of the study

that the evaluation instrument is comprehensive as it covers the different aspects of teaching due to using multiple sources of evaluation data as confirmed by 56.2% of the questionnaire respondents. To elucidate this, Participant 12 (school head) said, *“The system captures all aspects of teaching adequately. The teaching aspects and the co-curricular activities are considered although these different aspects contribute differently to the final rating”*. Given the complexity of the teaching exercise, it may be difficult for a single data source to be able to present an accurate picture of the teacher’s performance. The use of multiple sources of data enhances the reliability of the evaluation findings as the weaknesses of one source of data is compensated by the strength of another sources.

5.6.5 Discussion of quantitative and qualitative results on the design process inadequate stakeholder involvement in design of RBM

This study established that the design process was largely non-participatory. Neither the implementers nor their representatives in the form of the teachers’ unions participated in the design process and this resultantly affected buy-in and acceptance of the system. The teachers therefore have a negative attitude towards the evaluation system and have failed to embrace it even after it has been operational for nearly a decade. This finding resonates with that by Machingambi (2013:223) and Gutuza (2016:394) that in the Zimbabwean case, stakeholder involvement has been missing. Machingambi (2013:223) asserts that lack of teacher involvement breeds negativity that ultimately affects the overall effectiveness of the system. This position is declared by Musingafi et al. (2013:26) who posit that imposing a programme on the people guarantees failure and is an obvious recipe for disaster. This finding by the study confirms the proposition by the RADPS conceptual framework which similarly advocates for a collaborative approach to the development of acceptable performance evaluation systems. Research has proven that the involvement of stakeholders in the design phase of performance evaluation systems promotes buy-in and ownership as well as decreases the chances of resistance (Heneman et al. in Isoré, 2009:9; Mayne, 2007:3). This study thus confirmed the importance of stakeholder involvement in the design process as a way to promote buy-in thereby echoing the proposition of the RADPS conceptual framework.

It is argued that a system that is borne out of the contributions of all key stakeholders is better respected and embraced than one that is simply imposed on the users (Goe

et al., 2014:14; Stronge and Tucker, 1999:353; White et al., 2012:12). Danielson and McGreal weigh in by proffering that it is only the direct users that can certify the validity and relevance of an evaluation system and therefore the involvement of the users from the drawing table therefore helps to validate the system before it even gets off the ground which is commendable. This is reiterated by Stronge and Tucker (1999:341) when they postulate that stakeholder participation is critical in aiding the development of a sense of ownership by allowing expressions and discussions of conflicting ideas and negotiations before the implementation of policy which ultimately enhances the final product.

Usefulness of and availability of evaluation guidelines

This study revealed that the evaluation guidelines produced by the Public Service Commission are clear and useful. This resonates with the RADPS conceptual framework that proposes the development of clear and precise guidelines as a way of facilitating successful implementation. Haefele (1993:26) asserts that having guidelines that are clear and precise coupled with well-trained evaluators helps to enhance the consistency and integrity of the performance evaluation process. Availing evaluation guidelines is critical in facilitating successful implementation of the evaluation process. This study found out that although the Public Service Commission produced the evaluation guidelines, the Commission however seemed to have challenges of availing them to all the supposed users. It therefore defeats the purpose for which they were developed. The RADPS conceptual framework advocates for the provision of such documents to ensure effective implementation. The situation in Kwekwe district however is contrary to this proposition as a greater number of the schools that participated in the study do not have the guidelines. An enabling environment has to be created for a fruitful evaluation process. Institutional commitment has been suggested as one way of creating an enabling environment for proper implementation (McLaughlin, 1990:403). Conley (1987:64) proffers that institutional commitment can be shown in numerous ways like availing and allocating the necessary resources for implementation. Guidelines and manuals are critical resources for assisting the implementers to understand the evaluation procedure. Guidelines and manuals also help to standardise the whole procedure hence availing them is crucial in ensuring proper and uniform implementation across the schools. The unavailability of the guidelines in most of the schools in the study therefore

compromises the implementation process and may result in an implementation process that is not standard and uniform across the schools. This is also contrary to the proposal of the RADPS conceptual framework that it is important to provide clear guidelines to facilitate the proper implementation of the evaluation procedure.

Inadequacy of an evaluation instrument to capture performance accurately

It also emerged from the study that the teachers are of the view that the instrument is inappropriate and irrelevant, hence it fails to capture performance accurately. The implication is that the instrument was not well designed and this problem could be attributed to the lack of involvement of the teachers as the experts in their field who would obviously know how their actual performance can be captured. Goe et al. (2014:14) support this position by postulating that the teachers as the experts are better placed to tell that which can work or may not work in the teaching and learning process. As demonstrated by the RADPS conceptual framework, stakeholder involvement guarantees the relevance and appropriateness of the evaluation instrument derived from the exploitation of their expert knowledge. Machingambi (2013:223) also reiterates that the involvement of teachers in the design and development of the evaluation instruments and procedures not only motivates, empowers and helps to familiarise them to the evaluation process but can also help them to embrace the system to such an extent that they can even become the system advocates. The study finding thus confirms the assumptions of the RADPS conceptual framework on the design process that the process has to be participatory to enhance the appropriateness and relevance of the evaluation instruments.

This study also established that learner achievement is disregarded when assessing teacher performance using RBM. Participants however argue that disregarding learner achievement as a measure of teacher performance renders the instrument irrelevant as the evaluation process is supposed to be results-based. According to Haefele (1993:24), this instrument can be described as having measurement deficiency, meaning that it excludes some criterion that is considered being important and representative of teaching. In support of this position is the finding by Zhang and Ng (2017:208) who established that both the teachers and administrators in China also regard student results as being more factual and meaningful relative to other measures of teacher performance. This position however contradicts the finding by

Mavhiki et al. (2013:137) who posit that teachers have challenged the use of pass rate as a performance measure. Literature (TNTP, 2010:2; Santiago & Benavides, 2009:16) argues that the primary responsibility of a teacher is to assist the student to learn which therefore justifies the use of student achievement as a measure of teacher performance. The growing call to use student learning as a measure of teacher performance has in turn led to the use VAMs gaining ground as an acceptable measure of teacher performance (Darling-Hammond et al., 2012:8). The RADPS conceptual framework emphasises stakeholder involvement in the design process as a way to ensure relevance and appropriateness of evaluation instruments. Evaluation instruments are very technical and it requires the input of the teachers as the experts and duty bearers to develop an instrument which is relevant and can accurately capture their performance.

Use of multiple sources of evaluation information

On a more positive note, this study established that RBM does not rely on a single information source for its evaluation data but rather uses multiple sources of evaluation information. Due to the complexity of the teaching exercise, literature argues that no single data source can present a complete and accurate picture of the teacher's performance (TNTP, 2010:6; Santiago & Benavides, 2009:14; Peterson, 2000:91-102). Goe et al. (2014:19) add that the use of multiple sources of evaluation data helps to increase confidence in the results thereby enhancing the acceptability of evaluation results. The use of multiple sources of information also allows for compensation of weakness of information source (Hinchey, 2010:7) thereby preventing bias. The multiplicity of information sources allows the evaluation process to generate a comprehensive view of the teacher's performance which is in line with the suggestions of the RADPS conceptual framework. Relying on a single data source may present a distorted picture of the performance of the teacher thereby defeating the purpose of evaluation.

5.7 Objective 3: To appraise the implementation process of teacher performance evaluation system

One interest of this study was to determine how the implementation process affects the effectiveness of the teacher performance evaluation system in Kwekwe district.

5.7.1 Quantitative results on the implementation process

Questions in Table 5:10 were used to determine the opinion of the teachers on the implementation process.

Table 5:10 Perceptions of teachers on the implementation process

	Frequency /Percentage				
	S D	D	NS	A	S A
Implementation process of performance evaluation					
At my school, every teacher's performance is appraised.	15 5.1%	22 7.5%	40 13.7%	165 56.5%	50 17.1%
At my school, performance targets are set and agreed on by both the appraiser and appraisee at the beginning of the appraisal cycle.	11 3.8%	20 6.8%	20 6.8%	184 63%	57 19.5%
At my school, performance is appraised and rated on agreed targets only.	11 3.8%	30 10.3%	44 15.1%	168 57.5%	39 13.4%
At my school, performance appraisal actually measures the results of performance by the teachers.	25 8.7%	142 48.6%	51 17.5%	49 16.8%	25 8.6%
There is continuous monitoring of performance during the appraisal cycle at my school.	156 53.4%	44 15.1%	33 11.3%	18 6.2%	41 14%
Review discussions at my school are based on progress made towards the achievement of performance targets.	153 52.4%	43 14.7%	51 17.5%	18 6.2%	27 9.2%
I am given an opportunity to comment during performance reviews at the school.	29 9.9%	38 13%	26 8.9%	164 56.2%	35 12%
I can comment during the appraisal final rating process.	29 9.9%	50 17.1%	30 10.3%	153 52.4%	30 10.3%
I know the procedure to be taken if I fail to agree with the appraiser on the work plan for the appraisal process.	81 27.7%	56 19.2%	93 31.8%	43 14.7%	19 6.5%
I know the procedure to be taken if I fail to agree with the appraiser on the final rating.	84 28.8%	56 19.2%	86 29.5%	49 16.8%	17 5.8%
The performance appraisal at this school considers all aspects of teaching.	19 6.5%	37 12.7%	44 15.1%	158 54.1%	34 11.6%
The person who appraises me at the school is knowledgeable on teaching evaluation.	18 6.2%	32 11%	39 13.4%	149 51%	54 18.5%
My appraiser has adequate time to conduct the appraisal activities.	125 42.8%	61 20.9%	41 14%	30 10.3%	35 12%

The greater numbers of teachers, 241 (82.5%), agree that performance targets are set and agreed on by both the appraiser and appraisee at the beginning of the appraisal cycle. The majority of teachers in the study, 167 (57.3%), do not think that the performance appraisal process actually measures the results of the teacher performance. It is also evident from the information that there is limited continuous monitoring of performance during the appraisal cycle as indicated by 200 (68.5%) of the teachers. The observed facts mean that some other methods may be used to assess the teachers other than the set targets, or alternatively, it might mean that the ratings that are given at the end of the appraisal cycle are fictitious since there is no continuous monitoring. This is

also apparent as shown by 200 (67.1%) teachers, that review discussions are not based on progress made towards achieving the performance targets. The majority of the teachers, 137 (46.9%), did not have knowledge on the recourse they would take in the event that they do not agree with their appraisers on the work plans. This general lack of knowledge may imply that no one has been failed to render him or her to need the recourse for correction or that the teachers are not familiar with the appraisal policy as the grievance procedure is clearly laid out in Public Service Regulations (2000:50-52). It also follows from the information that teachers did not agree that the appraisal process actually measures their performance as 167 (57.2%) teachers do not agree while 75 (25,3%) agree. It is also noted that 186 (63.7%) of the teachers are of the opinion that their appraiser does not have adequate time to conduct the appraisal process which can explain the reason why there is likely no continuous monitoring of the teachers' performance.

However, the teachers agreed that they are given a chance to comment during the appraisal as well as having a say in their final rating as shown by the statistics 199 (68.2%) and 183 (62.7%) respectively. Many of the teachers 192 (65.7%), think that the teacher performance evaluation system considers all aspects of teaching and also believe that the appraisers have knowledge on teacher evaluation.

In summary, the teachers appreciate that:

- Nearly all teachers (82.5%) confirmed that performance targets are set and agreed on by both the appraiser and appraisee at the beginning of the appraisal cycle.
- A significant proportion of teachers (57,3%) refuted that the performance appraisal measures the teachers' performance.
- Most teachers (68.5%) argued that there is no continuous monitoring of teachers' performance during the evaluation cycle.
- To further support that there is no consistent monitoring, 67.1% of the teachers declared that the review discussions in the schools are not based on the progress made towards achievements of the set targets.
- There is limited knowledge as to the recourse to take in the event that there is a disagreement on the work plan or final rating. Only 22.6% of the teachers admit that they know the procedure to take while 48% confess that they do not know and 31.8% state that they are not sure. This effectively means that 79.8% of the

teachers are not aware of the recourse action to take in the event of disagreements during the appraisal process.

- Most teachers (68.2%) accepted that they can comment during appraisal and (62.7%) also admitted that they have a say on their final rating.
- A lot of teachers (63.7%) pointed out that the appraisers do not have adequate time for the appraisals processes.

5.7.2 Comparison of quantitative results on implementation based on location

Detailed comparative perspectives of participants on the different aspects of the implementation process are presented in Table 5.11.

Table 5:11 Comparisons of quantitative results on implementation process according to location

Implementation process of performance evaluation		SD	D	NS	A	SA	Total
At my school, every teacher's performance appraised.	Urban	11 5.8%	12 6.3%	31 16.4%	102 54%	33 17.5%	189
	Rural	4 3.9%	10 9.7%	9 8.7%	63 61.2%	17 16.5%	103
At my school, performance targets are set and agreed on by both the appraiser and appraisee at the beginning of the appraisal cycle.	Urban	8 4.2%	13 6.9%	10 5.3%	118 62.4%	40 21.2%	189
	Rural	3 2.9%	7 6.8%	10 9.7%	66 64.1%	17 16.5%	103
At my school, performance is appraised is rated on agreed targets only.	Urban	8 4.2%	14 7.4%	26 13.8%	119 63%	22 21.4%	189
	Rural	3 2.9%	16 15.5%	18 17.5%	49 47.6%	17 16.5%	103
At my school, performance appraisal measures the results of performance by the teachers.	Urban	21 11.1%	88 46.6%	33 17.5%	31 16.4%	16 8.5%	189
	Rural	4 3.9%	54 52.4%	18 36%	18 17.5%	9 8.7%	103
There is continuous monitoring of performance during the appraisal cycle at my school.	Urban	99 52.2%	27 14.3%	25 13.2%	14 7.4%	24 12.7%	189
	Rural	57 55.4%	17 16.5%	8 7.8%	4 3.9%	17 16.5%	103
Review discussions at my school are based progress made towards the achievement performance targets.	Urban	98 51.9%	31 16.4%	34 18%	11 5.8%	15 7.9%	189
	Rural	55 53.4%	12 11.7%	17 16.5%	7 6.8%	12 11.7%	103
I am given opportunity to comment during performance reviews at the school.	Urban	20 10.6%	25 13.2%	19 10.1%	104 55%	21 11.1%	189
	Rural	9 8.7%	13 12.6%	7 6.8%	60 58.3%	14 13.6%	103
I can comment during the appraisal final rating process.	Urban	20 10.6%	28 14.8%	18 9.5%	105 55.6%	18 9.5%	189
	Rural	9 8.7%	22 21.4%	12 11.7%	48 46.6%	12 11.7%	103
I know the procedure to be taken if I fail to agree with the appraiser on the work plan for the appraisal process.	Urban	56 29.6%	37 19.6%	56 29.6%	31 16.4%	9 4.8%	189
	Rural	25 24.3%	19 18.4%	37 35.9%	12 11.7%	10 9.7%	103
I know the procedure to be taken if I fail to agree with the appraiser on the final rating.	Urban	55 29.1%	41 21.7%	51 27%	33 17.5%	9 4.8%	189
	Rural	29 28.2%	15 14.6%	35 34%	16 15.5%	8 7.8%	103
The performance appraisal at this school considers all aspects of teaching	Urban	17 9%	25 13.2%	27 14.3%	96 50.8%	24 12.7%	189
	Rural	2 1.9%	12 11.7%	17 16.5%	62 60.2%	10 9.7%	103
The person who appraises me at the school is knowledgeable on teacher evaluation.	Urban	13 6.9%	23 12.2%	28 14.8%	91 48.1%	34 18%	189
	Rural	5 4.9%	9 8.7%	11 10.7%	58 56.3%	20 19.4%	103
My appraiser has adequate time to conduct the appraisal activities	Urban	79 41.8%	40 21.2%	26 13.8%	21 11.1%	23 12.2%	189
	Rural	46 44.7%	21 20.4%	15 14.6%	9 8.7%	12 11.7%	103

Results on the Table 5:11 show that 83.6% of the urban teachers compared to 80.6% of the rural teachers agree that performance targets are set and agreed to by both the appraiser and appraisee. In terms of commenting during the appraisal process, 66.1%

of the urban teachers against 71.9% of the rural teachers admitted that they can comment during the process. The majority of the teachers, 63.5% of the urban teachers against 69.9 % of the rural teachers confirmed that the appraisal process considers all aspects of teaching. Most of the teachers, 65.1% of the urban teachers contrary to 58.3% of the rural teachers affirmed that they can comment during the summative rating process. A significant number of the teachers, 57.7% of the urban teachers compared to 56.3% of the rural teachers, disputed that the appraisal process measures the actual performance of the teachers. With regards to the monitoring of performance, 66.3% of the urban teachers in contrast to 71.9% of the rural teachers refute that there is continuous monitoring of the teacher performance through the appraisal cycle. A substantial number of the teachers, 68.3% of the urban teachers versus 65.1% of the rural teachers disagree that review discussions are based on progress made towards achievement of performance targets. A greater proportion of the teachers, 63% urban compared to 65.1 % of the rural teachers, disprove that the appraiser has adequate time to conduct the appraisal process. The study revealed that a greater number of the teachers, 50.8% urban against 42.8% rural teachers are not familiar with the procedure to take in the events of disagreements during the appraisal process. It should also be noted that on the aspect of knowledge of procedure to take in event of differences, 27 % of the urban compared to 34 % of the rural teachers indicated that they are not sure which may indicate that a total of 77.8% of the urban versus 76.8% of the rural teachers are not knowledgeable of the procedure. It is evident from results presented here that the teachers in both the urban and rural settings have generally similar opinions on the different aspects of the implementation process.

5.7.3 Qualitative results on the implementation process

Interviews were conducted with twelve participants to find out their opinions, perceptions and experiences on the performance appraisal implementation process. A number of issues emerged from these discussions and these helped to shed light on the implementation process. These issues are outlined below.

Theme 1: Setting of performance targets

The majority of the participants indicated that the evaluation process starts by the evaluator and evaluatee discussing and agreeing on the work plan and performance targets for the particular appraisal cycle. Upon agreeing, the work plan is signed by the two parties. Participant 4 (teacher) pointed out that,

Each teacher has been allocated an appraiser who is the HOD of their teaching area. At the beginning of the appraisal cycle, the appraisee is supposed to draft a work plan which has to be discussed with the appraiser to agree on the performance targets. After the discussion and agreement on the targets, the two sign the work plan for that cycle.

This position was also supported by Participant 5 (head of department) who remarked that, “... *The teacher comes up with a work plan which he/she discusses with me so that we agree on the performance targets for that year. After discussing and agreeing we both sign the work plan for that appraisal cycle*”. These remarks indicate that there are discussions between the evaluator and evaluatee on the work plan and performance targets for a particular appraisal cycle. The remarks by the participants also point to the fact that performance targets are not pre-set but rather are discussed and agreed on by both parties involved in the evaluation process. It however also emerged from some participants that there is really no time for the discussions of the work plans and that teachers simply present to the evaluators the work plan which in some cases are simply copied from previous years. Participant 8 (head of department) commented that, “*Yaa, normally there is no time to sit and discuss the work plan. An individual will just sit down and does the work plan on his own, then go to the appraiser when it's already done and the appraiser will just sign...*”. To shed more light on the issue, participant 2 (teacher) also indicated that, “...*When we are told we just make sure that by that date we would have submitted our forms. What we simply do is to copy the work plan from the previous year and submit the form on the due date*” (laughs).

Other issues that emerged during the discussion on preparation of the evaluation process are that, in coming up with the work plan, the teacher is guided by basically three documents namely:

- the Departmental Integrated Performance Agreement (DIPA) formulated by the Provincial Education Director.
- The Director's Circular Number 36.
- The subject syllabus.
- The timetable.

It was however clear that the preparation process is not standard as there were some variations across the stations. The insufficiently standardised procedure could indicate confusion and absence of understanding and knowledge of the procedure and this can be attributable to inadequate training and possibly inadequate supervision and monitoring. The absence of evaluation guidelines in most of the schools as revealed by the study can also contribute to the absence of a standardised implementation procedure across the schools in the district.

Theme 2: Evaluation procedure, an event

Just like what emerged on the preparation process, it also emerged during the interviews that the process of implementation is not standard and varies across the schools. It appeared that in most schools, the evaluation process is simply done for compliance and it is more of an event than a process. Indications were that in most of the schools, the teachers simply write the work plans at the beginning of the appraisal cycle and only go back to the shelved work plans again when it is time to do the final rating procedure. The following sentiments of a participant bears testimony to the observation:

The exercise is simply done for compliance and due to challenges like time, work plans are submitted without much discussion between the appraiser and appraisee. Appraiser just write their work plan and submit. The review meetings are hardly conducted because of lack of time (Participant 11:school head).

To further explain on the implementation of the evaluation process, Participant 8 (head of department) added that, "...I just fill it at the beginning of the year and submit it at the end of the appraisal cycle. Nothing really happens between the beginning and the end of the cycle...". Remarks by the participants point out that the evaluation procedure is not a continuous process as it should be, but rather, more of an event or activity that is merely done to fulfil the policy requirement.

Theme 3: Objectivity and validity of rating process

Although a few participants acknowledged that the use of performance targets help to make the rating process objective, it however emerged from the majority of the participants that the rating process is neither objective nor valid. A number of reasons have been advanced for the opinion. The first reason that was given was that there is a bunching of teachers with the very low or high ratings being avoided. Participant 10 (school head) elucidated more on that issue by explaining that,

... teachers are just bunched either at 4 or 5 .The assessment therefore becomes subjective, the results which are being given right now are quite subjective .They do not reflect a true picture of the whole situation because people are just cooking marks .They sit and decide a mark which will be ok for the school or for the department, and they give the mark so it's not real.

From the interviews and discussions with the teachers, they said no teacher is failed as the whole process is managed. Teachers are awarded average ratings that are not too low or too high to put the authorities at bay. The implication is that the process is just a fallacy as it appears not to be evidence-based. Ratings seem to be given subjectively without a particular justification and ensuring that certain ratings are deliberately avoided as a way of avoiding the need for validation by external evaluators who are officials from the district. The quest to avoid external evaluators leads to the bunching of the teachers thereby intentionally avoiding extremely low ratings like 1 and 2 and also outstanding rating which is a 6.

The second reason given for disputing the objectivity and validity of the rating exercise was that in this case, the evaluation procedure was an event rather than a process so rating was not evidence-based since there was virtually no activity between target setting and rating. Participant 8 explained this in the remark that, “... *I just fill it at the beginning of the year and submit it at the end of the appraisal cycle. Nothing really happens between the beginning and the end of the cycle so some ratings are not reflective of what is actually happening on the ground*”. Another participant further elucidated the fact that the evaluation process is just an event,

... Like I said earlier on, there is nothing that actually happens between the beginning and end of the appraisal cycle. Its usual teaching without any reference to the work plans that we agreed on at the beginning of the appraisal cycle and if things have gone wrong during the cycle, then they will only be seen at the end when it's too late to do anything. There is no

review in between as is supposed to be the case. Although quarterly review dates are indicated but there is nothing that actually happens on the ground so at the end of the day, the marks that are given as the rating have no reflection whatsoever of the actual happening on the ground. It is difficult therefore to talk about objectivity and validity (Participant 7: head of department).

Participant 2 (teacher) also added that, *“It isn’t (laughs) because there isn’t supervision. There isn’t any RBM in the school, it’s just a fallacy. Yaa, nothing is happening on the ground”*. The implications of the remarks made by the participants suggested that the evaluation process did not have a formative function since there were indications that there was no monitoring and supervision of the performance of the teachers.

The third reason given for regarding the process as invalid was that it focuses on the outputs and ignores quality and outcomes. Participant 1 (teacher) argued that,

The process however fails to capture the quality aspect... That is why we say you can have a teacher producing 8 As and another teacher produces 1 A but the 8 A teacher might not have a very good mark when you are looking at the RBM. At the end of the day, the actual performance of that teacher is not captured.

The implication was that the validity of the process was questionable. The objectivity of the rating process has also been questioned. Participant 7 (head of department) also commented that, *“... it’s not objective at all (laughs). I have seen non performers being rated better than the higher performers. A person whom you know is a work horse getting a lower rating than someone known to be a dilly dally worker”*. The suggestion made here was that there was some form of favouritism during the rating process. The use of external evaluators or the use of multiple evaluators could be used to avert this problem.

Theme 4: Knowledge of evaluation policy

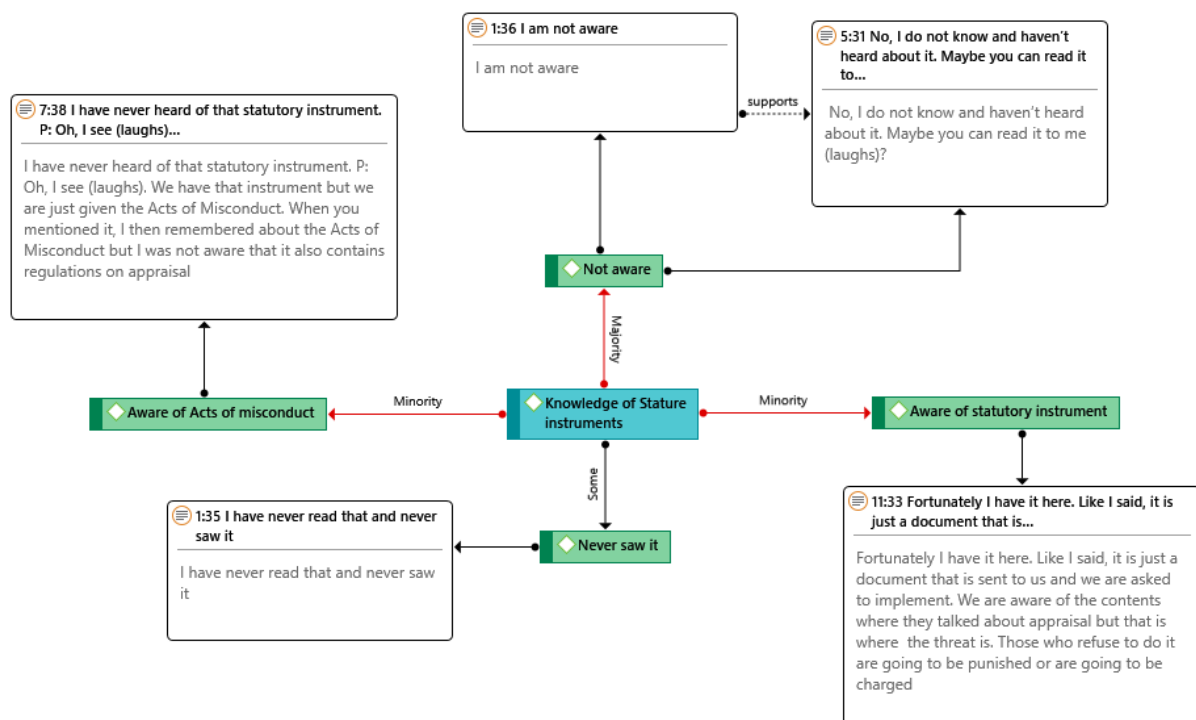


Figure 5:3 Participants' knowledge of the statute regulating appraisal process

Figure 5:3 shows that in terms of knowledge of the appraisal regulations, there were 4 groups of participants, those who were not aware of the existence of the regulations, those who had never seen or read the regulations and yet may have been aware of their existence, those who were only aware of the Acts of Misconduct and yet unaware of the appraisal regulations and those who were aware of the appraisal regulations. The study established that the majority of the participants were not aware of the existence of the statutory regulations on evaluation as shown in Figure 5:3. Remarks from participants like that from Participant 4 bears testimony to this. Participant 4 (teacher) said, *"I am not aware of the appraisal regulations. I may however been made to sign because that is what they normally do but I am really not aware of them"*.

Indications on the diagram also showed that some participants may have been aware of the existence of the regulations and yet they had neither read nor set eyes on them. Participant 1 (teacher) commented that, *"I have never read that and never saw it"*. The diagram further shows that there were a few participants who were aware of the Acts

of misconduct but unaware of the regulations on evaluation although these regulations are all contained in the Public Service Regulations of 2000. To clarify this fact, Participant 7 (head of department) said that, *“We have that instrument but we are just given the Acts of Misconduct. When you mentioned it, I then remembered about the Acts of Misconduct but I was not aware that it also contains regulations on appraisal”*. Very few participants indicated that they were aware of the existence of the regulations on appraisal as explained by Participant 12 (school head) that, *“Fortunately I have it here. Like I said, it is just a document that is sent to us and we are asked to implement. We are aware of the contents where they talked about appraisal...”*

From the indications on Figure 5:13, it is evident that the majority of the participants did not know the regulations on evaluation. This implies that the system could have been introduced without any attempts at marketing the programme since this would have entailed explaining the contents of the evaluation policy. The implication of this is that the implementers were just implementing the policy without actually knowing what it says. This has a lot of insinuations on both the implementers and the drivers. The fact that the majority of the participants including even some administrators were not aware of the existence of the evaluation policy, raised the issue of whether the implementers were even aware of the objectives and purpose, the character and nature, and the consequences of the evaluation exercise. The lack of awareness and lack of curiosity on the part of the teachers to know about the regulations governing some of their work practices raised questions on their professional standing. The situation could also be indicative of an intimidatory and dictatorial working atmosphere where the implementers were treated like a soldier who is simply told by his commander to jump and cannot ask why but only asks how high sir.

Theme 5: RBM implementation challenges

To have a complete understanding of the implementation process, participants were asked to explain any challenges that they faced in implementing RBM. The major challenges according to the participants were the aspect of time constraints, followed by the issue of resources. Participants indicated that there was so much pressure on time which affected their ability to implement the appraisal process effectively. The issue was worsened by double sessions in most of the schools especially the urban schools. The new curriculum had also placed further demands on the already limited

resource. The new curriculum had brought additional subjects and workload because the number of teachers in the schools had not been increased because of the current government employment freeze. One participant explained that, “... *The challenges include issues of time. The challenge of time is worsened by the fact that the school has double sessions and the new curriculum that has many more subjects which has made the timetable very congested*” (Participant 4: teacher). Participant 8 (head of department) further elucidated that, “*We do not have adequate time... That is also the other reason why at times there is no discussion of the work plan and there is no monitoring during the appraisal cycle*”. It was also explained from these remarks how the implementation process was being compromised due to the challenge of time. The challenge of time was reportedly making it difficult for work plans to be discussed as they should and also making it difficult for monitoring and for the supervision of teacher performance thereby making the process doubtful and lose credibility.

Inadequate resources

Material and financial resources were also cited as impediments to the implementation process. Schools were said to be facing a challenge of material resources in the form of text books and various teaching and learning materials which hindered the implementation process. Participant 10 (school head) explained that, “*At school level, the main challenge is the shortage of resources. It would be ideal if each learner could have his/her own textbook for the various subjects which we are unable to do*”. Another participant also added that, “*The other challenge concerns the issue of resources. Pupils do not pay school fees on time so the school is not able to buy the necessary requirements on time*” (Participant 4: teacher). The point being made here was that the learners did not have adequate learning material which in turn affected the ability of the teacher to discharge his/her duties effectively. The shortage of material resources extended to teaching requirements which are often procured late because students delayed in paying their fees. Such challenges resultantly led to a lot of compromises being made as teachers had justifiable explanations to not fulfilling their work plan.

Financial resource constraints also impeded the implementation of RBM in the schools. “*We have a challenge of financial resources that are needed for the effective implementation of the process*” added Participant 12 (school head). Finances were

needed to procure the requisite teaching-learning material, produce evaluation forms and support professional development programmes. It emerged that the relevant Ministry was not providing enough forms so the schools were now taking over this responsibility as they could not pass it to the teachers.

Other challenges that have been cited include:

Inconceivable evaluation form

- It was revealed from the interviews that a substantial proportion of the educators did not know or understand how the evaluation form should be filled and therefore they copied from the few who had some idea. Not knowing or understanding how the implementation should be done as stated by Participant 2 (teacher) that, *“We do not even know how it should be done. Very few teachers know and understand how it should be done so we just consult the few who have an idea... and we copy each other and complete the forms...”*. This is happening in spite of the Public Service Commission having produced the guidelines that were supposed to assist the educators in completing the evaluation form. The form had not been widely availed as only a few participants admitted to having it.

Too large teaching load

- Having to contend with very large classes, Participant 8 (head of department) pointed out that, *“We also have a challenge of classes that are too large”*. Indications from the teachers were that the classes were very large and therefore tended to distort the actual load of the teacher on the ground since a teacher may be having five maths classes officially whereas in fact, this is equivalent to ten classes due to the size of the classes. This resultantly affected the operations of the educator and his/her efficiency.

Excessive clerical work

- Participants highlighted that there is excessive clerical work in drafting the necessary documentation. To explain this, Participant 12 (school head) pointed out that, *“The paperwork is still very bulky and it’s a challenge, who is supposed*

to bear the cost of producing the paper work". The educators complained that the paperwork involved is too much and that it takes much of their time and that it is also costly on the schools since they cannot ask the educators to source their own forms since they should ideally be provided with the forms by the driving ministry.

Inadequate motivation

- Indications by most participants was that they were inadequately motivated because the exercise is insignificant and inconsequential. Participant 8 (head of department) clarified that, *"The main challenge is that people are just demotivated (laughs). It's an exercise in futility. They do not really enjoy filling the forms probably because there are no incentives attached to the process..."*

Administrative contradictions

- Participants proffered that contradicting goals existed when the school administrators were forced to send the students back home for fees which they needed to run the school programmes which on the other hand affected the teacher's teaching programme as students would not be available. This was said to be happening in spite of the government regulation against the sending of students home for fees. One participant explained that,

The challenges which we have is the issue of contradicting goals. The administration has their goals that mainly focus on the collection of and availing resources to be used. Teachers on the other hand have targets that they want to achieve which they fail to accomplish because the learner has been sent home for not being paid up on their fees (Participant 1: teacher)

The educators decried the fact that while the school administrators chased the pupils away from school for lack of fees payment in order to get finances to run the school, it also disrupted the teaching-learning programme. This led to the contradicting goal scenario where the administrator is chasing away the student from school to make the parents pay and on the other hand the teacher wants the student in the classroom to be able to work towards achievement of the set targets.

In summary, the qualitative findings were:

- Work plans and performance targets were discussed and agreed on by both evaluatee and evaluator. Participant 11 (school head) explained that, *“At the beginning of the year, we advise teachers to draft their work plans for the appraisal cycle and then come for discussion”*.
- There was no continuous monitoring of teacher performance. Participant 7 (head of department) asserted that, *“..., there is nothing that actually happens between the beginning and end of the appraisal cycle...”* to explain the absence of continuity of monitoring.
- Rating was neither objective nor valid. Participant 11 (school head) commented that, *“Aaaa, in my view it is subjective because it depends on who is rating you,”* and Participant 1 (teacher) added, *“The process however fails to capture the quality aspect,”*.
- Implementers were not cognisant with the evaluation policy. A comment by one of the participants, *“No, I have never seen the Statutory Instrument or heard about it”* (Participant 6: head of department) illustrated this point.
- The major impediments to implementation were time, resources, large classes and excessive clerical work. Participant 11 (school head) partly explained that, *“The challenges we have include the time factor, resources and our environment which can be prohibitive...”* and Participant 8 added, *“We also have a challenge of classes that are too large”*.

5.7.4 Comparing and contrasting quantitative and qualitative results on the implementation process

The quantitative and qualitative findings on the implementation corroborate and substantiate each other. Both the quantitative and qualitative strands of the study established that performance targets which are ultimately used as the standards for evaluation are not pre-set but rather are discussed and agreed upon by the evaluator and evaluatee. Most of the questionnaire respondents, 82.5% affirmed that performance targets are discussed and agreed to by the evaluator and evaluatee. This was corroborated in the interview that, *“At the beginning of the year, we advise*

teachers to draft their work plans for the appraisal cycle and then come for discussion. I then discuss the Key result areas and targets with each teacher. We agree and where there is a need to adjust we adjust" (Participant 11: school head). The study revealed that the goal-setting process was individualised to allow the teachers to work towards the attainment of the goal that is specific to his/her classroom scenario.

The majority of the questionnaire respondents (57.3%) in the quantitative strand also indicated that the performance evaluation does not really measure teacher performance. This finding was substantiated during the individual interviews that the reality on the ground is that in most cases nothing really takes place between the setting of targets and the rating process. Participant 2 (teacher) confirmed this in the statement that, *"It isn't (laughs) because there isn't supervision. There isn't any RBM in the school, it's just a fallacy. Yaa, nothing is happening on the ground"*. Indications are that the ratings that are given at the end of the appraisal cycle are just cooked up regardless of the reality on the ground since there is no continuous monitoring throughout the appraisal process which therefore makes the whole process doubtful.

The absence of continuous monitoring of performance is confirmed in both the quantitative and qualitative strands of the study. A high number of questionnaire respondents, 68.5% stated that there is no continuous monitoring. Participant 7 (head of department) clarified during the interviews that, *"... there is nothing that actually happens between the beginning and end of the appraisal cycle... There is no review in between as is supposed to be the case"*. The formative function of the process is therefore compromised.

The also study established that the implementers are not cognisant with the evaluation policy as explained during the individual interviews. The majority of the respondents (48%) had indicated in the questionnaire that they had no knowledge on recourse to take in the event of disagreements on the work plan or final rating of which the issues are enunciated in the evaluation policy. 31.8% of the respondents however pointed out that they were not sure. The interview participants admitted that, *"No and I have never seen the Statutory Instrument or heard about it"* (Participant 6: head of department). Observations made are that most of the participants are not aware of the appraisal regulations including some administrators.

Indications by 63.7% of the questionnaire respondents are that the evaluators do not have adequate time to conduct the evaluation activities and the interview participants explained that the time issue is a major impediment to the implementation of appraisal due to the extra time demands made on the implementers to satisfy the requirements of the new curriculum and the double sessions issue in most of the urban schools. Besides the challenge of time, it also emerged from the interviews that the lack of material resources, excessive clerical requirements of RBM and large classes impede the implementation procedure thereby resulting in the giving of ratings that are unsupported in an effort to comply with policy requirement.

5.7.5 Discussion of quantitative and qualitative results on the implementation process

Setting of performance targets

This study established that performance targets that were used as the basis of measurement of teacher performance were not pre-set but rather were discussed between the evaluator and evaluatee at the beginning of each evaluation cycle. The performance targets were the goals of the teaching process and these were set to direct the owner of the goal on what to do and at the same time help to focus the activity in a particular direction (Haefele, 1993:28). The goals which were the performance targets therefore helped to direct and focus the educator in the teaching process. The finding from this study is in line with the propositions of the RADPS conceptual framework which also advocates for the discussion and setting of performance goals by both the evaluator and the evaluatee. This effectively means that the evaluatee is in a way involved in deciding the standards of performance that are used to appraise him or her. The discussion of the targets by the two concerned parties facilitated critical reflection by both evaluator and evaluatee thereby ensuring that they were on the same page from the onset of the evaluation process (Santiago & Benavides, 2009:14). As highlighted in the RADPS conceptual framework, the performance goals have to be informed by the job description and also linked to the institutional goals as was revealed in the study. It was explained in the interviews that the performance targets were based on the evaluatee's job description and also the DIPA which enunciates the institutional goals.

Evaluation is an event

Continuous monitoring of teacher performance is a feedback mechanism aimed at reducing the variance between current performance and the intended goal, (Timperley & Hattie, 2007:86). Continuous monitoring should therefore be taken as an opportunity for the appraisers and appraisees to openly and constructively discuss deliverables and non-deliverables on the goals, thereby realigning the performance to the goals or alternatively, the goals can then be readjusted to conform to any discussed and agreed issues in the feedback discussions. The general lack of continuous monitoring of teachers' performance established in this study may therefore mean that there wasn't any self-regulation or self-correction by the system. It is therefore worth noting that monitoring of teacher performance should be continuous and regular in order for the evaluatees to remain focused on the intended targets. This study established on the contrary that there was basically no continuous monitoring of teacher performance in most participating schools. There was basically no activity between the target setting and rating events. It implied that there was basically one summative rating at the end of the appraisal cycle. This contradicted the assertion by TNTP (2010:8) that an evaluation process should not be limited to a single rating at the end of the evaluation cycle, but rather, cultivate a performance-focused culture by having regular and frequent monitoring and supervision of performance. Such an approach helps to ensure credibility of the whole process and at no point would the summative rating ever be a surprise to either the evaluator or evaluatee who would be having regular contact and keeping track of progress towards the attainment of the set performance targets. Frequent data collection has contributed to the evaluation system of China being regarded as credible even though it was designed in a non-participatory way. This thus helps to emphasise the role of frequent monitoring in enhancing credibility of the evaluation process.

The absence of continuous monitoring revealed by the study is also in contradiction with the propositions of the RADPS conceptual framework. The RADPS conceptual framework proposes that for a system to achieve a professional development function, then, regular, frequent observation and feedback should be given. Feedback that facilitates improvement should be accurate, clear and timely. Formative assessment facilitates the provision of timely, clear and accurate feedback as it is given on the

basis of a task that has just happened as compared to summative feedback. The situation on the ground as revealed by the participants thus contradicts the suggestions of the RADPS conceptual framework. The absence of continuous monitoring indicates that there is no frequent and regular feedback and this can contribute to making the final rating unacceptable thereby discrediting the evaluation process.

Inadequate knowledge of the statutory instrument on evaluation

The results of this study also showed that some implementers were not cognisant of the policy on evaluation. The lack of knowledge on policies is usually associated with poor dissemination of information, poor marketing of programme, low morale and lack of a reading culture in institutional structures and processes. Pretorius and Ngwenya (2008:156) emphasize the point that a programme has to be marketed to the implementers so that they are made aware of its purpose and benefits for them to embrace and accept it. This lack of familiarity with the evaluation policy could be indicative of a system that was hurriedly introduced without adequate marketing of the programme. The RADPS conceptual framework in the *Readiness Assessment component* suggests that demand for the programme should be ascertained. To ascertain demand, this requires that some form of marketing be done. Marketing of RBM would have entailed explaining the details of the programme which are derived from the evaluation policy. This is in line with the assertions by Mackay (2006:6) that low demand can be addressed by increasing awareness of the programme. Conducting readiness assessment would thus have entailed explaining the evaluation policy thereby ensuring that the implementers would be knowledgeable on the policy. The implication of this finding is that readiness assessment may not have been done which is contrary to the propositions of the RADPS conceptual framework.

The evaluation policy normally elucidates on issues such as the purpose and objectives of evaluation, its nature and procedure and the consequences of the process (Tucker & DeSander, 2011:5). Lack of familiarity of such critical issues may affect key issues such as the buy-in and understanding of the process thereby ultimately negatively affecting the implementation process like in this case. Heneman et al. (2006:6) proffer that for an evaluation system to survive and contribute to performance improvement, it must be accepted by those who use it, which is possible

only if people know and understand the system. Familiarity with the evaluation policy assists and contributes to an understanding of the system which ultimately leads to the system being appreciated. This study has however established that this was missing in the case of the implementers in Kwekwe district.

Lack of objectivity and validity of rating process

This study established that educators did not regard the rating process as objective and valid. This was attributed partly to the bunching of teachers. Bunching was done with the aim of deliberately avoiding giving either extremely low or high ratings that would need to be validated by external evaluators. The lack of monitoring and supervision of performance between the target setting and rating events was the other reason why objectivity was questioned as the rating could not be supported. The validity of the instrument was questioned because it was blamed of focusing more on the outputs at the expense of the outcomes.

An evaluation system should accurately evaluate teachers so that the ratings reflect the pattern of students' academic development and growth (TNTP, 2010:9). The fact that the evaluation process focused more on the outputs, which are the lessons and the activities given to the learners, at the expense of the learning outcomes shown through the student academic development and growth is contrary to the assertion by the TNTP. This concern from the teachers implied that the process may not be valid and this may be attributable to the exclusion of the teachers as the experts who are supposed to know how to measure effective teaching as highlighted by Goe et al. (2014:14). The concern by the teachers is in line with the assertion by Haefele (1993:24) that an evaluation instrument can have measurement deficiency. Measurement deficiency refers to the extent to which the evaluation instrument excludes some performance criteria considered important and representative of teaching. In this case, the argument advanced by the educators was that, the evaluation instrument has a measurement deficiency as it excluded student achievement as a measure of teacher performance. The educators on the other hand felt that students' results are a reliable and effective measure of teacher performance. This resonates with the view of both the teachers and administrators in China who regarded the students' results as being more factual and meaningful relative to the other measures of teacher performance (Zhang & Ng, 2017: 206).

The insufficient activity between the target setting and final rating exercises meant that there was no continuous monitoring that should generate information that is used as the basis of the rating process. The absence of continuous monitoring that was revealed by the participants is contrary to what is suggested by the RADPS conceptual framework that advocates for formative assessment as a way of enhancing acceptability of the summative rating and also to improve instructional practice. One purpose of feedback according to Hattie and Timperley (2007:86) is to lessen the discrepancy between the current performance and targeted performance. Giving feedback at the end of the appraisal cycle will therefore result in failure to fulfil this purpose. It is further argued that for feedback to be effective, it has to provide precise information of the task at hand (Hattie & Timperley, 2007:96). The feedback in this case will not be precise and targeted since it is summative and not directed to a specific task. Rating is a form of feedback since it is supposed to give information concerning an individual's performance. In this event where rating is not evidence-based then it fails to be objective and valid. Feedback is supposed to create an opportunity for the evaluator and evaluatee to openly and constructively discuss instructional practice to assist teachers to improve their competencies rather than being compliant-driven (TNTP, 2010:8).

Literature control on challenges to implementation

This study established that the implementation process was hound with a myriad of challenges. One of the challenges to the implementation process was the issue of time. Both evaluators and evaluatees claimed to have challenges of time to implement the evaluation procedure properly. Implementers complained that the new curriculum has brought in added work load further eroding the time which they were finding difficult to apportion due to an already overwhelming workload. In support of this finding, Derrington and Campbell (2015:312) postulate that time for the various evaluation procedures like observations, feedback and requisite developmental work may not be available since the evaluation responsibilities often add to a work load that is often already overwhelming. White et al. (2012:8) note that time demands could be improved by investing in the technology to reduce time and paperwork thereby freeing time for observations and provision of instructional leadership by the evaluators to the evaluatees.

The lack of material and financial resources was also cited as an impediment to the implementation process. This finding is stated by Mavhiki et al. (2013:138) when they proffer that the implementation of RBM across the various sectors in the civil service of Zimbabwe has been hampered by financial, human and physical resources. The fiscal challenges could also have affected other things like the provision of the initial training and the capacity of the driving ministry to impart the requisite skills. This position is confirmed by Machingambi (2013:223) who also made a similar finding that inadequate training had been given to the implementers due to constrained financial resources thereby compromising their ability to implement RBM.

The other impediment established by this study was that of excessive clerical work. This finding is authenticated by a similar finding by Gutuza (2016:394) who also established that teachers complained of RBM having too much paperwork which resultantly ate a lot into the time which could otherwise have been used for preparation of teaching and other related activities.

5.8 Objective 4: To Assess the significance of the teacher performance evaluation process

The study sought to establish the significance of the teacher performance evaluation system of Kwekwe district in promoting accountability and professional development.

5.8.1 Quantitative Results on significance of evaluation process

Questions in table 5:12 were used to get the opinions of the teachers on the significance of the evaluation process.

Table 5:12 Perception of teachers on the significance of evaluation process

Significance of the appraisal process	Frequency/Percentage				
	S D	D	NS	A	S A
The feedback on my appraisal helps me to improve my teaching	145 49.7%	52 17.8%	19 6.5%	33 11.3%	43 14.7%
Appraisal results inform decisions on promotion.	82 28.1%	60 20.5%	87 29.8%	50 17.1%	13 4.5%
Appraisal results inform decisions on regrading.	52 17.8%	59 20.2%	101 34.6%	64 21.9%	16 5.5%
Appraisal results inform decisions on whether one continues to be on probation.	57 19.5%	63 21.6%	117 40.1%	51 17.5%	4 1.4%
Appraisal results inform decisions on termination.	60 20.5%	62 21.2%	131 44.9%	34 11.6%	5 1.7%
Outstanding teachers are rewarded through teacher appraisal.	80 27.4%	79 27.1%	80 27.4%	46 15.8%	7 2.4%
Teachers at my school who perform poorly are mentored by more experienced and well-performing teachers.	79 27.1%	81 27.7%	61 20.9%	62 21.2%	9 3.1%
Appraised teachers that continue to perform poorly in spite of remediation have their employment terminated.	89 30.5%	105 36%	77 26.4%	17 5.8%	4 1.4%
Appraisal results help to decide what is covered in our school staff development programmes	106 36.3%	63 21.6%	71 24.3%	42 14.4%	10 3.4%
We are given time for collaborations as teachers to help us to improve performance.	120 41.1%	72 24.7%	42 14.4%	53 18.2%	5 1.7%
We are given manpower development support if appraisal results require training that cannot be provided within the school	67 22.9%	101 34.6%	51 17.5%	69 23.6%	4 1.4%
Teachers at my school are given a chance to say their contribution on how the performance appraisal system can be improved	70 24%	81 27.7%	44 15.1%	85 29.1%	12 4.1%

It has been noted that in this part of the questionnaire, a significantly high number of respondents indicated that they were not sure of most of the issues that were asked here. Cases of respondents admitting that they were not sure went as high as 44.9% and in most cases such respondents were well over 20% in most of the items asked. The majority of teachers, 197 (67.5%) thinks that feedback on appraisal does not help to improve the science and art of teaching for the teacher. Many teachers also felt that promotions were not based on appraisal results, 142 (48.6%) and this may demotivate and demoralise deserving teachers as well as those aspiring for higher positions. The decision for continuation on probation is also not dependent on a teacher's appraisal performance meaning that there may be no performance related targets to the contract of employment from probation to permanent employment or it can mean that those charged with supervising the probation workers neglect their duties. This could be the reason why many teachers, 122 (41.8%) thought that the results of the appraisal did not inform decisions on termination.

Teachers believed that outstanding performance does not receive any reward or recognition as shown by the figures 159 (54.5%), and at the same time 160 (54.8%) of

the teachers indicated that those who perform poorly are not likely to be mentored by more experienced and well-performing teachers. It was also observed that as many as 194 (66.4%) of the teachers appreciated that there were no sanctions to appraised teachers that continue to perform poorly in performance evaluation in spite of remediation. A greater number, 192 (65.8%) of the teachers, stated that there is no time given for collaborations for teachers to help each other to improve their performance. It was also evident that 168 (57.5%) of the teachers believed that further work related development which cannot be provided within the school was not supported.

In summary, it can be drawn that:

- Most teachers (67.5%), stated that feedback on appraisal was not helpful.
- Indications by 48.6% of the teachers were that promotion of a teacher is not dependent on his or her work related performance results. There was however a significant proportion of teachers, 29.8%, who said that they were not sure if promotion was performance based. It was also indicated that performance did not determine whether one continues to be on probation or not by 41.1% of the teachers while 40.1% said they were not sure.
- There was no reward for outstanding performance as confirmed by 54.5% of the teachers although 27.4% of the teachers admitted that they were not sure. Just as there were no rewards for outstanding performance, indications by 66.5% of the teachers were that there were no sanctions for continuous underperformance and on the other hand, 26.4% of the teachers were not sure if there were any sanctions.
- Many teachers, 54.8%, argued that poor performers were not mentored while 20.9% of the teachers claimed that they were not sure.
- It was confirmed by 57.5% of the teachers that those identifiable as needing further work related development were not given any support.
- According to 57.9% of the teachers' appraisal results did not inform staff development programmes. A significant section of the teachers, 24.3%, however indicated that they were not sure if staff development programmes were informed by the appraisal results.
- Most teachers, 65.8%, pointed out that there was no time for collaborations among teachers to help them improve.

5.8.2 Views of urban and rural teachers on significance of evaluation process

Table 5:13 Views of teachers urban and rural teachers on the significance of the evaluation process

Significance of the evaluation process		SD	D	NS	A	SA	Total
The feedback on my appraisal helps me to improve my teaching	Urban	86 45.5%	38 20.1%	10 5.3%	27 14.3%	28 14.8%	189
	Rural	59 57.3%	14 13.6%	9 8.7%	6 5.8%	15 14.6%	103
Appraisal results inform decisions on promotion.	Urban	55 29.1%	41 21.7%	57 30.2%	31 16.4%	5 2.6%	189
	Rural	27 26.2%	19 18.4%	30 29.1%	19 18.4%	8 7.8%	103
Appraisal results inform decisions on regrading.	Urban	35 18.5%	40 21.2%	66 34.9%	42 22.2%	6 3.2%	189
	Rural	17 16.5%	19 18.4%	35 34%	22 21.4%	10 9.7%	103
Appraisal results inform decisions on whether one continues to work on probation.	Urban	41 21.7%	42 22.2%	70 37%	33 17.5%	3 1.6%	189
	Rural	16 15.5%	21 20.4%	47 45.6%	18 17.5%	1 1%	103
Appraisal results inform decisions on termination.	Urban	43 22.8%	38 20.1%	82 43.4%	22 11.6%	4 2.1%	189
	Rural	17 16.5%	24 23.3%	49 47.6%	12 11.7%	1 1%	103
Outstanding teachers are rewarded through teacher appraisal.	Urban	58 30.7%	48 25.4%	56 29.6%	23 12.2%	4 2.1%	189
	Rural	22 21.4%	31 30.1%	24 23.3%	23 22.3%	3 2.9%	103
Teachers at my school who perform poorly are mentored by more experienced and well-performing teachers.	Urban	49 25.9%	48 25.4%	43 22.8%	44 23.3%	5 2.6%	189
	Rural	30 29.1%	33 32%	18 17.5%	18 17.5%	4 3.9%	103
Appraised teachers that continue to perform poorly in spite of remediation have their employment terminated.	Urban	62 32.8%	63 33.3%	51 27%	11 5.8%	2 1.1%	189
	Rural	27 26.2%	42 40.8%	26 25.2%	6 5.8%	2 1.9%	103
Appraisal results help to decide what is covered in our school's development programmes	Urban	60 31.7%	43 23.3%	48 25.4%	31 16.4%	7 3.7%	189
	Rural	46 44.7%	20 19.4%	23 22.3%	11 10.7%	3 2.9%	103
We are given time for collaborations as teachers to help us to improve performance.	Urban	76 40.2%	51 27%	22 11.6%	37 19.6%	3 1.6%	189
	Rural	44 42.7%	21 20.4%	20 19.4%	16 15.5%	2 1.9%	103
We are given manpower development support if appraisal results require training that cannot be provided within the school	Urban	43 22.8%	66 34.9%	32 16.9%	45 23.8%	3 1.6%	189
	Rural	24 23.3%	35 34%	19 18.4%	24 23.3%	1 1%	103
Teachers at my school are given a chance to say their contributions on how the performance appraisal system can be improved	Urban	46 24.3%	58 30.7%	30 15.9%	48 25.4%	7 3.7%	189
	Rural	24 23.3%	23 22.3%	14 13.6%	37 35.9%	5 4.9%	103

The results of the perceptions of the teachers on the significance of the appraisal process based on location are indicated in Table 5:13. A significant number of the teachers, 65.6% of the urban compared to 70.9% of the rural teachers highlighted that feedback on appraisal was not helpful. It was however noted that a substantial number of teachers indicated that they were not sure of most of the issues pertaining to the aspect of significance of the evaluation process. There was an abnormally high number of respondents that indicated that they were not sure on most aspects. Regarding the use of appraisal results to inform promotion, 50.8% of urban teachers against 44.6% of the rural teachers disagreed that promotion was made on the basis of evaluation findings. A substantial proportion of the teachers, 30.2% urban and 29.1% rural teachers nevertheless indicated that they were not sure if there was a link between promotion and appraisal results. On the aspect of rewards for outstanding teachers, 56.1% of the urban compared to 51.5% of the rural teachers refuted that there were any rewards for teachers that have outstanding performance. There were however 29.6% urban and 23.3% rural who claimed that they were not sure of the issue of rewards. Regarding termination of employment for teachers that continue to perform poorly in spite of remediation, 66.1% of the urban relative to 67% of the rural teachers disproved that this happened. A greater proportion of the rural teachers (61.1%) versus 51.3% urban teachers negated that teachers who perform poorly were mentored.

Concerning the provision of manpower development support, 57.7% urban compared to 57.3% rural teachers disputed that manpower development support was availed to the teachers who needed it. A significant proportion of teachers in the urban (67.2%) against 63.1% of the rural teachers refuted that teachers were given time for collaborations to help them improve their instructional practice. A greater number of the rural teachers (64.1%) compared to 55% of the urban teachers denied that appraisal findings informed staff development programmes in the schools. A closer examination of the results shows that the variations in percentages in most cases is around 5% or less which indicate a similarity in the opinions of the teachers in the two geographical locations. Both urban and rural teachers showed that they perceived the appraisal process to be insignificant.

5.8.3 Qualitative results on significance of appraisal process

An evaluation process should have meaningful implications for it to earn support from the implementers (TNTP, 2010:8). This study therefore sought to establish the opinions of the implementers on the significance of the appraisal process. A number of views emerged from the interviews and these basically pointed to the fact that the evaluation process was meaningless and perfunctory since it was inconsequential and thus the implementers had basically no respect for the process.

Theme 1: Evaluation process does not promote accountability

It emerged from the interviews that the majority of the participants felt that the process was inconsequential. The participants stipulated that the whole process was just a facade which in reality was meaningless since whether one got an outstanding or dismal rating, nothing really mattered. The following remarks by one of the participants demonstrated this feeling:

That is what I said earlier on that I haven't seen any tangible results from this system. One getting a 5 and excelling and another one getting a 1, it's the same thing because nothing happens to either of the two. I have never seen anyone going for retraining although I have seen people who deserve to be sent for retraining due to their level of performance. I have never seen anyone being charged because of incompetence or negligence of duty or anyone being rewarded for outstanding performance either. So we just think it's an instrument meant to while up time because no one is being rewarded and no one is being assisted to improve (Participant 3: teacher)

The sentiment by the participant showed that the process had lost respect from the participants. It was regarded as meaningless as it appeared not to serve a purpose either of holding the teachers to account or even assisting them to improve. Another participant reiterated that,

Nothing of that sort has happened because like I mentioned earlier on when we started we were concerned about the rating that we would get because we thought that something would come out of the process. We thought that probably we were going to be remunerated or to be called in for an explanation if one performed poorly but nothing actually happened. It's just a matter of filling in the forms and nothing happens thereafter. No one has ever been called for a warning or even counselling. No one has ever been remunerated for good performance either? All we do is to make sure that someone gets a mark around 4 so that they won't be a need for explanations to the external bosses. The fact that nothing really happens in

spite of the rating that one gets has made this process ineffective and lose any semblance of respect that I had (Participant 2: teacher).

It was revealed that when RBM was introduced, educators initially cared about the rating they would be awarded as they believed that it would have consequences. Indications are that participants were expectant that the process would have either positive or negative repercussions relative to one's level of performance. On realising that the evaluation process had no outcome, the participants admitted that they had lost any respect of the process that they had. Participant 7 (head of department) also weighed in saying,

Nothing, whether you achieve very high or you do not achieve at all, there are no consequences. There are no rewards, there is nothing. It has a very negative impact on the whole process because as long as I know that nothing will happen in spite of my performance, then, I won't have respect for the process at all. The fact that there is nothing that is attached to it makes it useless, actually. When the system was introduced, we were told that we would be remunerated based on the results of the appraisal process.

The above sentiments by the participants were evident of the fact that the whole process had completely lost respect from the implementers and according to the participants, the loss of respect emanated from the fact that the process is inconsequential and is void of any repercussions for the non-performers or rewards of any form for the outstanding performers.

Theme 2: Evaluation process relativity with professional development activities

It also emerged from some participants that the appraisal process was linked to professional development activities. This meant that the evaluation results were used to inform professional development programmes targeted at enhancing and improving teacher performance. In support of this observation, Participant 8 (head of department) said that, *"... some teachers are not teacher-trained so they do not know some teaching methodologies and issues like planning etc, so staff development is informed by what came out of the appraisal process..."*. To further clarify on the link between the evaluation process and professional development activities, Participant 11 (school head) added, *"We can however say appraisal informs staff development in a way. During a lesson or exercise book supervision you can note a common trend then*

organise for staff development in that area". Participant 12 (school head) also weighed in supporting the existence of the link. He stated that, *"... if I say I am not good at financial management and one or two other heads also make similar indications, the district office will pick it and that is how they organise workshops and invite resource persons to come and facilitate"*.

Indications were that in some schools there was a link between the evaluation process and the professional development programmes. The observation made was that there was no uniformity in the way that RBM was implemented across the schools which necessitated a monitoring mechanism to be put in place for standardisation.

Some participants however argued that the appraisal process was not linked to professional development programmes. Participant 9 (school head) pointed out that, *"No, not even a single day did we sit down and say let's look at our KRA, let's revisit those areas that we feel need improvement. We are not using it to be very honest. So I think, yes as a recommendation, this is supposed to be done"*. The sentiments of this participant suggest that although there is no link currently, such a link would however be appreciated. Participant 2 (teacher) also supported the absence of a link between the professional development programmes and the evaluation process by highlighting that, *"We haven't had any workshops on issues that arise from the appraisal process since it's just a fallacy..."* Indications by this participant is that the link is absent as the appraisal process is none existent in reality.

In summary, the following issues on significance of the evaluation process emerged from the individual interviews:

- The evaluation process was inconsequential and did not promote accountability since there were neither rewards nor sanctions given on the basis of performance. A remark by one of the participants illustrated this, *"Nothing, whether you achieve very high or you do not achieve at all, there are no consequences. There are no rewards, there is nothing. It has very negative impact on the whole process..."* (Participant 7: head of department)
- There was a bit of contradiction with some participants indicating that evaluation results inform professional development programmes as indicated by Participant 11 (school head) that, *"We can however say appraisal informs staff*

development in a way” while on the other hand some participants also argued that there was no connection whatsoever between the two as expressed by Participant 3 (teacher) that, “... I haven’t seen that link or connection... staff development is only done if there is something from the district or regional office”. It was however noted that this variation was across and not within a school which could indicate absence of a standardised approach to the process.

5.8.4 Comparing and contrasting quantitative and qualitative results on significance of appraisal process

The quantitative and qualitative findings on the significance of the evaluation process corroborated and validated each other. All two strands of the study established that the process is meaningless and unhelpful. Feedback is neither helpful nor likely to be used since the system has lost respect from the implementers. The majority of the questionnaire respondents, 67.5%, indicated that feedback does not really help them to improve their teaching practice. This finding was confirmed during the individual interviews that the reality was that in most cases nothing was happening on the ground so there was actually no feedback as baseless ratings were often given to the appraisee. This is illustrated in the following quotation from Participant 7 (head of department), *“... there is nothing that actually happens between the beginning and end of the appraisal cycle. It’s usual teaching... and if things have gone wrong during the cycle, then they will only be seen at the end when it’s too late to do anything”*. Indications are that the implementers have been so disillusioned that they simply implement the procedure just to fulfil a statutory obligation by simply completing the forms for submission without any monitoring and supervision of their performance being done.

Indications were that the implementers have since lost the respect that they had for the process because of its inconsequentiality and insignificance. Indications from both the quantitative and qualitative strand showed that the general feeling was that the evaluation process was simply a perfunctory and ritualistic exercise since there were neither rewards nor sanctions attached to the outcome of the process. From the quantitative strand, 54.5% of the respondents indicated that there was no reward for outstanding performance while 66.5% of the respondents also pointed out that there were no sanctions for teachers who continuously under-performed. This was

confirmed by sentiments from Participant 6 (head of department) who said, *“Teachers do not respect the RBM because there are no consequences. Whether you get a 5 or you get nothing, nothing is done to you so it doesn’t have respect as far as teachers are concerned...”* Both strands of the study established that there were neither rewards nor incentives given for outstanding performance or any sanctions on the other hand for poor performance. Implementers were just complying with policy otherwise there was no motivation to conduct the exercise.

Indications from both the quantitative and qualitative strands of the study were that no manpower development support was given. A significant number of the questionnaire respondents, 57.5%, affirmed that there was no manpower development support to those who need it. It emerged from the interviews that even though a developmental need might be identified through the evaluation process, the challenge was that some of these identified developmental needs could not be addressed within the institution so an individual would need to get that from other institutions which was only possible if one was given manpower development support in the form of paid manpower development leave. Indications were that teachers no longer went for manpower development programmes because they could not afford to do so since it was no longer supported. Participant 10 (school head) clarified that, *“Yea, there is no manpower development support... Many people have families and they can’t go to upgrade themselves without a salary... with the current levels of teachers’ salaries, it is not easy to spare anything towards that”*.

Majority of questionnaire respondents, 68.5%, indicated that there was no time that is given for collaborative activities which can help the teachers to improve their teaching practice. There was however a contradiction between the participants with some saying that the appraisal process helps to inform the professional development programmes within a school while on the other hand other participants disputed that there is such a link. In support of the claim that a link between evaluation and professional development processes exist, Participant 11 (school head) commented that, *“We can however say appraisal informs staff development in a way. During a lesson or exercise book supervision you can note a common trend then organise for staff development in that area.”* Participant 2 (teacher) however contradicted this by saying, *“We haven’t had any workshops on issues that arise from the appraisal*

process since it's just a fallacy...". The majority of the questionnaire respondents, 57.9%, disagreed that there was some link between the evaluation process and the professional development programmes that were conducted in the schools. It was however clear that this variation is across and not within the schools which could be indicative of the absence of a standardised operation procedure as observed earlier on. It would appear that the schools do not really have a standard way of implementing the appraisal process. This can be attributed to lack of adequate training leading to people having a different understanding and knowledge of how it was to be done or it could also be a result of lack of monitoring by the supervisors who should be in a position to identify and address these variations if proper monitoring was done. The findings from the two strands however generally corroborated and validated each other.

5.8.5 Discussion of quantitative and qualitative results on the significance of the process

For an evaluation process to earn support and respect from the implementers, it has to have meaningful implications (TNTP: 2010:8). An evaluation process with meaningful implications will be driven towards the fulfilment of the two major functions of evaluation which are the improvement and accountability functions.

Significance of appraisal on promoting professional development

The improvement function entails ascertaining the strengths and weaknesses of a teacher as well as the shortcomings of the teaching processes to design responsive professional development programmes (Hinchey, 2010:6; Santiago & Benavides, 2009:8). Once these are established, strengths are enforced while weaknesses and shortcomings are corrected through professional development planning as organised responses to the evaluation process (Santiago and Benavides, 2009:19). An evaluation system with a well-designed and professional development program aligned to evaluation findings will therefore enhance teacher quality which ultimately raises student achievement and growth (Looney, 2011:440).

One of the objectives of this study was to assess the significance of the evaluation system of Kwekwe district in promoting professional development as well as accountability. This study found out that there was a bit of controversy on the ability of

the system to promote professional development. There were some implementers who felt that the evaluation system link to professional development was one of its key strengths while on the other hand others argue that no such link existed. The majority of the questionnaire respondents (57.9%), refuted that a link existed between appraisal and professional development activities and this was supported by some participants from the interviews while on the other hand more participants from the interviews and 17.8% of the questionnaire respondents agreed that a link existed. Smylie (2014:98) postulates that professional development can be viewed from two perspectives, the 'weak link problem' and the 'weak quality problem'. What was evident from the study was that while the 'weak link problem' may exist, what is not very clear however is if the 'weak quality problem' was addressed by having relevant and good programmes that facilitated improvement of teacher instructional practice. The RADPS conceptual framework proposes that there should be a strong link between the evaluation and professional development processes. Professional development programmes at both school and individual levels have to be informed by evaluation findings in cases where a strong link exists. This proposition by the RADPS conceptual framework echoes Goe et al. (2014:43) who assert that the most significant phase of the appraisal cycle is the stage at which evaluation results are used to inform and support professional development. Further studies could be conducted to ascertain the quality of the professional development programmes that are provided where such links have been claimed to exist.

Significance of appraisal on promoting accountability

In terms of the accountability function, the study observed that the implementers have generally lost the respect and motivation that they had for the evaluation system as they have come to the realisation that the evaluation process has no consequences. Implementers believed that the process was simply a routine compliance exercise which has neither benefits nor sanctions in spite of the summative rating that one gets. Zvavahera (2014:6) also made a similar finding that the employees did not appreciate the system because of its lack of concrete benefits. These findings from the two studies contradicts the suggestions of the RADPS conceptual framework which advocates for a connection between evaluation findings and personnel decisions. The RADPS framework proposes that evaluation findings should inform decisions on

promotion, type of tenure given to teacher, performance related awards and sanctions for the under-performers. This is in line with assertions by other researchers that evaluation results should inform personnel decisions such as career advancement, rewards or sanctions (Danielson, 2011:36; Isoré, 2009:15; Santiago & Benavides, 2009:19).

It also emerged from the study that there was no meaningful relationship between the actions listed in the Public Service Regulations (2000:12-13) as possible consequences of the appraisal process with what is actually happening on the ground. It emerged from both strands of the study that the appraisal process was inconsequential since it is linked to neither rewards nor sanctions. Indications were that, since the introduction of this performance evaluation system, the rare occasions that salary increment had been given, it had not been linked to RBM rating but had been given to all irrespective of the summative ratings. Promotions and advancements were not being done transparently with a clear connection to the RBM rating as should be done as stated in the appraisal policy. This has resultantly made the implementers disillusioned, disinterested and indifferent to the evaluation process. In support of this observation, Pham et al. (2017:3) asserts that, incentives like merit pay, motivate teachers to improve their performance which in turn improves the performance of the learner due to an improved learning experience. The issue of lack of incentives being an impediment to the implementation of RBM has also been confirmed by Mavhiki et al. (2013:137). A study conducted in India also established that linking a part of the teacher's pay to objective measures of performance is popular with the teachers and likely to result in improved student learning (Muralidharan & Sundararaman (2011:402). The popularity of linking performance to pay has also been observed in this study. Zvavahera (2014:7) further cautions that there is no likelihood of improved service delivery in the absence of performance related reward.

5.9 General opinions on RBM

The study sought to find out the overall recommendations of the educators on RBM as a teacher performance evaluation system.

5.9.1 Quantitative results on teachers' general opinion on the evaluation system

The perceptions of teachers on what is to be done with RBM was established through their responses in Table 5:14.

Table 5:14 Perceptions of teachers on what is to be done with RBM

Recommendations	Frequency (Percentage)				
	S D	D	NS	A	S A
RBM needs to be improved to suit the Ministry of Education environment.	22 7.5%	13 4.5%	19 6.5%	105 36%	133 45.5%
RBM should be removed as a performance appraisal system for teachers.	33 11,3%	41 14%	27 9.2%	65 22.3%	126 43.2%

The majority of the teachers, 238 (81.5%) recommended that RBM should be improved to suit the Zimbabwean education sector and at the same time there was also popular call by 191 (65.5%) that the RBM system should be disbanded as a performance appraisal system for teachers. The high proportion of the teachers calling for the review of RBM to enhance its relevance to education shows that the implementers currently have no confidence in the system's relevance and ability to fulfil its purpose. Teachers were also calling for the removal of the system because a greater section of them felt that the system was more suitable for the industrial rather than the services sector like education.

5.9.2 Comparison of quantitative results on teachers' general opinion on the evaluation process

The comparative perceptions of urban and rural teachers on what is to be done with RBM as a teacher performance evaluation system are presented in Table 5:16.

Table 5:15 Comparison of quantitative results on teachers' general opinion on the evaluation process based on location

Recommendation		SD	D	NS	A	SA	Total
Recommendations RBM needs to be improved	Urban	16 8.5%	8 4.2%	11 5.8%	61 32.3%	93 49.2%	189
	Rural	6 5.8%	5 4.9%	8 7.8%	44 42.7%	40 38.8%	103
Recommendations RBM should be removed	Urban	19 10.1%	20 10.6%	21 11.1%	39 20.6%	90 47.6%	189
	Rural	14 13.6%	21 20.4%	6 5.8%	26 25.2%	36 35%	103

It was evident that teachers in both the urban (81.5%) and rural (81.5%) schools under this study strongly recommended a variation of RBM to improve it so that it becomes relevant to their situation. The majority of urban (68.2%) and rural (60.2%) suggested that RBM should be removed as a teacher appraisal system. It is evident from these results that teachers in both urban and rural schools have generally similar sentiments on RBM as there is no marked variation in the proportion of the teachers from the two locations with comparable perceptions.

5.9.3 Qualitative results on educators' general opinion on an evaluation system

A number of suggestions were made by the participants on RBM as an appraisal system. The suggestions made included mainly reviewing the system or removing it completely.

Theme 1: Improving the evaluation system by incentivising it

It emerged during the individual interviews that the majority of the participants felt that the performance evaluation system needs to be improved by reviewing it. The most popular suggestion from the participants was that the system should be improved by incentivising it so that it regains respect from the implementers. Participant 5 (head of department) explained that, *"The system needs to be improved somehow. Like the issue of incentives, it needs to be addressed so that teachers can respect the system"*. Participant 8 (head of department) added that, *"... maybe if they had a way of motivating teachers... but as it is, the system lacks impact and people have no respect for it and just take it as something just done for compliance. There is a need to incentivise the system"*.

The remarks above emphasize the need to make the evaluation system consequential by incentivising it to restore its respectability from the implementers. The impression given was that the lack of rewards and sanctions was the major challenge why the system is disrespected and that reviewing the system in this area would restore the system respectability among the implementers and make it more acceptable.

Theme 2: Improving the evaluation system by reviewing the evaluation instrument

It also emerged from the interviews that there was a need to review the evaluation instrument itself. The following remark by Participant 12 (school head) illustrated this:

Yes, my comment is that the instrument is not a prescription. It needs improvement regularly. Officials could get the stakeholders, say every one or two years, to suggest any areas of possible improvement on the instrument and review the instrument accordingly. I think that would help to improve the instrument into something that is acceptable to everyone and also enhance its appropriateness.

Similar to the observation by Participant 12 (school head), Participant 1 (teacher) also added that, *“The instrument has to be improved so that it can achieve its purpose... The appraisal process would not be important at all if it is not reviewed. There is a need to involve the teachers because we have not accepted...”*

It is evident from the above sentiments that there is a feeling that the evaluation instrument needs to be revised. The system was said to be inappropriate for the services sector and more suited to the industrial sector. Participant 7 (head of department) clarified through the statement that, *“I also feel that it is more suited to the manufacturing industrial sector than it is to the service industry like education. It is more suited to the manufacturing sector because those have tangible targets...”*. It was also revealed that the review of the system should involve stakeholders to promote ownership and acceptance by the implementers.

Theme 3: Removal of RBM as a teacher evaluation system

Some participants suggested that the RBM should just be removed. They argued that the system had not brought any benefits or improvement to the way that people operate since its introduction. The remarks below illustrated this point:

I think it should be stopped forthwith (laughs), and maybe we just continue the way that we have been functioning before. All these things were already being supervised before RBM came in. I do not think that it is necessary at all in the schools. It wasn't in the schools when I joined service and we have always been working hard because we have set targets in the form of the departmental policies (Participant 2: teacher).

The implication by the above statement is that RBM has not brought any improvement to what was existing before it was introduced. It therefore makes no sense to the participants to add a load which simply burdens the implementers without having any positive impact.

5.9.4 Comparing and contrasting of quantitative and qualitative results on what should be done with RBM

It emerged from both the quantitative and qualitative strands of the study that the majority of the educators were of the opinion that that RBM needs to be improved by making the system consequential. The majority of the questionnaire respondents (81.5%) indicated that RBM has to be improved. The participants in the interviews confirmed this by clarifying that the system can be improved by incentivising it. Participant 5 (head of department) explained that, *“The system needs to be improved somehow. Like the issue of incentives, it needs to be addressed so that teachers can respect the system”*. From the sentiments of the participants, making the evaluation process consequential is one key way of improving it and making it regain the respectability that it lost.

The educators also felt that the instrument needed to be customised to the education sector as there was a general view that the system is currently more suited to the industrial sector than it is to the services sector. Participant 12 (school head) elucidated that, *“... It needs improvement regularly. Officials could get the stakeholders, ... to suggest any areas of possible improvement on the instrument and review the instrument accordingly ... the review process should be participatory so that stakeholder ownership can be promoted”*. The view of the participants was that the instrument should be reviewed to make it more relevant and appropriate for the education sector. The process should however be participatory so that the input of the stakeholders is incorporated.

Results from the quantitative and qualitative strands also indicated that some educators felt that RBM as an appraisal system should be removed. A total of 65.5% of the questionnaire respondents are of the opinion that RBM should be removed as a teacher appraisal system. The participants explained that they did not see the benefits of the system as nothing had changed in terms of the way they operated before the introduction of the system and the way they operate now except that they

now have an added work load brought about by the evaluation responsibilities. To clarify the point, Participant 2 (teacher) said, “... *I do not think that it is necessary at all in the schools. It wasn't in the schools when I joined service and we have always been working hard because we have set targets in the form of the departmental policies*”. The majority were of the view that if the system cannot be improved, then it should be removed completely. Opinions from the quantitative strands were thus corroborated and explained by sentiments from the individual interviews.

5.9.5 Discussion of quantitative and qualitative results on recommendations on RBM

Indications from the educators were that RBM should be improved as a teacher performance evaluation system. The educators felt that RBM is not really appropriate for the services sector, a position which is also supported by Gutuza (2016:395), Machingambi (2013:224) and Mavhiki et al. (2013:138). Educators felt that their involvement could have enhanced the relevance and appropriateness of the evaluation system. This was supported by the assertion that for a system to be viewed as useful, responsive and fair and to be respected and embraced by the stakeholders, it has to be developed in a participatory way rather than being imposed (Goe et al., 2014:14; Machingambi, 2013:224; Pretorius & Ngwenya, 2008:162; Stronge & Tucker, 1999:353). Danielson and McGreal (2000:66) also weigh in, by positing that the relevance and validity of an evaluation system can only be certified by the direct users which thus makes stakeholder involvement from the drawing board imperative. The finding of this study is in line with the proposition of the RADPS conceptual framework that advocates for a continual monitoring of the evaluation system to facilitate its review in line with the inputs from the stakeholders. The RADPS conceptual framework proposes stakeholder involvement as a way of enhancing validity and relevance of the system. It also emerged from the study that the system should be improved by making it consequential. The system can be made consequential by incentivising it and sanctioning under-performing teachers to make it more respectable. This suggestion resonates with one given by Gutuza (2016:395) and Madhekeni (2012:126) that performance information should be utilised for the process to be respectable. The RADPS conceptual framework also proposes that an evaluation process should be significant or consequential for it to be respectable.

The argument advanced by some educators was that if the system cannot be improved, then it should be removed. This is suggestive for a review of the evaluation system which resonates with the RADPS' proposition for ongoing monitoring and evaluation of the system to facilitate continued improvement. Some participants however argued that the system should just be removed since it has not brought any improvement or benefits since its inception. Heneman et al. (2006:6) assert that for an evaluation system to survive and contribute to performance improvement, it must be accepted by those who use it. Evidence from this study was that this system has neither been embraced nor accepted by the implementers. The likely chances of it having a positive effect were therefore greatly hampered.

5.10 Recommendations by the participants on how to improve RBM

The participants made a number of recommendations that were on the different components of the evaluation system from timing of introduction, design process, implementation process and finally significance of process.

5.10.1 Recommendations to enhance the appropriateness of the timing of the introduction

Participants gave a number of suggestions on how the appropriateness of timing of introduction could have been enhanced. The most popular suggestion by the participants was that the government should have sourced and secured adequate funding first before attempting to introduce a new programme. Adequate fiscal resources were required to facilitate adequate initial training, procurement of the requisite material resources and supporting the system through the provision of associated incentives. The need to source fiscal resources is illustrated in Participant 1 (teacher)'s remark that, *"...whoever came up with this should have first sourced funds to make sure during the holidays, teachers in batches are accommodated at education institutes being to be trained"*. This remark suggests that training could have been hampered by lack of funding so the sourcing of adequate fiscal resources first would have facilitated adequate pre-implementation training. On the aspect of material resources, Participant 5 (head of department) posited that, *"... that could have been helpful and maybe sourcing as many textbooks as possible for the pupils so that teachers would have adequate resources to facilitate teachers to achieve their targets... we need adequate material to successfully implement"*. This remark

emphasises the need for adequate material resources for the successful implementation of the evaluation programme. Participant 3 (teacher) added that, *“There was a need for the system to have monetary incentives because without such incentives, the teachers cannot be motivated...If they wanted to introduce the system, they should have seen to it that they had adequate resources first”*. The emphasis being made is that the programme drivers should have sourced adequate funding to allow them to incentivise the system so that teachers would be motivated.

The second most popular suggestion was that there should have been active stakeholder consultation from the onset. Participant 8 (head of department) pointed out that, *“I think they should have consulted us as stakeholders because we are the ones that are on the ground. After consulting us, they should...”*.

The third common suggestions were that adequate training should have been provided prior to introducing the performance evaluation system and that piloting of the system should have been done before full scale implementation. Piloting would allow the necessary reviews to be effected before full scale implementation. On the aspect of training, Participant 8 (head of department) explained that, *“Maybe if they could have taken time to train the teachers about the system before it was implemented... If training had been more thorough, maybe people would know more about the whole system and be more prepared”*. Indications are that the educators were still not very knowledgeable about how implementation should be done to this day. The idea of piloting the system was suggested by a number of participants.

Participant 7 (head of department) said, *“... they should have piloted it first before full scale implementation although the government usually claims that it has piloted... I have never been part of a government pilot program in my 24 years of service although many new programs and new things have been introduced during this time”*. Participant 12 (school head) added,

Yes, they could have sampled schools, for instance, we have ten provinces in this country where schools are located. Some schools in these provinces are rural, some are urban, some are in farming areas and others in mining areas making the schools contextually different. They could have sampled heads from all the different categories of the schools and crafted the instrument with them. They could have further sampled schools after having a draft instrument and could have tried the instrument in the sampled schools before full scale implementation. That could have improved or

lessened the burden that came with the instrument because they would have gotten information from these schools to improve the instrument.

The participant had the view that piloting the system first would have assisted in coming up with a more appropriate and relevant system. Piloting would create a forum for the stakeholders to input their contributions thereby enhancing the appropriateness of the system and also enhancing its acceptability.

Other suggestions that were given included: conducting a readiness assessment first, aggressively marketing the programme and introducing the different aspects of RBM in a gradual approach and alternatively shelving the introduction of the new programme until the existence of an economically conducive environment.

5.10.2 Recommendations to improve the design process

The participants offered a number of suggestions that could be used to improve the design process of the system. One of the suggestions that were given was on stakeholder involvement. Participant 12 (school head) proposed, *“There should be more stakeholder involvement and participation in the design process. It helps the people to accept the product better”*. Besides involving stakeholders at design phase only, Participant 4 (teacher) further clarified, *“The design process of an appraisal system can be improved by putting in place an all stake holder committee at cluster level so that this committee gets the views of the people, the challenges of implementation and suggestions of improvement”*. The design process is not a once off event as it should be ongoing as highlighted by participant 4 (teacher). The system should be reviewed and improved on a continual basis and putting in place an all stakeholder committee to coordinate this process would assist in ensuring a well-coordinated process.

Participants also suggested that the evaluation instrument should be customised to the education sector. Participants felt that the current instrument was too general and that a tailor-made instrument for the Ministry of Education will serve a better purpose. This is illustrated in the remark by Participant 9 (school head) who stated, *“... there is a need to make the form more suitable to the education sector since right now there is just a general form”*.

Other suggestions that were given are that:

- The evaluation cycle should be changed so that it allows the factoring in of public examination results as one of the measures of teacher performance.
- There should be consultation to come up with a shared definition of effective teaching.
- That the different aspects of teaching should be given equal weighting.
- The evaluation policy should be more open to allow the implementers to design the finer and more specific details of the evaluation system in line with their contextual situation but within a national policy framework.

5.10.3 Recommendations to improve the implementation process

A number of suggestions on how to improve the implementation process were given by the interview participants. Most of the participants suggested that involving the stakeholders would help to improve the implementation process. The implication is that the implementation process is hindered by the negative attitude of the implementers who have failed to embrace the system because it was simply imposed on them. Participant 2 (teacher) explained,

I think implementers should be taken seriously by ensuring that things are not imposed on them by people who are not on the ground. I therefore think that the concerned parties should be convened so that they discuss and come up with something tangible. The major issue is that the implementers were not consulted beforehand so the attitude of the implementers is grossly negative. They are failing to see anything positive about the appraisal process. Probably if we had been involved from the beginning, we could probably implement the whole thing properly.

Participant 7 (head of department) further added, *“It is difficult to accept something that has been imposed on you but if it’s something that I am involved in right from the beginning, it is easier to embrace”*. The implications from the remarks are that implementation is hampered by lack of buy-in which can be addressed through stakeholder involvement even at this stage.

The provision of material and financial resources was also given as a key way of improving the implementation process. Participant 10 (school head) explained, *“I think there is a need to ensure that the schools are well resourced and have the necessary infrastructure... Materials are scarce so it becomes an impediment to implementation”*.

To further elucidate on the aspect of resources and how their provision can improve the implementation process, another participant added,

Implementation can be improved by availing resources so that we can build more classrooms so that we can have a single session instead of two sessions. That will help to avail more time for other activities like the appraisal. Material resources are also inadequate which makes it difficult for teachers to discharge their duties properly (Participant 9: school head).

The availing of financial resources would lead to infrastructural development and procurement of material resources which would in turn improve the implementation process.

Training has also been suggested as another important way that can be used to improve the implementation process. To clarify on training, Participant 3 (teacher) stated, *“Maybe they could have taken teachers in batches going for say a seminar or a workshop for 2 or 3 days, inducting them exactly on what is expected rather than inducting our superiors so that they come and induct us”*. Indications were that most of the implementers were not very clear and knowledgeable on how the RBM should be implemented. New members have since joined the profession and were not part of the contingent that got initial training which has been said to be inadequate since some implementers were left in a state of confusion and others did not get the opportunity to get any training at all. Other suggestions that have been made include:

- Reducing the class sizes.
- Incentivising the system as a way of motivating the teachers.
- Employing more teachers to assume the additional load created through the introduction of the new curriculum.
- Using learner achievement as one of the measures of teacher performance.
- Reviewing the evaluation system in line with the current developments in the ministry.

5.10.4 Recommendations to enhance the significance of the appraisal process

After indications by the majority of the participants that the evaluation process was insignificant, participants were asked to suggest ways which could be used to enhance the significance of the evaluation process. A variety of suggestions were given. The

most popular suggestion given by the participants was that the process should be incentive/reward linked. Participants felt that pay should be somehow linked to performance and this could be on a sliding scale based on one's performance rating. As a suggestion to enhance the significance of the appraisal process, Participant 11 (school head) said, *"For it to be effective, then the people need to be remunerated. Our ratings range from 0 to 6 so there would be a need to have a clear scale which shows differing rewards for all the ratings"*. Some participants have suggested that non-monetary rewards can also be used as a way of motivating the implementers and making the process respectable. In support of this point, Participant 10 (school head) said, *"I think the hard-working teachers should be rewarded by monetary or non-monetary rewards. Those who are lagging behind should be given what is suitable to them"*. From the suggestions given by the participant, it was evident that they believed that linking the process to incentives and rewards was one sure way of enhancing the significance of the process.

Participants also felt that underperformers should be exposed to some form of sanction. Suggestions ranging from withholding of an annual bonus, preferring charges, deduction of earnings and even termination have been proffered. Participant 7 (head of department) suggested, *"The non-performers should have negative reinforcement used on them like withholding annual bonus"*

It was also been suggested that underperformers could be made to undergo mandatory in-service training or even retraining before dismissal if there was no improvement. Participant 11 (school head) suggested, *"Non performers should be made to undergo in-service training or staff development which can be at school, cluster, circuit or district levels"*. In addition to this, Participant 12 (school head) also suggested, *"...the teachers that persistently perform poorly in- spite of mentoring and staff development programs should be sent for retraining or even dismissal"*. The participants emphasized the need to make the evaluation process significant by making it consequential.

5.11 Chapter summary

In this chapter, I presented empirical evidence that I collected through questionnaires and individual face-to-face interviews on the effectiveness of the teacher performance evaluation system in Kwekwe district. Detailed descriptions of the quantitative and

qualitative samples were given. I then presented the results of the regression analysis in the form of a model summary, ANOVA and Beta coefficient. The purpose of the model summary and the ANOVA was to explain the suitability of the regression model for the analysis and the acceptability of the model from a statistical perspective, respectively. The Beta coefficient explained the strength of the relationship between the independent and the dependent variables.

In the chapter I also presented, analysed, interpreted, and discussed the results to the four research questions outlined in chapters one and four. Findings from this study indicate that the timing of introduction of the system was inappropriate. Educators felt that the system was introduced prematurely before adequate preparations had been done to facilitate a smooth introduction of RBM. Suggestions on what could have been done to enhance the appropriateness of the introduction of the system were given and these included securing adequate funding first, marketing the programme and including the stakeholders in the marketing process, conducting readiness assessment, providing adequate initial training and piloting the programme.

It also emerged from this study that the educators felt that the design process was faulty as it was non-participatory. The feeling was that their input through their involvement in the design process could have enhanced the relevance and validity of the system. They found it difficult to accept something that was simply imposed on them and which they did not really understand. Educators suggested that the design process could be improved by involving the stakeholders by seeking their input, having a general evaluation framework and letting districts or schools design the finer details of the system, review the system, change the instrument and giving equal weighting to the different aspects of teaching.

Facts that also emerged from the study are that there was no proper implementation in a substantial number of schools that participated in the study. Participants explained that there was generally no monitoring and supervision of teacher performance at all and that the evaluation process was limited to the submission of the work plans at the beginning and rating at the end of the evaluation cycle. Indications were that the evaluation instrument is just completed for compliance purposes although there were no associated activities on the ground. There were indications that there were a myriad of impediments hounding the implementation process but the participants also

suggested ways that could be used to improve the process and these included involving the stakeholder by considering their suggestions, providing adequate resources and training, reducing class sizes and using more teachers could improve the implementation process.

The other finding of this study was that the educators appeared to have lost respect for the evaluation process. The majority of the participants viewed it as more of a burden and a waste of time than anything else. The fact that nothing happens to either the high performers or the poor performers had turned the process into simply a perfunctory and ritualistic exercise. The whole process was regarded as insignificant and educators had become indifferent to the ratings that they were given at the end of the evaluation cycle. There however appeared to be no consensus on the issue of the capacity of evaluation process to promote professional development. Some educators argued that at their schools professional development activities were informed by evaluation findings. There was however another section of educators who argued that there was no link or connection whatsoever between the evaluation findings and professional development. Indications made by some educators were that the evaluation process did not help them to improve their instructional practice. The different effects of the evaluation process on professional development could be indicative of an unstandardised implementation process across the schools. Participants suggested that the significance of the system can be enhanced by incentivising the system, sanctioning the underperformers and linking the evaluation process to professional development activities by providing targeted staff development. The general feeling of the educators was that the RBM system needs to be improved by customising it to the education sector and making it consequential. The feeling was that if it cannot be improved, then, it should be removed totally as it was currently merely a compliance exercise which was serving no purpose besides being a burden to the implementers.

CHAPTER SIX

Summary, Conclusions and Recommendations

6.1 Chapter Overview

This chapter presents a summary of the research process and gives an overview of the major research findings. Next follows conclusions made based on empirical evidence and critiqued by relevant literature for each of the four research objectives. The significance of the study is then discussed as well as the new knowledge that has been generated. Informed by the major findings and conclusions of the research, I then provide a list of implications of the study that can be considered to enhance effectiveness of teacher performance evaluation systems.

6.2 Introduction

The aim of this study was to assess the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. The aim of the study was achieved through the following specific objectives that intended to:

- Investigate the relationship between the timing of the introduction of the teacher performance evaluation system and its effectiveness.
- Examine the design process of the teacher performance evaluation system in Kwekwe District.
- Appraise the implementation process of teacher performance evaluation system in Kwekwe district.
- Assess the significance of the teacher performance evaluation process in Kwekwe District in promoting professional development and accountability.

6.3 Summary of the research process

The study was a mixed method study with a convergent parallel design. It was underpinned by the Pragmatic philosophical world view. The study was premised on the RADPS conceptual framework on performance evaluation system effectiveness. The four components of the conceptual framework are *Readiness Assessment, Design, Process, and Significance* (RADPS). This study had two samples for the quantitative and qualitative strands of the study. For the quantitative strand, stratified purposive sampling technique was used to select a total of ten schools constituting five urban and five rural secondary schools. From the ten schools, a total of 292

teachers were selected through simple random sampling to make up the quantitative sample. For the qualitative strand, stratified purposive sampling technique was used to select four schools constituting two urban and two rural schools. Purposive sampling technique was then used to select three key informants in the form of the Head or deputy head of the school, a head of department and a teacher from each of the four schools to come up with a total of 12 participants for the qualitative sample. Questionnaires were used to collect quantitative data and individual face-to-face interviews were used to collect qualitative data. Quantitative data were analysed using SPSS version 25 while the qualitative data was analysed using ATLAS. ti 8. Quantitative and qualitative results were presented separately before merging them to compare, contrast and corroborate results from the different strands. The merged findings were then discussed by placing them in the broader context of existing related literature.

6.4 How the study responded to the research objectives

The aim of the study was to assess the effectiveness of the teacher performance evaluation system of Kwekwe district. This aim was achieved through four objectives as follows:

6.4.1 Objective 1: To investigate the relationship between the timing of the introduction and effectiveness of the system

Through this objective, the study was investigating the readiness to implement the evaluation system. Because the system was already in place when this study was conducted, the intention was therefore simply to determine the appropriateness of the timing of the introduction and how this could have influenced the effectiveness of the performance evaluation system. Teachers' perceptions and opinions about the timing were used to respond to this objective. It was established that the timing of introduction of the teacher performance evaluation system negatively impacted on the performance system effectiveness. Evidence to support this assertion is presented in sections 5.5.1 and 5.5.3 of this report: Findings from both the quantitative and qualitative strands of the study indicate that the lack of consultation led to the dearth of buy-in by the educators. A total of 54.5% of the questionnaire respondents highlighted that no consultations were made prior the introduction of RBM and this was corroborated by participants in the interviews one of whom commented that, "*It's something which was just imposed on us without knowing...*" (Participant 6: head of

department). Participant 4 (teacher) added that, “ *There was need for the ministry to assess if the stakeholders embraced the system before introducing it*”. The comments show that the system had been imposed and readiness assessment was not done to ascertain buy-in. It also emerged that the educators had no knowledge and understanding of how the system was supposed to be implemented. This was affirmed by 65.4% of the respondents who admitted that training had been inadequate. Participant 2 (teacher) substantiated this, “... *only a few people were coached and when they came back to the schools, they didn’t actually understand what they had learnt. We were all confused about what we were supposed to do so we were simply copying and pasting from each other*”. Testimonies by the participants show that the educators did not have adequate understanding and knowledge of the system that they had to implement. The interviews further revealed that the driving ministry had no adequate resources to support the introduction of the evaluation system. Participant 2 (teacher) proffered that, “*I don’t think the ministry had enough resources to support the introduction of the new system then, and this could have led them not to train everyone on the ground*”. This explains why 65.5% of the questionnaire respondents had refuted that pre-implementation training had been adequate. Financial resources were needed to facilitate marketing and pre-implementation training as well as implementation of the system.

Participants preferred that timing of introduction could have been made appropriate by first sourcing adequate funding, providing adequate training, involving stakeholders in aspects like marketing of the programme and piloting the programme before full implementation. Readiness assessment could then have been conducted to ascertain the readiness of the educators and the schools to implement the programme.

Further evidence on this claim is also presented in section 5.4.1.3 of the report where the Beta coefficient table is presented. The table gives the statistical significance of timing on teacher performance evaluation effectiveness and also shows the Beta coefficient value of timing. The significance value of timing is shown to be 0.000 which means that the relationship between timing and the effectiveness of teacher performance evaluation system is statistically significant. The Beta coefficient value of timing is 0.213. This Beta value shows that timing is in third position relative to the other three independent variables of design, implementation and significance of

process in terms of its influence on the effectiveness of teacher performance evaluation system. In view of the findings made, this objective was therefore fully met.

6.4.2 Objective 2: To Examine the design process of the teacher performance evaluation system in Kwekwe District

This objective sought to investigate the way the teacher performance evaluation system was designed by assessing the level of stakeholder involvement. In exploring the design, other factors like the relevance and suitability of the evaluation instrument as well as its comprehensiveness were considered. The study also sought to find out the utility of the guidelines for completing the appraisal form. The views and opinions of the educators were used as the basis of answering the objective. The study established that the design process of the teacher performance evaluation system was faulty because it was non-participatory and lacked the input of the teachers as the experts in the field. The empirical evidence to support this claim is presented in sections 5.6.1 and 5.6.3 of this report: Although the teachers (81.3%), acknowledged that they know about what constitutes good teaching and 55.5% admitted they are cognisant with the means to measure teacher performance accurately, the study revealed that the educators had been excluded from the design process as pointed out by 59.9% of the respondents and also confirmed in the interviews where Participant 6 (head of department) remarked that, *“We were not consulted or involved...”*. It also emerged from the study that although the educators admit that the instrument is comprehensive as confirmed by Participant 5 (head of department) that, *“...it adequately captures all the aspects of teaching”*, Participant 6 (head of department) however argues that, *“...I don’t think it captures the teacher’s performance accurately...”* showing cynicism in the ability of the instrument to capture actual teacher performance. The educators indicate that the instrument is more of activity-based when on the other hand it is purported to be result-based and that it excludes student achievement in its measures of teacher performance thereby making it deficient. Suggestions from the participants on improving the design process included involving stakeholders by seeking their input, developing a shared definition of effective teaching, reviewing the system and having a general national evaluation framework and leaving the specific details to be developed at local level. Further evidence to how this objective was achieved is presented in sections 5.4.1.3 on the Beta coefficient table where the Beta coefficient and the statistical significance of the

design variable are given. The significance value of the design variable is given as 0.000 meaning that the relation between design and effectiveness of teacher performance evaluation system is statistically significant. The Beta value is 0.204 showing that the design process is the least influential variable relative to the other three independent variables of timing, implementation and significance of process on the effectiveness of teacher performance evaluation systems.

6.4.3 Objective 3: To Appraise the implementation process of teacher performance evaluation system in Kwekwe district

The implementation process was assessed through this objective by considering issues such as the objective setting process, the quality of evaluation feedback, knowledge of the appraisal policy and the challenges to the implementation process. To investigate to the objective, the views and opinions of the educators on the implementation process were considered. Evidence to how the study responded to objective 3 is presented in sections 5.7.1 and 5.7.3 of the report: The indication from 82.5% of the teachers was that the appraisal cycle starts by the evaluator and evaluatee discussing and agreeing on the work plan for the particular appraisal. To endorse the indications by the questionnaire respondents, Participant 3 (teacher) posited that, “ *Work plans are discussed and agreed on at the beginning of the year*”. Indications are that the evaluation process is a fallacy in most of the sites because the reality is that there is no activity between the development of the work plan at the beginning of the cycle and the summative rating at the end of the cycle. To justify the claim, Participant 2 (teacher) substantiated that, “ *Due to lack of time you find teachers and HODs sitting down and just agreeing on the mark to be awarded to the appraisee...*”. Although review dates are set at the beginning of the appraisal cycle, it was however revealed from the interviews that it is merely for compliance purposes. Indications are that teacher performance is not monitored or supervised in most of the sites thereby rendering the formative function (responsible for improving instructional practice) of the evaluation system null and void. In the absence of continuous monitoring and supervision, the summative ratings are then debatable. It was also revealed that participants are not familiar with the appraisal policy and in support, Participant 6 (head of department) confessed that, “ *I have never seen the Statutory Instrument or heard about it*”. Participants explained that shortage of time is a major impediment to the implementation process. Lack of material resources, excessive

clerical work, and existence of abnormally large classes have been cited as key challenges to implementation. Participant 11 (school head) elucidated that, *“The challenges we have include time factor, resources and our environment which can be prohibitive...”*. It also emerged that the failure to embrace the system by the educators also impedes the implementation process as professed by Participant 2 (teacher) that, *“The system was imposed on us and we are not keen to learn”*. Reviewing the system and providing a platform for the educators to give their input was suggested as a way of promoting buy-in and ownership and likely to improve implementation. Participant 10 (school head) suggested that, *“The appraisal system needs to be reviewed...”* to which Participant 2 (teacher) added, *“...parties should be convened so that they discuss and come up with something tangible”*. It was also revealed that the provision of adequate resources (finance, material and infrastructure), training on RBM and reduction of class sizes would improve implementation.

The Beta coefficient table presented in section 5.4.1.3 further proved how the study had achieved the objective. The table showed the significance value of implementation process as 0.000 which showed that the relationship between implementation and the effectiveness of teacher performance evaluation system was statistically significant. The table also presented the Beta value of implementation was 0.368 showing that implementation was at par with the significance of evaluation process in terms of their influence on the effectiveness of teacher performance evaluation system.

6.4.4 Objective 4: To Assess the significance of the teacher performance evaluation process in Kwekwe District in promoting professional development and accountability

This objective guided the study in evaluating the consequentiality of the evaluation process. In making this assessment, two issues were considered, the capacity of the evaluation process in fostering professional development and accountability. The perceptions and opinions of the educators were used to answer to the objective and the evidence is shown in sections 5.8.1 and 5.8.3 of this report: The study revealed that there are some schools where educators claim that evaluation findings are linked to professional development activities and on the other hand, schools where it was revealed that no link exists at all. To validate the professional development link, Participant 4 (teacher) asserted that, *“ We try to do this and focus our programs towards identified weaknesses”* but on the other hand, Participant 2 (teacher) argued

that, “...I haven’t seen that link or connection”. This study has revealed there is no uniformity in the implementation of RBM across the different sites hence the variations. In terms of accountability, the study revealed that there is no connection between what is stated as the possible consequences of the evaluation ratings in the Zimbabwe Public Service Regulations (2000:12-13) and what is on the ground. Indications are that the process is insignificant as it has neither rewards nor sanctions to the outstanding or under-performer, respectively. Participant 8 (head of department) explains that, “...there are no consequences afterwards, nothing happens so the process is not significant at all”. Section 5.4.1.3 of this report presents further proof on how the study achieved this objective. The Beta coefficient table in that section gives the significance value of the significance of appraisal process as 0.000 denoting the existence of a statistically significant relationship between the significance of the appraisal process and the effectiveness of teacher performance evaluation system. The table also gives the Beta value of the significance of appraisal process as 0.368 putting it at par with the implementation process as the most influential variables on the effectiveness of teacher performance evaluation system relative to timing and design. Taking into account the findings made, it is fair to say that the objective was fully met.

6.5 Overview of the major study findings

The study, which sought to establish the effectiveness of the teacher performance evaluation system in Kwekwe district, came up with major findings relating to each of the four subsidiary objectives.

6.5.1 Challenges with regards to the introduction of RBM in Kwekwe

This study established a number of findings regarding the timing of the introduction of the teacher performance evaluation system. The first of the findings that emerged was that the performance evaluation system was simply imposed without engaging or consulting the educators as the key stakeholders. A remark by Participant 3 (teacher) that, “I personally feel that this was imposed on us. There wasn’t any dialogue” illustrated how the educators regarded the introduction of the system. In support, Participant 12 (school head) added, “The information is one way, from the top to the bottom. It was just an instruction given to the school heads to carry out”. Educators revealed that the driving ministry just forced the programme on the implementers without even attempting to market it.

The second issue that was revealed was that the performance evaluation system was introduced at a time when there were no resources to support the introduction of a new performance evaluation system. A larger section of the teachers, 55.8% admitted that they had not been trained and 65.4% indicated that they did not know how they were supposed to implement the system. Many educators attribute the failure by the responsible ministry to provide adequate pre-implementation training, market the programme, link the programme to incentives and rewards, pilot the programme and conduct a readiness assessment as proof that the resources were not adequate enough to support such an endeavour then. To clarify on this assertion, Participant 2 (teacher) said, *“I don’t think the ministry had enough resources to support the introduction of a new ministry then and this could have led them not to train everyone on the ground”*. Participant 3 (teacher) also added, *“...no-one really convinced the teachers that this thing was going to help us because most of still see it as a waste of time and a waste of resources...”*. This implied that the programme had not been marketed to the educators so they did not appreciate it as they did not understand its benefits.

Besides revealing their concerns and observations on the timing of the introduction of RBM as a teacher appraisal system, the educators also made suggestions on how the appropriateness of timing of introduction could have been enhanced. Educators suggested that the responsible authority should have sourced adequate fiscal resources first to enable it to market the new programme, provide adequate pre-implementation training, have funds to allow it to support the programme staff development programmes and provide incentives and rewards, and conduct a readiness assessment exercise to ascertain acceptance of the programme. To emphasize the idea of marketing the programme prior to introducing it, Participant 1 (teacher) said, *“...to sell the idea to them so that they also accept and understand the system”*. Educators also suggested that the programme should have been piloted first to provide for stakeholder input and facilitate improvement of the final product. A comment by Participant 7 (head of department) that, *“I think they should have piloted it first before full scale implementation...”* echoed this sentiment from the educators.

6.5.2 Inadequate stakeholder involvement in the design process

In examining the design process of the evaluation system, the study established a number of findings. The first of the findings was that the teacher performance

evaluation system was designed in a no-participatory way. The stakeholders were excluded from the design process as specified by 59.9% of the teachers. A comment by Participant 6 (head of department) that, *"We were not consulted or involved...this thing is just an animal which was just imposed on us..."* illustrates the sentiments of the educators on the development of the teacher performance evaluation system. To further clarify on the aspect of their exclusion from the design process, Participant 4 (teacher) added, *"The ministry should have instituted a two way communication and not forced the programme down our throat..."*. The two remarks both illustrated that there was no buy-in or acceptance. The exclusion of the educators from the design process did not help to promote acceptance as illustrated by the use of phrases such as, 'this thing is just an animal', and 'forced the programme down our throat'.

The second finding that emerged on the design process was that the educators consider the evaluation instrument to be inappropriate and having no capacity to capture the actual teacher performance. The majority of teachers, 57.3% stated that the appraisal process did not measure teacher performance accurately. Participant 9 remarked that, *"I do not think it manages to capture the performance level accurately..."*. Educators are of the view that their exclusion may have contributed to the development of inappropriate tools as posited by Participant 2 (teacher) that, *"I think it takes someone on the ground to design the tools because if someone who is outside tries to bring their own mechanism, those on the ground might find it difficult to implement..."*. This was also supported by Participant 7 (head of department) who said, *"Remember we are teaching learners with different learning needs and the instrument that we are not cater for such learners"*. This implies that the educators feel that the instrument is defective or inappropriate and thus fails to capture performance accurately. Educators were of the opinion that the instrument has a measurement deficiency (Haefele, 1993:24) because it leaves out certain criterion that they believe is necessary in measuring teacher performance. Participant 8 (head of department) argued that, *"...There is no way therefore that the form can really capture the actual performance unless the rating was done after the results are out"*. The sentiments by the educators emphasise the need for stakeholder involvement in the design process.

The third finding that was revealed concerned the evaluation guidelines. A total of 63% of the teachers stipulated that there were no clear evaluation guidelines. It was explained from the interviews that although the evaluation guidelines are quite useful, they are however not available in most of the schools. The usefulness was confirmed by such comments as, “Yes, *the guidelines are quite helpful in completing the appraisal instrument*” (Participant 12: school head). Many educators were not even aware of the existence of the guidelines as pointed by Participant 2 (teacher) that, “The guidelines? We do not even know the guidelines”. There were a few teachers who pointed out that although they do have the guidelines, they however do not bother to use them as confessed by Participant 7 (head of department) that, “...*because at the end of the day what they want is a completed form and I will produce it and give it to them so the guidelines might not be necessary for me*” (laughs). This shows a carefree attitude which can occur as a result of someone not embracing the system or someone that knows that the process does not have consequences. Some of the few teachers who had the guidelines had also indicated that “...*at times we don’t even bother to go back to them and refer because sometimes we do not understand them*” (Participant 4: teacher). The same participant added that, “*There is need for more education on the RBM, more trainings on the system and how to use the guidelines so that we are clear*”. The remarks emphasise the need for further training to ensure clarity.

The fourth finding that emerged was that it is comprehensive although the challenge is that the system was not implemented as it should as asserted by Participant 11 (school head) that, “...*it looks at the academic and co-curricular aspects. It therefore tends to cover all aspects of teaching*”. A total of 56.2% of the teachers confirmed that multiple sources of information are used to ensure comprehensiveness. Participant 8 (head of department) however acknowledged that, “...*if it were done properly, it is quite comprehensive but then in actual fact it is not done as it should*”. This implies that although the instrument is comprehensive, it might not serve the intended purpose due to improper implementation.

The educators also gave suggestions on improving the design process. One of the suggestions was that the teacher performance evaluation system design can be enhanced by reviewing the system periodically in line with the observations and in-put

of the implementers as asserted by Participant 10 (school head) that, *“The design process can be improved by creating a platform where stakeholders can input their ideas on how to improve the system”*. This would allow the educators’ ideas to be incorporated and in turn likely to improve acceptability. Educators propose that the evaluation instrument should be customised as suggested by Participant 9 (school head) that, *“...there is need to make the form more suitable to the education sector since right now there is just a general form”*. Another proposal suggested was that there should be a general national evaluation framework which allows the educators to design the fine details in line with their context but ensuring that they remain within the national framework. Participant 11 (school head) posited that, *“...it should leave it discretionary on the implementers to decide on the finer details”*. Another suggestion made by the educators was that there is a need to have a shared definition of what ‘effective teaching’ is like what happened in the case of Chile where they developed a shared definition through consultative means (Avalos-Bevan, 2018:299; Isoré, 2009:36).

6.5.3 Uneven implementation process of RBM

The study established a number of findings on the implementation process of the teacher evaluation system. The first finding that was revealed on the implementation process was that there are no pre-set standards of performance although there are guidelines like the Director’s Circular Number 36. The performance targets that were used as the standards for the evaluation process were discussed and agreed to by the evaluator and evaluatee as confirmed by 82.5% of the teachers.

The second issue that emerged from the study was that the evaluation process was not continuous and ongoing throughout the evaluation cycle. Indications by 68.5% of the teachers were that there is no continuous monitoring of the process. The observation made was that in most of the schools, evaluation activities are limited to the target setting and summative rating events as stipulated by Participant 2 (teacher) that, *“...there isn’t supervision...nothing is happening on the ground”*. The implication is that the evaluation process does not serve a formative function as there is no supervision through the appraisal cycle. The appraisal process is not continuous in spite of the appraisal regulations being clear that the evaluation process is meant to continually review and improve quality of service provided (Public Service Regulations,

2000:12). The summative rating given is therefore not justifiable thus making the process lose integrity and instead become a joke.

The third issue that was revealed was that the educators are not familiar with the regulations on the appraisal process. This is the situation on the ground although the Public Regulations of 2000 contains information on the purpose of evaluation, character and nature of evaluation and the consequences of the evaluation process. This therefore implies that the educators are unaware of these critical issues which might therefore make it difficult for them to appreciate the appraisal process as they do not know or understand its purpose and consequences.

The fourth revelation made on the implementation process was that the objectivity and validity of the rating process is contentious. The educators explained that although the rating process is supposed to be objective as it is based on clear performance targets, it is however not objective. Participant 7 (head of department) elucidated that, “...*there is nothing that actually happens between the beginning and the end of the appraisal cycle*” to justify the summative rating that is given. In support of Participant 7 (head of department), Participant 10 (school head) added, “...*teachers are just bunched either at 4 or 5...people are just cooking marks...*”. Educators also argued that the process is not valid because, “*The process however fails to capture the quality aspect*”, reasoned Participant 1 (teacher).

The educators also revealed that there are many challenges bedevilling the implementation process. The challenges include lack of time as asserted by Participant 4 (teacher) that, “...*the challenges include issues of time. The challenge of time is worsened by the fact that the school has double sessions and the new curriculum that has many more subjects...*”. Time is especially a challenge in the urban schools where there are double sessions compounded by demands of the new curriculum. A sum of 63.7% of the teachers confirmed that the appraisers do not have adequate time to conduct the appraisal. The lack of adequate resources has also been cited as a challenge to the implementation process. Participant 12 (school head) explained that, “*We have a challenge of financial resources that are needed for the effective implementation of the process*”. Educators indicated that they also have a challenge of classes that are too large as clarified by Participant 8 (head of department) that, “*We also have a challenge of classes that are too large*”. The

abnormally large classes make it difficult for the teachers to effectively discharge their duties and fulfil the appraisal requirements. The educators also revealed that the paperwork is bulky and the fact that the driving ministry is not providing enough forms worsens the situation. Participant 12 (school head) complained that, *“The paperwork is still bulky and it’s a challenge. Who is supposed to bear the cost of producing the paperwork?”* This appears to be a challenge as some schools indicated that they have to provide the forms as they cannot expect teachers to do that on their meagre salaries.

The educators also proffered suggestions on improving the implementation process. One of the suggestions they gave was that the provision of resources would help to mitigate some of the challenges. Participant 10 (school head) postulated that, “...there is need to ensure that the schools are well resourced and have the necessary infrastructure”. Schools have challenges of shortage of classrooms and teaching-learning materials which hinder the implementation process. Participant 9 (school head) also recommended that, *“Training should be conducted for the new members and old members who have challenges of implementation”* to allow all educators to have a good understanding and knowledge of how they are supposed to implement the system. It has also been suggested that class sizes should be reduced by addressing the teacher-pupil ratio. Participant 10 (school head) posited that, *“...some classes have up to 60 learners...the effectiveness of the teaching-learning process becomes compromised”*. The abnormally large classes have put a strain on the teachers and this has in turn affected their effectiveness.

6.5.4 Insignificant effect of RBM on professional teacher development and accountability

In assessing the significance of the teacher performance evaluation process in Kwekwe district, several issues were raised. There is controversy on the role that the system is playing in promoting professional development with some educators saying evaluation findings inform professional development activities while on the other hand some educators say that there is virtually no link or connection. The majority of the teachers, 67.5%, attested that the feedback from the appraisal process is not helpful. The teachers, 57.9% also admitted that the appraisal results do not inform professional development programmes. Some teachers have argued that, *“...checking of what is happening is done through supervision of schemes and lesson observations which*

allow shortfalls of teachers to be realised and in turn inform staff development programmes at departmental or individual level” (Participant 8: head of department) thus implying that there is a link between the appraisal process and staff development activities.

In terms of promoting accountability, educators argued that the process is ineffective as it does not have consequences. A participant commented that, *“Teachers do not respect RBM because there are no consequences. Whether you get a 5 or you get nothing, nothing is done to you...”* (Participant 6: head of department). Indications by 54.5% of the teachers were that there are no rewards for outstanding teachers and 66.5% of the teachers confirmed that there are no sanctions for the teachers that continuously underperform. It has also emerged that there is no meaningful link between the actions listed in the Public Service Regulations (2000:12-13) as possible consequences of the appraisal process with what is actually happening on the ground.

One of the possible consequences of the appraisal process according to the Public Service Regulations (2000:13) is to grant or withhold rewards of which 54.5% pointed out that there are no rewards associated with the process and 66.5% admitted that there are no sanctions for persistently underperforming teachers. The Public Service Regulations also state that possible action can also include demotion or discharge subject to, and in accordance with the disciplinary procedure. A total of 41.7% of the teachers however disagreed that the appraisal results inform decisions on termination while 44.9% of the teachers claimed not to be sure. The high percentage of teachers who indicated that they were not sure might be taken to mean that no dismissal occurs due to appraisal findings because if there was such a connection, then, the teachers would definitely know.

Several suggestions were given by the educators on enhancing the significance of the teacher performance evaluation system. One way that was suggested was to link the appraisal process to an incentive/reward scheme. Participant 11 (school head) suggested that, *“...our ratings range from 0 to 6 so there would be need to have a clear scale which shows differing rewards for all the ratings...”*. The idea being advanced here is that rewards should be offered to all on a sliding scale. Participant 9 (school head) emphasised, *“...There is need for policymakers to make the process have consequences so that both the achievers and non-achievers are catered for in*

the policy". Participants emphasized the need to make the process consequential to allow the system to gain respect from the implementers. Suggestions that include withholding annual bonus, deducting a certain amount from the salary, charging the educator with negligence and even dismissal were given for the non-performers. To illustrate these sentiments, Participant 7 (head of department) declared that, *"The non-performers should have negative reinforcement used like withholding annual bonus or charging them for poor performance"*.

Educators also suggested linking the appraisal process to staff development programmes as a way of enhancing the significance of RBM. Participant 11 (school head) clarified that, *"non-performers should be made to undergo in-service training or staff development..."* and Participant 12 (school head) added, *"The teachers that persistently perform poorly...should be sent for retraining or even dismissal"*. The sentiments show that the educators feel that remediation efforts should be made to improve instructional practice as a way of enhancing the significance of the appraisal process.

6.5.5 General position of educators on the way forward on RBM

The educators' general opinions on the evaluation system are that RBM needs to be improved by making it consequential and customising the instrument to the education sector. The educators suggest that the improvements made should emanate from a participatory and consultative process. A total of 81.5% of the teachers affirmed that RBM needs to be improved to suit the Ministry of Education requirements in agreement with this quantitative result, participant 1 (teacher) explained that, *"...the responsible ministry should review the whole system and try to involve the implementers so that they have their input and accept it as their own"*. Some of the teachers, 65.5% declared that RBM should be removed as teacher appraisal system. This was supported by people like Participants 2 (teacher) who proclaimed that, *"I think it should be stopped forthwith...I do not think it's necessary at all in the schools"*. The feeling of the participants was that RBM is not serving any purpose save for giving them an extra load.

6.6 Study conclusions

The aim of this study was to ascertain the effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe. In view of the findings, the study

drew conclusions on *readiness assessment, design, process and significance* in line with the components of the RADPS conceptual framework on which the study was premised.

6.6.1 Inadequate readiness for assessment

This study concludes that the timing of the introduction of RBM as a teacher evaluation system was not appropriate. The introduction of the system was actually mistimed in view of the fact that the drivers of the programme had not acquired adequate financial resources to support the introduction of a new system. The system was therefore not marketed to the stakeholders so that they could appreciate its benefits and buy-in. There was also inadequate pre-implementation training provided which resulted in most of the educators not knowing or understanding how the system is supposed to be implemented. Some educators indicated that they had never been given any form of training on the system at all. The observation was that there is no buy-in or acceptance of the programme. Indications are that the system was introduced prematurely without adequate planning and preparation. It can therefore be concluded that the timing of the introduction of RBM thus negatively impacted on its effectiveness as a teacher performance evaluation system.

6.6.2 Non- participatory design process

This study concludes that the design process of RBM was generally faulty as it was largely non participatory. Indications from the educators are that neither the educators nor their representatives in the form of the teachers' unions were involved in the design aspect of RBM. Educators were excluded from the design process in spite of their knowledge and expertise on what constitutes effective teaching and how it can be measured. The educators thus had no confidence in the ability of the evaluation instrument to capture the performance of the teacher accurately. The study therefore concludes that the instrument was not piloted to allow the teachers to make their own contributions which could have resulted in a more appropriate and credible evaluation instrument. The good features of RBM were overlooked by the educators because they have never really embraced the system. The educators had not accepted the system years after it has been introduced. There was no ownership of the system by the educators. They perceived the performance evaluation system as 'their thing' referring to the Ministry of Public Service, Labour and Social Welfare as the driving

ministry. It can thus be concluded that the non-participatory, top-down design approach negatively impacted on the effectiveness of RBM in Kwekwe.

6.6.3 Impurities in the implementation process

Indications from the educators were that the evaluation activities are limited to the performance target setting and the summative rating. There is no much activity in between the two events in spite of the review dates being set at the beginning of the appraisal process in most of the sites. It thus implies that there is no monitoring of the teachers' performance in most of the schools that participated in the study. Teachers also revealed that the rating process was neither objective nor valid as the summative rating ratings given are often not evidence-based but are just cooked when they are due. The study therefore concludes that the evaluation process is just a compliance exercise. Educators simply did what is necessary to satisfy the policy requirements. The evaluation process is consequently not serving its purpose of improving instructional practice and in turn quality of education. Educators were not even aware of the existence of the evaluation policy and its contents. This points to a weakness on the part of the programme drivers and the administrators. Just like effort was made for each teacher to be familiar with the Acts of Misconduct, the same approach should be done with the appraisal policy. This as a result implies that the programme drivers and administrators were not eager and supportive of the programme. The study hence concludes that there was no political will to drive the programme. In view of the myriad of the impediments hounding the implementation process, the study also concludes that unless these are addressed, then it may be far-fetched to think that evaluation process will be implemented properly. The study accordingly concludes that improper implementation has negatively impacted on the effectiveness of the performance evaluation system in this district.

6.6.4 Insignificance of appraisal process on teacher professional development and accountability

The study found that educators perceived the evaluation process as inconsequential because there are neither rewards nor sanctions for outstanding and under-performers, respectively. It is evident therefore that the educators have lost respect for the evaluation process and have become indifferent to the whole procedure due to its insignificance in terms of consequences. Educators were divided on the existence of a link between the evaluation findings and the professional development activities.

Some educators claimed that there is a strong link between evaluation findings and professional development programmes while on the other hand, there is a group that claimed that no such link exists whatsoever. The conclusion that the study draws on this issue is that it is evident that the operation procedure of the performance evaluation activities is not standard across the sites within the district. The lack of standardisation of the implementation process may be the result of the inadequate pre-implementation training that was provided. It is concluded that there is no serious monitoring of what happens on the ground by the education officials at district level. In the case where there is active monitoring, such implementation variations in the district should have been identified and addressed. This study also notes that the appraisal process is not being enforced at all since there is no connection with what is written in the Public Service Regulations as the consequences of the evaluation process and what happens in reality. The study hence concludes that the insignificance of the evaluation process has negatively impacted on the effectiveness of the teacher performance evaluation system in this district.

Given the above conclusions in the four key aspects of RBM, this study concludes that if RBM is to gain respect from the educators, then it has to be consequential. Currently the educators are indifferent to the ratings that they get and do not take the process seriously. It is nothing more than a compliance exercise. This study can thus conclude that there is little chance of the evaluation process being implemented properly and serving its purpose if it remains inconsequential.

6.6.5 The status quo on teacher performance evaluation in Kwekwe district

Figure 6.1 presents the status quo as regards teacher performance evaluation in Kwekwe district taking into consideration the findings of the study and the conclusions made thereon. It should be noted that the arrows on figure 6.1 indicate that the weakness on each aspect of RBM affects the whole RBM as well as affecting each sub-aspect.

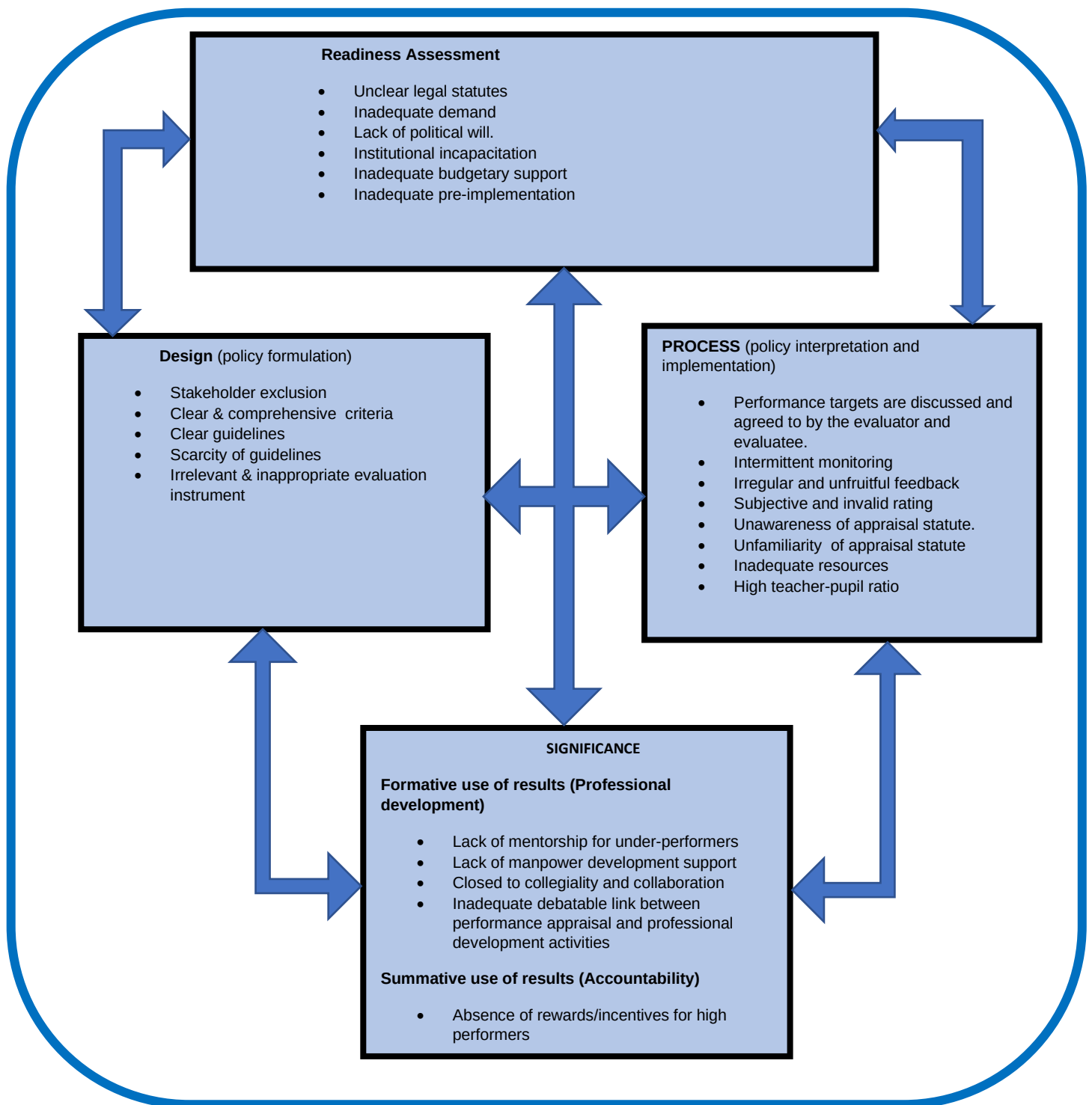


Figure 6.1 The status quo on teacher performance evaluation in Kwekwe district

6.7 Recommendations

In view of the findings of the study and the conclusions drawn thereon, the following recommendations are made:

6.7.1 Recommendations for Policy

The participants revealed that the evaluation policy was formulated in a non-participatory and exclusive way which resultantly led to lack of ownership and buy-in from the educators. This study therefore concluded that the top-down approach used for the design process negatively influenced the effectiveness of RBM. In view of this finding and conclusion reached thereon, this study recommends that policy formulation should be consultative and participatory. The study advocates that policy formulation should be consultative being spearheaded by a multi technical committee constituted of the target key stakeholders and other relevant experts. Teachers are a valuable resource in policy implementation, it is prudent therefore that they should be included in policy formulation as they are a valuable resource in the implementation process. This recommendation is advanced on the basis of the proposition of the RADPS conceptual framework, extant literature and also the research participants themselves' proposal. The RADPS conceptual framework proposes active stakeholder involvement in the policy formulation process. The involvement of a wide array of stakeholders including government officials, school administrators, teachers and their unions, and educational researchers cannot be overemphasized (Isoré, 2009:8-9). Research has established that a system that is developed from the contributions of all key stakeholders has better chances of being embraced and respected compared to one that is imposed (Goe et al., 2014:14; Stronge & Tucker, 1999:353). McLaughlin (1990:406) asserts that the exclusion of teachers from the design process promotes division and discord which is detrimental to the evaluation process. The involvement of teachers on the other hand has been found to promote the development of a sense of ownership (Stronge & Tucker, 1999:341). Chile is a good example of a country that used a participatory approach to the development of its teacher performance evaluation system successfully. Although the process took fairly long, it paid off in the end because it promoted buy-in and at the same time minimised resistance. The participatory nature of Chile's teacher performance evaluation has been highlighted as one of its key strength. Other studies that have been conducted on RBM have also

recommended that a participatory approach should be adopted to promote ownership (Machingambi, 2013:224; Pretorius & Ngwenya, 2008:162).

The study also established there is a single evaluation policy governing performance appraisal for all the government workers. Educators added that the evaluation instrument that is used is inappropriate and fails to capture the actual teacher performance. The study in turn concluded that the design process was faulty and the instrument inappropriate. In view of the complexity of the teaching task, it is therefore recommended that a policy customised to the education sector be developed rather than have a one-size-fits-all across the sectors. This recommendation resonates with the recommendation that has been made by other researchers like Gutuza (2016:395); Machingambi (2013:224) and Musingafi et al. (2013:26). The educators themselves have also highlighted the need to develop a model that is customised to the education sector as the perception is that the current model is more relevant to the production sector rather than the service industry. Participant 7 (head of department) observed that, *“...there is room for improvement or modification to make it more appropriate to the education sector. It is currently more appropriate for production industry than for the service industry such as the education sector”*. With regard to the validity and appropriateness of the evaluation instrument, the study recommends that the evaluation system should be assessed and monitored continually for validity and relevance. This will allow the system to be modified and reviewed in line with the observations made and can also create a platform for the educators' inputs to be incorporated which can in turn help to promote buy-in and ownership. Danielson and McGreal (2000:64) posit that it is only the users of an evaluation system that can certify its validity and relevance. As such, the involvement of the stakeholders from the drawing board can assist to validate the system from the onset like what happened in the case of Chile. A valuable lesson can further be learnt from the case of Chile, which used a bottom-up approach to develop the teacher performance evaluation system. The Ministry of Education of Chile however still contracted a University Measurement Centre assigned with the responsibility to continually assess and review the teacher performance evaluation system in a bid to maintain validity and relevance. The RADPS conceptual framework also proposes that continuous monitoring and review of the evaluation system should be done as a way of ensuring validity and relevance.

Another finding was that there is no link between the possible evaluation consequences listed in the Public Service Regulation (2000:12-13) and the reality on the ground. The study noted that there is no enforcement of the policy and no-one to hold to account as the policy is not explicit on who is supposed to take responsive action to the appraisal findings. The policy merely states that Commission or head of Ministry or department, as the case may be, shall decide on the action to be taken, **if any**, on the basis of the appraisal. This appears to be a deliberate way of avoiding accountability on the part of the policy makers. It also shows a lack of political will and commitment to the programme on the part of the government. In light of these findings, the study therefore concluded that the insignificance or inconsequentiality of the appraisal of the evaluation process affected the effectiveness of the teacher performance evaluation system negatively. Given these findings and subsequent conclusions made, the study thus recommends that the appraisal policy should specify the authority responsible for actioning the consequences. The policy should also clearly articulate the consequences for the different levels of performance rather than leaving it open and discretionary upon an authority that has not been specified. This is contrary to the assertion by Goe et al. (2014:43) that it is important to clarify from the outset the decisions on which specified action in response to evaluation findings will be actuated. The cases of China and Chile are good reference points on this issue. The teacher evaluation policies in these countries have clear consequences for the different ratings, and the period of actuation for the consequences is clearly stated from the outset.

6.7.2 Recommendations for practice

The study found out that there was inadequate pre-implementation training and no marketing of the appraisal system to the educators before the performance evaluation system was introduced. Participants attributed the inadequate training and lack of marketing to lack of adequate budgetary support for the new programme. The study therefore concluded that RBM was introduced prematurely and this ultimately negatively affected the effectiveness of the performance evaluation system. Taking into consideration these findings and the resultant conclusion, this study recommends that readiness assessment should be conducted before introducing a new programme to ascertain feasibility in view of institutional capacity and buy-in by the stakeholders. The RADPS conceptual framework proposes that readiness assessment should be

conducted before introducing a new programme. Readiness assessment considers institutional capacity, fiscal resources availability, demand for the programme and availability of technical skills and expertise (Kusek & Rist, 2002:152-153; Kusek & Rist, 2004:40; Lehman, Greener & Simpson, 2002:198). The participants also suggested that readiness should have been conducted to ascertain feasibility of introducing RBM. Participant 4 said, “They should also have done thorough research before rushing to introduce the system”, in support of readiness assessment.

The majority of the teachers (55.8%) claimed that they had not received adequate training on RBM. The educators admitted that they do not have adequate knowledge and understanding of the system, years after it has been introduced. This study hence concluded that the lack of knowledge and understanding of the system by the educators has therefore led to improper implementation which has subsequently impacted negatively on the effectiveness of the performance evaluation system. This study consequently recommends that training for both the evaluators and evaluatees should be ongoing to cater for any members having challenges with the system and the new members joining the ministry. This recommendation is backed by the assertion of White et al. (2012:8) that training should be ongoing to accommodate both the new teachers and old teachers who might need clarification on issues that arise during the implementation process. Other researchers like Zvavahera (2014:7), Madhekeni (2012:126), and Mavhiki et al. (2013:138) also support that training should be provided for members at all levels to promote understanding. The participants themselves also suggested that, *“Training should be conducted for the new members and old members who are having challenges of implementation”* (Participant 9: school head).

The participants also revealed that educators were excluded from the design process in spite of their expertise and knowledge on what constitutes good teaching and the means by which it can be measured. This resultantly led to the development of an evaluation instrument which was inappropriate. In light of these findings, this study concluded that the evaluation instrument was not pilot tested prior to full implementation. Taking into consideration the findings and the conclusions made thereon, this study recommends that a programme should be piloted first before full scale implementation. Literature proffers that a new evaluation policy should be piloted

first to provide a platform for the necessary modifications and improvements in answer to the pilot feedback (White et al., 2012:9). The benefits of piloting can be seen in the case of South Korea. South Korea attempted to use the bottom-up approach to designing the teacher performance evaluation system but abandoned it when stakeholders failed to reach a consensus. The Ministry of Education Science and Technology proceeded to design two versions of teacher evaluation systems and proceeded to pilot them. The procedure of piloting created a platform to get feedback from the stakeholders which was then incorporated. The piloting therefore served two purposes: one of getting stakeholder input which was used to improve and modify the instruments, and secondly, promoting buy-in and a sense of ownership. Madhekeni (2012:126) and Mavhiki et al. (2013:138) also emphasise the importance of piloting a programme first before full scale implementation.

Another issue revealed in the study was that the rating process is not objective as the summative ratings are often not evidence based because there is no monitoring and supervision of teacher performance throughout the appraisal cycle. It was exposed that the teachers in a department are often bunched between 4 and 5 ratings irrespective of the actual performance of the teacher as a way of avoiding external officials who are expected to validate the ratings which are either too low or too high. This study therefore concluded that the evaluation process is a mere façade. The evaluation procedure is not serving a formative function to improve instructional practice and the summative ratings given are just numbers without any meaning. Given these findings and conclusions, the study recommends that the responsible ministry should employ more education inspectors to allow active monitoring of the performance evaluation implementation process in the schools and ensure its functionality. There is need to standardise the implementation procedure which was observed to vary across the schools. To guarantee objectivity and avert issues of favouritism that were implied, this study recommends that there should be multiple evaluators including external evaluators as a way to certify objectivity and validate the summative ratings. The use of multiple evaluators has been used as a way of enhancing transparency, objectivity and credibility of results in countries like Chile and China. The RADPS conceptual framework also proposes the use of multiple evaluators rather than relying on a single evaluator as a way of making certain of rigor in the evaluation process (Hinchey, 2010:11; Santiago & Benavides, 2009:17).

6.8 Significance of the study

The findings of this study have implications regarding the implementation and effectiveness of a teacher performance evaluation system from different standpoints as follows:

6.8.1 Significance on research approach

In view of the complexity of the research problem, this study used the convergent parallel mixed method design. This was done to facilitate an exhaustive examination of the problem and also to ascertain collaboration of the results. Quantitative and qualitative data were collected for breadth and depth of findings. A sample size of 292 respondents (29.6%) of the total population was used for the quantitative strand. This sample size is adequate to be representative and allow a generalisation of the results to the total population. A total of 99.3% of the quantitative strand were teachers who are over 26 years of age and 92.5% of the teachers had over 6 years of teaching experience. Of the 292 teachers, only 14% had a non-teaching qualification and the least of these was a diploma while the remaining 86% had teaching qualifications ranging from a Diploma to Masters in Education. This means that the sample was thus made up of mature, qualified and experienced teachers who are likely to give reliable and informed views and opinions on the research issue. The qualitative sample was composed of twelve participants with a minimum of twelve years of teaching experience. The qualitative sample was constituted of teachers, heads of department and school heads. This allowed me to get a wider range of perspectives from the views of the different groups of participants. The informants were experienced educators who were purposively selected for their rich knowledge and experience on RBM since the system was introduced when they were already in service. On that account, it implies that the knowledge generated by the study is deep and rich. Quantitative and qualitative findings were compared for confirmation and collaborativeness. The data and findings generated by this study can be relied on and thus the study can be used as a springboard for further studies on the area of teacher performance evaluation systems.

6.8.2 Study significance on RADPS conceptual framework

The study was premised on the RADPS conceptual framework (*Readiness assessment, Design, Process and Significance*) on performance evaluation system

effectiveness. I developed this framework from extant literature and personal experience. The RADPS framework can be used to predict the effectiveness of teacher evaluation systems and also a guide in the examination and explanation of the performance evaluation system effectiveness. In view of this development, this study has thus contributed to new knowledge. A revised RADPS framework in line with the study findings is presented in Figure 6.1. The original version of the model is presented in Figure 2.1 in Chapter 2 of the report. Explanations on the different components of the conceptual framework, *Readiness Assessment*, *Design*, *Process* and *Significance* are presented in sections 2.5.1.1 to 2.5.1.4. Modifications that have been done on the RADPS conceptual framework are presented in red. These modifications were informed by the findings from the study. In the initial model of the conceptual framework presented in Figure 2.1, the arrows connecting readiness assessment and design component and the one connecting readiness assessment and process were uni-directional implying a one-way link. The arrow connecting the design and process components was also uni-directional, that is, from the design component to the implementation process. These arrows have now been modified to show a bi-directional link between the components. The process of evaluation is an iterative process which involves moving back and forth between the components hence the need to have the bi-directional arrows to show this reality. Another notable modification is the addition of two factors in the *Readiness Assessment* component. The study revealed that a programme has to be piloted first to ascertain the readiness of the implementers before full scale implementation. The need to consult stakeholders to ascertain their readiness and acceptance and involving them in marketing of the product as a way to enhance acceptance and ownership was also suggested.

Improving performance evaluation system effectiveness

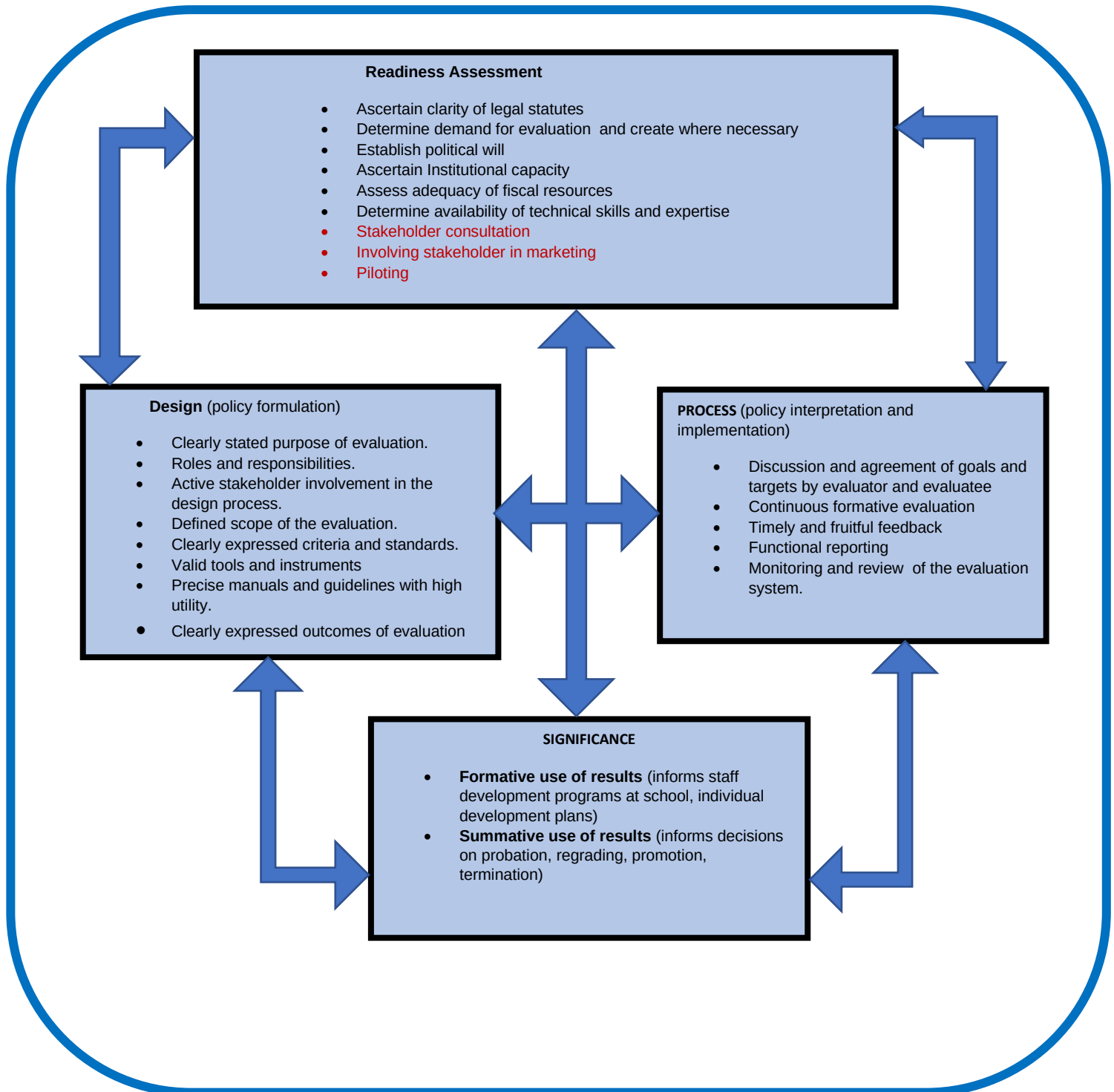


Figure 6.2 Recommended RADPS Conceptual Framework for RBM in Kwekwe district

6.8.3 Significance on practice and policy

The findings from the study also have implications on practice. The knowledge generated by the study can be used to improve the different aspects of the current teacher performance evaluation from the aspects of *design*, *process* and *significance* since it is already being implemented. The design process should be participatory and inclusive, the implementation process should be continuous to be formative, and the evaluation process should have consequences to gain the respect of implementers. The knowledge generated by the study can also be used to inform the development of new programmes in future.

The study findings also have implications on policy. The study reveals the need for policy makers to relook the policy formulation approach with a view of making the process consultative and participatory so as to promote ownership. A policy developed through an inclusive approach has greater chances of being accepted and certified as valid from the drawing board (Goe et al., 2014:14; Stronge & Tucker, 1999:353).

6.9 Limitations of study

The data collection procedure was conducted during the time when Zimbabwe was preparing for the 2018 harmonised elections. The timing of the data collection procedure which coincided with the election preparation period and active campaigning could have compromised the data to a certain extent. The issue of teacher performance appraisal is rather political in nature, and so could have been a sensitive topic at that particular time. This may have made the participants not to answer as truthfully as they should. The use of the 5 point Likert scale may also have affected the findings to a certain extent. There were incidences where a significantly high number of the respondents opted for the 'Not Sure' option. This was quite glaring on some items where for example, over 30% of the respondents took this option. One of the quality measures taken to ensure the credibility and validity of study findings was to make use of established research methods (Nieuwenhuis, 2016:123; Shenton, 2004:64. The use of the mixed method approach and convergent design allowed the validation of results of the qualitative strand with those of the quantitative strand and vice versa, thereby addressing these two limitations.

Although the study included teachers, heads of departments and school heads who are the key stakeholders at school level, other key stakeholders such as the district

and provincial inspectors could have been included for a wider perspective on the effectiveness of the teacher performance evaluation system. They were however excluded because of the need to delimit the study to a specific scope for thoroughness. The perception of other stakeholders on the effectiveness of the teacher performance evaluation system can however be looked at in a another study as suggested in the section for further studies.

6.10 Suggestions for further research

This study should be taken as a springboard to further research in the area of performance evaluation system effectiveness. The study covered wide ground from readiness assessment, design, process and significance of the system. Some of the aspects of the study were not thoroughly examined and some interesting issues which need further examination also cropped up in the process of reviewing literature and analysis of research results. The study hence suggests the following areas for further research:

- An in-depth examination of the technical design of the teacher performance evaluation system.
- The use of student surveys as a source of teacher performance information.
- The effectiveness of collaborative and collegial approach to teaching.
- The quality of professional development programmes provided to the teachers.
- The perception of policy makers on the effectiveness of the teacher performance evaluation system.

6.11 Chapter summary and conclusion

This chapter presents a synopsis of the whole study. It highlights the key issues that emerged from the study pertaining to the effectiveness of the teacher performance evaluation system in Kwekwe district. The chapter starts by presenting a summary of the research process to give an outline of the procedures that were conducted from the research design, methodology and analysis of the data. Explanations on how the four sub-objectives of the study were met are presented and reference to the section of the report containing the relevant evidence to this effect is given.

A summary of major findings on appropriateness of timing of introduction, design process, implementation process and the significance of the study is presented. It was revealed that there was no dialogue with the educators but the system was simply

imposed and this affected buy-in. It emerged that there was inadequate budgetary support for the programme as the driving ministry failed to provide adequate pre-implementation training and failed to market the programme to the users. The performance evaluation system was designed in a non-participatory way. The evaluation instrument is regarded as inappropriate and incapable of measuring the actual teacher performance. The performance standards that are used as the basis of appraisal are set and agreed to by both the evaluator and evaluatee. The educators revealed that there is no continuous monitoring and supervision of teacher performance in most of the schools and challenges to the implementation process include inadequate time, lack of material and infrastructural resources, existence of large classes and excessive clerical work. The educators also exposed that the evaluation process is inconsequential as there are neither rewards nor sanctions for both the top and under performers. It is evident that there is no meaningful link between the actions listed in the appraisal regulations as possible consequences of the appraisal process and what is happening on the ground.

Conclusions were drawn on each of the four objectives of the study in accordance with the major findings. The general conclusions drawn were that the timing of the introduction of RBM was inappropriate, the design process was faulty, the implementation process is a fallacy and the evaluation process is insignificant. Also presented in this chapter are recommendations that are informed by the findings and conclusions of the study.

Recommendations for policy and practice are made. Key recommendations made on policy are that policy formulation should be participatory and inclusive. It is also suggested that instead of having a general evaluation model for all sectors of government, there should be a customised model for the education sector. In terms of clarity, the current policy has been found wanting, so it is recommended that the policy should specify the authority responsible for actioning the consequences of the evaluation findings and also specify the consequences for the different levels of performance. In terms of practice, it is recommended that readiness assessment should be conducted before introducing a new programme to ascertain the feasibility of the programme. A new programme should be pilot tested before full scale implementation to give room for stakeholder feedback which can then be used to

improve the programme and enhance ownership. Training should not be a one off event but should be ongoing to accommodate those having challenges with the implementation aspect and at the same time accommodate the new members. There should be multiple evaluators including external ones for objectivity in the rating process and validation of the summative ratings. The performance evaluation system has to be monitored and assessed continually for validity and reviewed in line with the observations and inputs of the users.

The findings of the study have implications on the effectiveness of teacher performance evaluation system and these are given from different stand-points as implications on research, implications on theory, implications on practice and implications on policy. The value of the study in relation to each of the listed factors is given.

In view of the findings that were made, the study is concluded by asserting that the timing of introducing a system, the design process, the implementation process and the significance of the evaluation process are key factors that influence the effectiveness of a teacher performance evaluation system.

References

- Archer, J., Kerr, K.A. & Pianta, R.C. 2014. 'Why measure effective teaching?' in Kane T.J, Kerr, K.A & Pianta, and R.C (Eds.) *Designing Teacher Evaluation Systems: New Guidance from the Measures of Effective Teaching Project*, [Online] San Francisco: Jossey Bass. Retrieved from: <http://www.metproject.org/downloads/designing.pdf&sa> [Accessed on 20/10/15].
- Armstrong, M., 1996. *A Handbook of Personnel Management Practice*. New Delhi: Kalyani Publishers.
- Arvey, R.D. and Murphy, K.R., 1998. Performance evaluation in work settings. *Annual review of psychology*, 49(1), pp.141-168.
- Avalos-Bevan, B., 2018. Teacher evaluation in Chile: highlights and complexities in 13 years of experience. *Teachers and Teaching*, 24(3), pp.297-311.
- Babbie, E. 2014. *The practice of social science research*. 14th ed. Boston: Cengage Learning.
- Bagraim, B., 2016. 'Motivation and engagement of the South African workforce' in Werner, A. (ed) *Organisational behaviour: A contemporary South African perspective*. 4th ed. Pretoria: VanSchaik Publishers, pp.92-132 .
- Baker, T.L., 1988. *Doing social research*. Singapore: McGraw-Hill College.
- Barry, C.A., 1998. Choosing qualitative data analysis software: Atlas/ti and Nudist compared. *Sociological research online*, 3(3), pp.1-13.
- Beerens, D.R., 2000. *Evaluating Teachers for Professional Growth: Creating a Culture of Motivation and Learning*. Thousand Oaks, CA: Corwin Press.
- Berger, R., 2015. Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative research*, 15(2), pp.219-234.
- Bertram, C. and Christiansen, I., 2014. *Understanding research: An introduction to reading research*. 1st Ed. Pretoria: Van Schaik Publishers.
- Bertram, C. and Christiansen, I., 2015. *Understanding research: An introduction to reading research*. 4th Ed. Pretoria: Van Schaik Publishers.
- Birks, M., Chapman, Y. and Francis, K., 2008. Memoing in qualitative research: Probing data and processes. *Journal of research in nursing*, 13(1), pp.68-75.
- Blaikie, N., 2000. *Designing Social Research*. Cambridge: Polity.
- Blanton, L. P., Sindelar, P. T., Correa, V., Hardman, M., McDonnell, J., & Kuhel, K. (2003). Conceptions of beginning teacher quality: Models for conducting research (COPSSE Doc. No. RS-6). Gainesville: Center on Personnel Studies in Special Education (COPSSE), University of Florida. Retrieved from <http://www.coe.ufl.edu/copsse/docs/RS-6/1/RS-6.pdf> Accessed on 15/10/15
- Bryman, A., 2001. *Social Research Methods*. Oxford: Oxford University Press.

Burton, N., Brundrett, M. and Jones, M., 2014. *Doing your education research project*. Thousand oaks: Sage.

Callahan, K. and Sadeghi, L., 2015. Teacher Perceptions of the Value of Teacher Evaluations: New Jersey's ACHIEVE NJ. *International Journal of Educational Leadership Preparation*, 10(1), pp.46-59.

Cardy, R.L. and Dobbins, G., 1994. *Performance Appraisal: Alternative Perspectives*. Cincinnati, OH: South-Western Publishing Co.

Carroll, D.H., 1997. A Comparison of Clinical Supervision and Evaluation. *Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX, January 23-26, 1997*.

Casanave, C. and Li, Y., 2015. Novices' struggles with conceptual and theoretical framing in writing dissertations and papers for publication. *Publications*, 3(2), pp.104-119.

Castillo-Montoya, M., 2016. Preparing for interview research: The interview protocol refinement framework. *The Qualitative Report*, 21(5), pp.811-831.

Celik, S., 2011. Characteristics and competencies for teacher educators: Addressing the need for improved professional standards in turkey. *Australian Journal of Teacher Education*, 36(4), pp.73-87.

Centre for Development and Enterprise. 2015. Teacher evaluation: Lessons from other countries. Retrieved from: <https://cde.org.za/teacher-evaluation-lessons-from-other-countries/> Accessed 23/10/18

Chivore, B.R.S., 1992. Pre-service teacher education at a distance: The case of Zimbabwe. *Distance education in Anglophone Africa: Experience with secondary education and teacher training*, pp.103-115.

Choi, H.J. and Park, J.H., 2016. An analysis of critical issues in Korean teacher evaluation systems. *Center for Educational Policy Studies Journal*, 6(2), pp.151-171.

Classen, S., Lopez, E.D., Winter, S., Awadzi, K.D., Ferree, N. and Garvan, C.W., 2007. Population-based health promotion perspective for older driver safety: conceptual framework to intervention plan. *Clinical interventions in aging*, 2(4), p.677-693.

Coe, R., Aloisi, C., Higgins, S. and Major, L.E., 2014. What makes great teaching? Review of the underpinning research. Retrieved from: www.suttontrust.com/wp-content/uploads/2014 [Accessed on 18/10/15].

Cohen, L. and Manion, L., 1989. *Research methods in education*. 3rd Ed. London and New York: Routledge.

Cohen, L. Manion, L & Morrison, K. 2007. *Research methods in education*. London: Routledge Falmer.

Cohen, L., Manion, L., and Morrison, K., 2011. *Research methods in education*. 7th Ed. London and New York: Routledge.

- Collins, K.M., Onwuegbuzie, A.J. and Jiao, Q.G., 2007. A mixed methods investigation of mixed methods sampling designs in social and health science research. *Journal of mixed methods research*, 1(3), pp.267-294.
- Conley, D.T. 1987. Critical attributes of effective evaluation systems. *Educational Leadership*, 44(7), 60-64.
- Cook, J.W. and Tonurist, P., 2016. From Transactional to Strategic: Systems approaches to public service challenges. Retrieved from: https://www.oecd.org/media/oecdorg/satellitesites/opsi/contents/images/h2020_systemsthinking-fin.pdf Accessed on 15/01/19
- Crano, W.D. and Brewer, M.B., 2002. *Principles and methods of social research*. 2nd ed. New Jersey: Lawrence Erlbaum Associates Publishers, Inc.
- Creswell, J.W., 2008. The selection of a research design. *Research Design- Qualitative, Quantitative, and Mixed Method Approaches-*, pp.3-22.
- Creswell, J.W., 2014 *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 4th ed. Thousand Oaks: Sage Publications, Inc.
- Creswell, J.W. and Plano Clark, V.L., 2011. *Designing and conducting mixed research methods*. 2nd ed. Thousand Oaks: Sage Publications, Inc.
- Creswell, J.W., 2003. *Research design: Qualitative, quantitative, and mixed methods design*. 2nd ed. Thousand Oaks: Sage Publications, Inc.
- Creswell, J.W., 2013. *Qualitative inquiry and research design: Choosing among five approaches*. 3rd ed. Thousand Oaks: Sage Publications, Inc.
- Creswell, J.W., 2014. *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks: Sage Publications, Inc.
- Creswell, J.W., 2015. *A concise introduction to mixed methods research*. Thousand Oaks: Sage Publications, Inc.
- Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3), pp.297-334.
- Crowe, M., Inder, M. and Porter, R., 2015. Conducting qualitative research in mental health: Thematic and content analyses. *Australian & New Zealand Journal of Psychiatry*, 49(7), pp.616-623.
- Cutcliffe, J.R., 2003. Reconsidering reflexivity: Introducing the case for intellectual entrepreneurship. *Qualitative health research*, 13(1), pp.136-148.
- Danielson, C. and McGreal, T.L., 2000. *Teacher evaluation to enhance professional practice*. Ascd.
- Danielson, C., 2011. Evaluations that help teachers learn. *Educational leadership*, 68(4), pp.35-39.
- Danielson, C., 2013. The Framework for Teaching Evaluation Instrument, 2013 Instructionally Focused Edition. Retrieved from: <http://www.loccsd.ca/~div15/wp-content/uploads/2015/09/2013-framework-for-teaching-evaluation-instrument.pdf> Accessed 5/10/18.

Darling-Hammond, L., 2013. *Getting teacher evaluation right: What really matters for effectiveness and improvement*. San Francisco: Jossey-Bass.

Darling-Hammond, L. 2000. Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis*, 8(1), pp.15-17.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. and Rothstein, J., 2012. Evaluating teacher evaluation. *Phi Delta Kappan*, 93(6), pp.8-15.

Darling-Hammond, L., Cook, C., Jaquith, A. and Hamilton, M., 2012. *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

Dawson, C., 2002. *Practical research methods*. Oxford: How To Books Ltd.

Denscombe, M., 2008. Communities of practice: A research paradigm for the mixed methods approach. *Journal of mixed methods research*, 2(3), pp.270-283.

Dent, E.B. and Umpleby, S.A., 1998. Underlying assumptions of several traditions in systems theory and cybernetics. *Cybernetics and systems*, 29, pp.513-18.

Derrington, M.L. and Campbell, J.W., 2015. Implementing new teacher evaluation systems: Principals' concerns and supervisor support. *Journal of educational change*, 16(3), pp.305-326.

DiCicco-Bloom, B. and Crabtree, B.F., 2006. The qualitative research interview. *Medical education*, 40(4), pp.314-321.

Donaldson, M.L., 2009. So long, Lake Wobegon? Using teacher evaluation to raise teacher quality. *Center for American Progress*, pp.1-32.

Driscoll, D.L., Appiah-Yeboah, A., Salib, P. and Rupert, D.J., 2007. Merging qualitative and quantitative data in mixed methods research: How to and why not. *Ecological and environmental Anthropology*, 3(1). Retrieved from: <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article%3D1012&context%3Ddicwdmee&sa=1> Accessed on 19/02/16.

Eisenhart, M., 1991. Conceptual frameworks for research circa 1991: Ideas from a cultural anthropologist; implications for mathematics education researchers. *Paper presented at the Psychology of Mathematics education-North America meeting, October 16-19, 1991*.

Elliott, K., 2015. Teacher Performance Appraisal: More about Performance or Development?. *Australian Journal of teacher education*, 40(9), p.n9.

Feilzer, Y.M., 2010. Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of mixed methods research*, 4(1), pp.6-16.

Ferris, G.R., Munyon, T.P., Basik, K. and Buckley, M.R., 2008. The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management Review*, 18(3), pp.146-163.

Finnigan, K.S. and Gross, B., 2007. Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal*, 44(3), pp.594-630.

- Frase, L.E. and Streshly, W., 1994. Lack of accuracy, feedback, and commitment in teacher evaluation. *Journal of Personnel Evaluation in Education*, 8(1), pp.47-57.
- Freshwater, D. and Cahill, J. 2014. 'Mixed method research' in Mantosh-Scott, A., Mason, T., Mason-Whitehead, E., and Coyle, D. (eds.) *Key concepts in nursing and health care research*. Thousand Oaks: Sage Publications, Inc.
- Goddard, W. and Melville, S., 2004. *Research methodology: An introduction*. 2nd ed. Lansdowne: Juta and Company Ltd.
- Goe, L., Bell, C. and Little, O., 2008. Approaches to Evaluating Teacher Effectiveness: A Research Synthesis. *National Comprehensive Center for Teacher Quality*. Retrieved from: www.tgsources.org Accessed on 26/08/17.
- Goe, L., Holdheide, L. and Miller, T., 2014. Practical guide to designing comprehensive teacher evaluation systems, revised edition. Retrieved from: <https://files.eric.ed.gov> Accessed on 27/11/16.
- Golafshani, N., 2003. Understanding reliability and validity in qualitative research. *The qualitative report*, 8(4), pp.597-606.
- Goldkuhl, G., 2012. Pragmatism vs interpretivism in qualitative information systems research. *European journal of information systems*, 21(2), pp.135-146.
- Grant, C. and Osanloo, A., 2014. Understanding, selecting, and integrating a theoretical framework in dissertation research: Creating the blueprint for your "house". *Administrative Issues Journal*, 4(2), pp.12-26.
- Greene, J.C., 2008. Is mixed methods social inquiry a distinctive methodology?. *Journal of mixed methods research*, 2(1), pp.7-22.
- Greene, J.C., Kreider, H. and Mayer, E., 2005. Combining qualitative and quantitative methods in social inquiry. *Research methods in the social sciences*, pp.274-281.
- Greener, S., 2008. *Business research methods*. BookBoon.
- Grix, J., 2002. Introducing students to the generic terminology of social research. *Politics*, 22(3), pp.175-186.
- Guba, E. G., & Lincoln, Y. S., 1994. Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications, Inc, pp. 105-117.
- Guest, G., Bunce, A. and Johnson, L., 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods*, 18(1), pp.59-82.
- Gutuza, R.F., 2016. An investigation on the challenges faced by schools in implementation of the Result Based Management system in Mutasa district of Zimbabwe. *Global Journal of Advanced Research*, 3(5), pp.389-396.
- Haefele, D.L., 1993. Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education*, 7(1), pp.21-31.
- Hattie, J. and Timperley, H., 2007. The power of feedback. *Review of educational research*, 77(1), pp.81-112.

Heneman III, H.G., Milanowski, A., Kimball, S. and Odden, A., 2006. Standards-based teacher evaluation as a foundation for knowledge-and skill-based pay. *CPRE Policy Briefs*. Retrieved from http://repository.upenn.edu/cpre_policybriefs/33 Accessed on 5/10/18.

Hinchey, P.H., 2010. Getting Teacher Assessment Right: What Policymakers Can Learn from Research. *National Education Policy Center*. Retrieved from: <http://nepc.colorado.edu> Accessed on 28/08/17.

Hull, J., 2013. Trends in teacher evaluation: How states are measuring teacher performance. *Center for Public Education*. Retrieved from: <http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/Trends-in-Teacher> Accessed on 20/10/15.

Hunter, M., 1988. Effecting a reconciliation between supervision and evaluation: A reply to Popham. *Journal of Personnel Evaluation in Education*, 1(3), pp.275-279.

Hwang, S., 2008. Utilizing qualitative data analysis software: A review of Atlas.ti. *Social Science Computer Review*, 26(4), pp.519-527.

Imenda, S., 2014. Is there a conceptual difference between theoretical and conceptual frameworks?. *Journal of Social Sciences*, 38(2), pp.185-195.

Ingvarson, L., 2002. Strengthening the profession? A comparison of recent reforms in the UK and the USA. Retrieved from: http://research.acer.edu.au/teaching_standards/10 Accessed on 04/10/17.

Isoré, M., 2009. "Teacher Evaluation: Current Practices in OECD Countries and a Literature Review", *OECD Education Working Papers*, No. 23, Paris: OECD Publishing. Available from: <http://dx.doi.org/10.1787/223283631428> Accessed on 4/10/17.

Ivankova, N.V., Creswell, J. W., and Plano Clark, V. L., 2016. 'Foundations and approaches to mixed method research' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 305-336.

Jabareen, Y., 2009. Building a conceptual framework: philosophy, definitions, and procedure. *International journal of qualitative methods*, 8(4), pp.49-62.

Johnson, B.L. 1997. An organizational analysis of multiple perspectives of effective teaching: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, pp. 69-87.

Johnson, R.B. and Onwuegbuzie, A.J., 2004. Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), pp.14-26.

Johnson, R.B., Onwuegbuzie, A.J. and Turner, L.A., 2007. Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), pp.112-133.

Joint Committee on Standards for Educational Evaluation, 1994. *The program evaluation standards: How to assess evaluations of educational programs*. Sage Publications.

Kang, N.H., 2013. Teacher evaluation policy development in South Korea. In *Teacher reforms around the world: Implementations and outcomes* (pp. 147-177). Emerald Group Publishing Limited.

Kanyongo, G.Y., 2005. Zimbabwe's Public Education System Reforms: Successes and Challenges. *International Education Journal*, 6(1), pp.65-74.

Kennedy, M. (2010). The uncertain relationship between teacher assessment and teacher quality. In M.M. Kennedy (Ed.), *Teacher assessment and the quest for teacher quality: A handbook*. San Francisco: Jossey-Bass, pp. 1-6.

Kim, P.S., 2002. Cultural Change in Government: Promoting a High-Performance Culture A Review of Ten Years of Modernisation: The HRM Perspective. In *Human Resources Management (HRM) Working Party Meeting Paris: OECD*. Retrieved from: <http://www.oilis.oecd.org/olis/2002doc.nsf/LinkTo/PUMA-HRM> (2002) (Vol. 11). Accessed on 17/02/17

Kimberlin, C.L. and Winterstein, A.G., 2008. Validity and reliability of measurement instruments used in research. *American journal of health-system pharmacy*, 65(23), pp.2276-2284.

King, L., 2007. SET—the student evaluation of teaching staff-in secondary education. Retrieved from: <http://www.taolearn.com/articles/article17.pdf&sa..6> Accessed on 17/02/16.

Konopásek, Z., 2007. Making thinking visible with Atlas. ti: Computer assisted qualitative analysis as textual practices. *Historical Social Research, Supplement*, (19), pp.276-298.

Korthagen, F.A., 2004. In search of the essence of a good teacher: Towards a more holistic approach in teacher education. *Teaching and teacher education*, 20(1), pp.77-97.

Kothari, C., R., and Garg, G. 2014. *Research methodology methods and technique*. 3rd Ed. New Delhi: New Age International Limited Publishers.

Krauss, S.E., 2005. Research paradigms and meaning making: A primer. *The qualitative report*, 10(4), pp.758-770.

Krejcie, R.V. and Morgan, D.W., 1970. Determining sample size for research activities. *Educational and psychological measurement*, 30(3), pp.607-610.

Kusek, J.Z. and Rist, R.C., 2004. *Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners*. Retrieved from: <https://openknowledge.worldbank.org/bitstream/handle/10986/14926/296720PAPER0100steps.pdf> Accessed on 19/12/17

Leedy, P.D. and Ormrod, J.E., 2010. *Practical Research: Planning and Design*. Boston: Pearson Education Limited.

Leedy, P.D. and Ormrod, J.E., 2014. *Practical Research: Planning and Design*. London: Pearson Education Limited.

Lehman, W.E., Greener, J.M. and Simpson, D.D., 2002. Assessing organizational readiness for change. *Journal of substance abuse treatment*, 22(4), pp.197-209.

Lester, F.K., 2005. On the theoretical, conceptual, and philosophical foundations for research in mathematics education. *Zdm*, 37(6), pp.457-467.

- Lewis, R.B., 2004. NVivo 2.0 and ATLAS. ti 5.0: A comparative review of two popular qualitative data-analysis programs. *Field methods*, 16(4), pp.439-464.
- Liehr, P. & Smith, M.J.1999.Middle Range Theory: Spinning Research and Practice to Create Knowledge for the New Millennium, Advances in Nursing Science, Vol.21, No.4 pp. 81-91.
- Lincoln, Y.S. and Guba, E.G., 1985. *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Liu, S. and Zhao, D., 2013. Teacher evaluation in China: latest trends and future directions. *Educational Assessment, Evaluation and Accountability*, 25(3), pp.231-250.
- Liu, S., Zhao, D. and Xie, W., 2016. Chinese teachers' attitudes toward performance pay: the cases of three schools. *International Journal of Educational Management*, 30(6), pp.791-808.
- Locke, E.A. and Latham, G.P., 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9), pp.705-717.
- Looney, J., 2011. Developing High-Quality Teachers: teacher evaluation for improvement. *European Journal of Education*, 46(4), pp.440-455.
- López, M.G. and Tashakkori, A., 2006. Differential outcomes of two bilingual education programs on English language learners. *Bilingual Research Journal*, 30(1), pp.123-145.
- Lunenburg, F.C., 2010. Schools as open systems. *Schooling*, 1(1), pp.1-5.
- Machingambi, S., 2013. Teachers' perceptions on the implementation of the performance management system in Zimbabwe. *International Journal of Educational Sciences*, 5(3), pp.217-225.
- Mack, L., 2010. The philosophical underpinnings of educational research. *Polyglossia*, 19, pp. 5-11.
- Mackay, K., 2006. *Institutionalization of monitoring and evaluation systems to improve public sector management*. World Bank. Retrieved from: http://siteresources.worldbank.org/INTEVACAPDEV/Resources/4585664-1254406777526/monitoring_evaluation_psm.pdf Accessed on 22/12/2017.
- Madhekeni, A., 2012. Implementing results-based management systems in Zimbabwe: Context and implications for the public sector. *International Journal of Humanities and Social Science*, 2(8), pp.122-129.
- Malterud, K., Siersma, V.D. and Guassora, A.D., 2016. Sample size in qualitative interview studies: guided by information power. *Qualitative health research*, 26(13), pp.1753-1760.
- Maree, K. and Pietersen, J., 2016. 'Sampling' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 191-202.
- Maree, K. and Pietersen, J., 2016. 'Surveys and the use of questionnaire' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 174-190.

- Marshall, C. & Rossman, G. B. 1995. *Designing Qualitative Research*. London: Sage Publications, Inc.
- Marshall, K., 2005. It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86(10), pp.727-735.
- Marzano, R.J., 2012. Teacher evaluation. *Educational Leadership*, 70(3), pp.14-19.
- Matunhu, J. and Matunhu, V., 2014. A review of the implementation of public service reforms by Zimbabwe Revenue Authority (2011). *Journal of Social Sciences*, 38(1), pp.63-70.
- Mavhiki, S., Nyamwanza, T. and Dhoru, L., 2013. An evaluation of RBM implementation in the civil service sector in Zimbabwe. *European Journal of Business and Management*, 5(32), pp.135-139.
- Maxwell, J.A., 2008. 'Designing a qualitative study' in Bickman, L. and Rog, D.J. (eds) *The SAGE handbook of applied social research methods*. Sage publications, pp. 214-253.
- Mayne, J., 2007. Challenges and lessons in implementing results-based management. *Evaluation*, 13(1), pp.87-109.
- Mayne, J., 2010. Building an Evaluative culture: The key to effective evaluation and results management. *The Canadian Journal of Program Evaluation*, 24(2), pp.1-30.
- McGaghie, W.C., Bordage, G. and Shea, J.A., 2001. Problem statement, conceptual framework, and research question. *Academic Medicine*, 76(9), pp.923-924.
- McLaughlin, M.W. (1990). Embracing contraries: Implementing and sustaining teacher evaluation. In J. Millman, and L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers*. Newbury Park, CA: Sage Publications.
- Miles, M., and A. Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. 2nd Ed. Thousand Oaks, CA: Sage.
- Miles, M.B., Huberman, A.M. and Saldana, J., 2014. *Qualitative data analysis: A methods sourcebook*. 3rd Ed. Thousand Oaks, CA: Sage Publications, Inc
- Morgan, D.L., 2014. *Integrating qualitative and quantitative methods: A pragmatic approach*. Thousand Oaks: Sage Publications, Inc.
- Morse, J.M., Barrett, M., Mayan, M., Olson, K. and Spiers, J., 2002. Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, 1(2), pp.13-22.
- Muijs, D. 2004. *Doing Quantitative Research in Education with SPSS*. London:Sage.
- Muralidharan, K. and Sundararaman, V., 2011. Teacher opinions on performance pay: Evidence from India. *Economics of Education Review*, 30(3), pp.394-403.
- Murphy, J., Hallinger, P. and Heck, R.H., 2013. Leading via teacher evaluation: The case of the missing clothes?. *Educational Researcher*, 42(6), pp.349-354.

Musingafi, M.C., Dumbu, E., and Chadamoyo, P., 2013. Improving Performance in the African Civil Service: Empirical Evidence from Zimbabwe. *Public Policy and Administration Research*, 3(2), pp.19-26.

Neuman, L.W., 2014. *Social Research Methods: Qualitative and quantitative approaches*. 7th ed. London: Pearson Education Limited.

Ngulube, P., Mathipa, E.R. and Gumbo, M.T., 2015. Theoretical and conceptual frameworks in the social and management sciences. *Addressing research challenges: Making headway in developing researchers*, pp.43-66.

Nieuwenhuis, J., 2016. 'Analysing qualitative data ' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 103-131.

Nieuwenhuis, J., 2016. 'Qualitative research designs and data gathering techniques ' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 71-102.

OCHA. 2012. Midlands Province : Natural Farming Regions. [Online] Available from: <http://ochaonline.un.org/zimbabwe> [Accessed on 17/09/18].

OCHA. 2012. Midlands Province: Natural Farming Regions [Online] Available from: <https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/> Accessed on 08/10/18.

Olatoye, R.A. and Aanu, E.M., 2011. Senior secondary school science teachers' perception of using students to evaluate teaching effectiveness. *Journal of Emerging Trends in Educational Research and Policy Studies*, 2(3), pp.164-170.

Onwuegbuzie, A.J. and Collins, K.M., 2007. A typology of mixed methods sampling designs in social science research. *The qualitative report*, 12(2), pp.281-316.

Onwuegbuzie, A.J. and Johnson, R.B., 2006. The validity issue in mixed research. *Research in the Schools*, 13(1), pp.48-63.

Onwuegbuzie, A.J. and Johnson, R.B., 2006. The validity issue in mixed research. *Research in the Schools*, 13(1), pp.48-63.

Onwuegbuzie, A.J. and Leech, N.L., 2005. On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International journal of social research methodology*, 8(5), pp.375-387.

Orb, A., Eisenhauer, L. and Wynaden, D., 2001. Ethics in qualitative research. *Journal of nursing scholarship*, 33(1), pp.93-96.

Pandit, N.R., 1996. The creation of theory: A recent application of the grounded theory method. *The qualitative report*, 2(4), pp.1-15.

Patton, M.Q., 2002. Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative social work*, 1(3), pp.261-283.

Peterson, K.D. , 2000. *Teacher Evaluation: A Comprehensive Guide to New Directions and Practices*, 2nd ed. Thousand Oaks, CA: Corwin Press.

Pham, L.D., Nguyen, T.D. and Springer, M.G., 2017. Teacher merit pay and student test scores: A meta-analysis. *Vanderbilt University*, April, 3.

Pietersen, J. and Maree, K., 2016. 'Standardisation of a questionnaire ' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 237-247.

Pietersen, J., and Maree, K., 2016. 'Overview of some of the most popular statistical techniques' in Maree, K. (ed) *First steps in research 2*. 2nd Edition. Pretoria: Van Schaik Publishers, pp. 249-304.

Plano Clark, V.L., Huddleston-Casas, C.A., Churchill, S.L., O'Neil Green, D. and Garrett, A.L., 2008. Mixed methods approaches in family science research. *Journal of Family Issues*, 29(11), pp.1543-1566.

Popham, W.J., 1988. The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education*, 1(3), pp.269-273.

Popkewitz, T.S., Tabachnick, B.R. and Zeichner, K.M., 1979. Dulling the senses: Research in teacher education. *Journal of Teacher Education*, 30(5), pp.52-60.

Pretorius, S.G. and Ngwenya, V.C., 2008. Teachers' perceptions of and attitudes towards performance appraisal in Zimbabwean schools. *Africa Education Review*, 5(1), pp.144-164.

Public Service Commission. 2000. Public Service Regulations: Statutory Instrument 1 of 2000. Harare: Government Printers.

Quinlan, C., Babin, B.J., Carr, J., Griffin, M. and Zikmund, W.G., 2011. *Business research methods*. Hampshire: South-Western Cengage Learning.

Rajasekar, S., Philominathan, K., and Chinnathambi, V., 2013. Research methodology. [Online] Available from: <https://arxiv.org/pdf/physics/0601009.pdf> Accessed on 02/02/2018.

Rao, V. and Woolcock, M., 2003. Integrating qualitative and quantitative approaches in program evaluation. *The impact of economic policies on poverty and income distribution: Evaluation techniques and tools*, pp.165-190.

Riordan, J., Shakman, K., Chang, Q., Lacireno-Paquet, N. and Bocala, C., 2015. Redesigning Teacher Evaluations: Lessons from a Pilot Implementation. Stated Briefly. REL 2016-101. *Regional Educational Laboratory Northeast & Islands*. pp.1-9.

Rogers, G. and Badham, L., 1992. *Getting Started on Training and Implementation*. London: Routledge.

Rosen, R. and Parise, L.M., 2017. Using Evaluation Systems for Teacher Improvement: Are School Districts Ready to Meet New Federal Goals?. *MDRC*.

Sandelowski, M., 2000. Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in nursing & health*, 23(3), pp.246-255.

Sanders, W.L., Wright, S.P. and Horn, S.P., 1997. Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of personnel evaluation in education*, 11(1), pp.57-67.

Santiago, P., Benavides, F., Danielson, C., Goe, L. and Nusche, D., 2013. Teacher evaluation in Chile. *OECD Reviews of Evaluation and Assessment in Education*). Santiago de Chile: OCDE.

Santiago, P. and Benavides, F., 2009. Teacher evaluation: A conceptual framework and examples of country practices. *Paper for presentation at the OECD Mexico*. Retrieved from: <http://www.oecd.org/edu/school/44568106.pdf> Accessed on 17/09/17

Saunders, M., Lewis, P. & Thornhill, A., 2009. *Research Methods for Business Students*, 5th Ed. England: Pearson Education Limited.

Save the Children (n.d). *Monitoring, Evaluation, accountability and Learning (MEAL: 6 Methods of Data collection and Analysis*. Retrieved from: <http://www.open.edu/openlearnworks/moa/resource/view.phd%3Fid%3D52658&sa> [Accessed on 19/02/16].

Scholtes, V.A., Terwee, C.B. and Poolman, R.W., 2011. What makes a measurement instrument valid and reliable?. *Injury*, 42(3), pp.236-240.

Scriven, M., 1994. Duties of the teacher. *Journal of personnel evaluation in education*, 8(2), pp.151-184.

Seo, K., 2012. Lessons from Korea. *Educational Leadership*, 70(3), pp.75-78.

Senge, P.M., 2006. *The fifth discipline: The art and practice of the learning organization* (Rev. ed.). New York, NY: Currency Doubleday.

Shenton, A.K., 2004. Strategies for ensuring trustworthiness in qualitative research projects. *Education for information*, 22(2), pp.63-75.

Shinkfield, A. J. and Stufflebeam, D.L., 1995. *Teacher Evaluation: Guide to Effective Practice*. Boston: Kluwer Academic Publishers. Retrieved from: <http://files.eric.ed.gov/fulltext/ED435632.pdf> Accessed on 17/09/17.

Silverman, D., 2005. *Doing qualitative research: A practical handbook*. 2nd ed. London: Sage Publications, Inc.

Smit, B., 2002. Atlas. ti for qualitative data analysis. *Perspectives in Education*, 20(3), pp.65-75.

Smith, K., 2005. Teacher educators' expertise: what do novice teachers and teacher educators say? *Teaching and Teacher Education*, 21(2), pp.177-192.

Smith, K.G., Locke, E.A. and Barry, D., 1990. Goal setting, planning, and organizational performance: An experimental simulation. *Organizational Behavior and Human Decision Processes*, 46(1), pp.118-134.

Smylie, M.A., 2014. Teacher Evaluation and the Problem of Professional Development. *Mid-Western Educational Researcher*, 26(2).

Statistics Canada. 2010. Survey methods and practices. Retrieved from: <http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf> Accessed on 24/05/18

Stiggins, R.J. and Duke, D.L., 1988. *The case for commitment to teacher growth: Research on teacher evaluation*. New York: State University of New York Press .

Stronge, J.H., 2011. 'Teacher evaluation and school improvement: Improving the educational landscape' in Stronge, J.H., *Evaluating teaching*, pp.2-24.

Stronge, J.H. and Tucker, P.D., 2003. *Handbook on Teacher Evaluation: Assessing and Improving Performance*. Larchmont, NY: Eye On Education.

Stronge, J.H. and Tucker, P.D., 1999. The politics of teacher evaluation: A case study of new system design and implementation. *Journal of Personnel Evaluation in Education*, 13(4), pp.339-359.

Stronge, J.H., 1995. Balancing individual and institutional goals in educational personnel evaluation: A conceptual framework. *Studies in Educational Evaluation*, 21(2), pp.131-151.

Stufflebeam, D.L. and Coryn, C.L., 2014. *Evaluation theory, models, and applications*. 2nd ed. San Francisco: Jossey Bass.

Stufflebeam, D.L. and Shinkfield, A.J., 1988. *Systematic evaluation: A self-instructional guide to theory and practice*. 3rd ed. Boston: Kluwer-Nijhoff Publishing.

Svinicki, M.D., 2010. A guidebook on conceptual frameworks for research in engineering education. *Rigorous Research in Engineering Education*, pp.1-53.

Taut, S. and Sun, Y., 2014. The development and implementation of a national, standards-based, multi-method teacher performance assessment system in Chile. *Education policy analysis archives*, 22(71), pp.1-33.

Teddlie, C. and Tashakkori, A., 2009. *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks: Sage Publications, Inc.

Teddlie, C. and Yu, F., 2007. Mixed methods sampling: A typology with examples. *Journal of mixed methods research*, 1(1), pp.77-100.

The New Teacher Project (TNTP). 2010. *Evaluation 2.0* [Online] Available from: www.tntp.org/assets/documents/Teacher-Evaluation-Oct10F.pdf Accessed on 18/10/15.

Tucker, P.D and DeSander, M.K., 2011. Legal Considerations in Designing Teacher Evaluation Systems. In Stronge, J.H. *Evaluating Teaching*. Thousand Oaks: SAGE Publications, Inc

Tuli, F., 2010. The basis of distinction between qualitative and quantitative research in social science: Reflection on ontological, epistemological and methodological perspectives. *Ethiopian Journal of Education and Sciences*, 6(1), pp.97-108.

Tuytens, M. and Devos, G., 2014. How to activate teachers through teacher evaluation?. *School effectiveness and school improvement*, 25(4), pp.509-530.

UNESCO, 2015. Results-based programming, management, monitoring and reporting (RBM) approach as applied at UNESCO: Guiding principles [Online] Available from: <http://unesdoc.unesco.org/images/0017/001775/177568E.pdf> Accessed on 19/12/17

- Van Griensven, H., Moore, A.P. and Hall, V., 2014. Mixed methods research—The best of both worlds?. *Manual therapy*, 19(5), pp.367-371.
- Wang, L., Lai, M. and Lo, L.N.K., 2014. Teacher professionalism under the recent reform of performance pay in Mainland China. *Prospects*, 44(3), pp.429-443.
- White, B.R., Cowhy, J., Stevens, W.D. and Sporte, S.E., 2012. *Designing and Implementing the Next Generation of Teacher Evaluation Systems: Lessons Learned from Case Studies in Five Illinois Districts. Research Brief*. Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.
- Williams, J.H. and Engel, L.C., 2012. How do other countries evaluate teachers?. *Phi Delta Kappan*, 94(4), pp.53-57.
- Wilson, N. and McClean, S., 1994. *Questionnaire design: A practical introduction*. Newtownabbey: University of Ulster.
- Yilmaz, K., 2013. Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education*, 48(2), pp.311-325.
- Yin, R. K., 2016. *Qualitative Research from Start to Finish*. 2nd ed. New York: The Guilford Press.
- Yoo, J., 2018. Evaluating the new teacher evaluation system in South Korea: Case studies of successful implementation, adaptation, and transformation of mandated policy. *Policy Futures in Education*, 16(3), pp.277-290.
- Zhang, X.F. and Ng, H.M., 2017. An effective model of teacher appraisal: Evidence from secondary schools in Shanghai, China. *Educational Management Administration & Leadership*, 45(2), pp.196-218.
- Zigora, P., and Chigwamba, S., 2000. Management of Human Resources in Africa: Challenges for the Third Millenium. *Paper presented in Workshop on Performance Management in the Public Service*, Harare, Public Service of Zimbabwe, August 14 to 17, 1999.
- Zimbabwe Government. 2005. National Acton Plan of Zimbabwe: Education for all towards 2015. [Online] Available from: <http://uil.unesco.org/i/doc/lifelong-learning/policies/zimbabwe-national-action-plan-of-zimbabwe-education-for-lifelong-learning-towards-2015.pdf> Accessed on 12/06/18.
- ZIMSTAT. 2012. Provincial Report: Midlands. [Online]. Available from: <http://www.zimstat.co.zw/sites/default/files/img/publications/Census/CensusResults2012/Midlands.pdf> Accessed on 12/06/18.
- Zvavahera, P., 2014. An evaluation of the effectiveness of performance management systems on service delivery in the Zimbabwean civil service. *Journal of Management and Marketing Research*, 14, pp.1-8.
- Zvobgo, R.J., 1998. *The Post-colonial State and Educational Reform:(Zimbabwe, Zambia, and Botswana)*. Zimbabwe Publishing House.

Appendices

Appendix 1: University of Venda Higher Degrees Committee project approval

UNIVERSITY OF VENDA

OFFICE OF THE DEPUTY VICE-CHANCELLOR: ACADEMIC

TO : MR/MS B.R MUSODZA
SCHOOL OF EDUCATION

FROM: PROF J.E. CRAFFORD
DEPUTY VICE-CHANCELLOR: ACADEMIC

DATE : 18 MAY 2017

DECISIONS TAKEN BY UHDC OF 16TH MAY 2017

Application for approval of Thesis research proposal in Education: **B.R Musodza (16013601)**

Topic: "Effectiveness of the teacher Performance Evaluation System in Kwekwe District of Zimbabwe."

Promoter UNIVEN
Co-promoter UNIVEN

Dr. M. Mpeta
Prof. T. Runhare

UHDC approved Thesis proposal



Prof J.E. CRAFFORD
DEPUTY VICE-CHANCELLOR: ACADEMIC

Appendix 2.1: Recommendations from ethics committee



University of Venda

Directorate of Research and Innovation

Date : 07 August 2017

Research Ethics recommendations on Application for **Mrs BR Musodza** on
Project title *"Effectiveness of the teacher performance evaluation system in Kwekwe
District of Zimbabwe."*

The following comments and recommendations were made

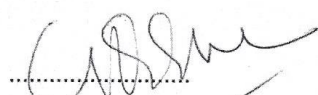
Informed Consent Forms:

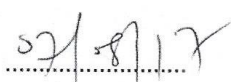
- These forms have been attached and are relevant; however, there is no consent form for the interview schedule. This is an important document and needs to be attached.

Recommendations

The application was provisional approved subject to compliance to the above comments to the satisfaction of the reviewers and supervisors/promoters.

Thank you


.....
PROF G.E. EKOSSE


.....
DATE

Appendix 2.2: Ethical clearance certificate

RESEARCH AND INNOVATION
OFFICE OF THE DIRECTOR

NAME OF RESEARCHER/INVESTIGATOR:

Mrs BR Musodza

Student No:

16013601

PROJECT TITLE: **Effectiveness of the teacher performance evaluation system in Kwekwe District of Zimbabwe.**

PROJECT NO: SEDU/17/CSEM/08/0708

SUPERVISORS/ CO-RESEARCHERS/ CO-INVESTIGATORS

NAME	INSTITUTION & DEPARTMENT	ROLE
Dr M Mpeta	University of Venda	Promoter
Prof T Runhare	University of Venda	Co- Promoter
Mrs BR Musodza	University of Venda	Investigator – Student

ISSUED BY:

UNIVERSITY OF VENDA, RESEARCH ETHICS COMMITTEE

Date Considered: August 2017

Decision by Ethical Clearance Committee Granted

Signature of Chairperson of the Committee:

Name of the Chairperson of the Committee: Prof. G.E. Ekosse



University of Venda

PRIVATE BAG X5050, THOHOYANDOU, 0950. LIMPOPO PROVINCE, SOUTH AFRICA
TELEPHONE (015) 962 8504/8313 FAX (015) 962 9060

"A quality driven financially sustainable, rural-based Comprehensive University"

Appendix 3.1: Permission from Ministry of Primary and Secondary Education of Zimbabwe

*All communications should be addressed to
"The Secretary for Primary and Secondary
Education"*

Telephone: 799914 and 705153

Telegraphic address : "EDUCATION"

Fax: 791923



ZIMBABWE

Reference: C/426/Midlands
Ministry of Primary and
Secondary Education
P.O Box CY 121
Causeway
ZIMBABWE

13 October 2017

Musodza Rindai Belinda
School of Education
P Bag x5050
Thohoyandou
0950
Limpopo

**RE: PERMISSION TO CARRY OUT RESEARCH IN MIDLANDS PROVINCE:
KWEKWE DISTRICT; KWEKWE, MANUNURE, AMAVENI, GLOBEAND PHOENIX,
SHUNGU MARIST BROTHERS, MUNYATI, SIDHAKENI, SILOBELA, LORETO
AND FATIMA HIGH SCHOOLS**

Reference is made to your application to carry out a research in the above mentioned schools in Midlands Province on the research title:

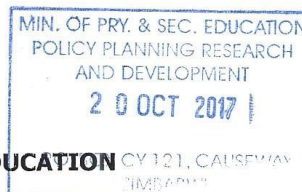
**"EFFECTIVENESS OF TEACHER PERFORMANCE EVALUATION SYSTEM IN
KWEKWE DISTRICT OF ZIMBABWE"**

Permission is hereby granted. However, you are required to liaise with the Provincial Education Director Midlands Province, who is responsible for the schools which you want to involve in your research. You should ensure that your research work does not disrupt the normal operations of the school. Where student are involved, parental consent is required

You are required to provide a copy of your final report to the Secretary for Primary and Secondary Education.

Dr. S.J Utete-Masango

SECRETARY FOR PRIMARY AND SECONDARY EDUCATION



Appendix 3.2: Permission from Provincial office

All communications should be addressed to "The Provincial Education Director"
Telephone: 054- 222460

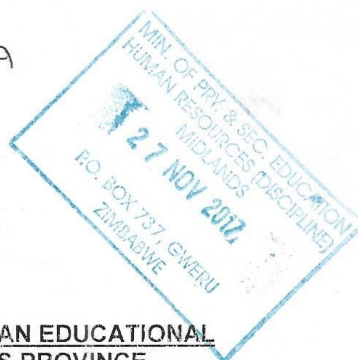
Fax: 054- 226482



Ministry of Primary and Secondary Education
P.O Box 737
GWERU

27 November 2017

Mr/Mrs/Miss: BELINDA R MUSONZA
University of Venda
School of Education
P Bag X 5050 Thohoyandou



Dear Sir/Madam

APPLICATION FOR PERMISSION TO CARRY OUT AN EDUCATIONAL RESEARCH IN SELECTED SCHOOLS IN MIDLANDS PROVINCE

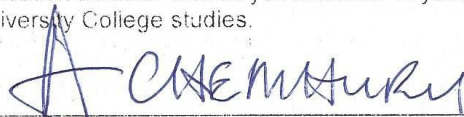
Permission to carry out a Research on:-

"Effectiveness of the teacher performance evaluation system in Kwekwe District of Zimbabwe"

In the Midlands Province has been granted on these conditions.

1. That in carrying out this you do not disturb the learning and teaching programmes in schools.
2. That you avail the Ministry of Primary and Secondary Education with a copy of your research findings.
3. That this permission can be withdrawn at anytime by the Provincial Education Director or by any higher officer.

The Education Director wishes you success in your research work and in your University College studies.



For: PROVINCIAL EDUCATION DIRECTOR: MIDLANDS PROVINCE

Appendix 3.3: A letter stamped by the District Schools Inspector

All communications should be addressed to "The Provincial Education Director"
Telephone: 054- 222460.

Fax: 054- 226482



Ministry of Primary and Secondary Education
P.O Box 737
GWERU

27 November 2017

Mr/Mrs/Miss: BELINDA R. MUSONZA

University of Venda

School of Education

P Bag X 5050 Thohoyandou

Dear Sir/Madam

APPLICATION FOR PERMISSION TO CARRY OUT AN EDUCATIONAL
RESEARCH IN SELECTED SCHOOLS IN MIDLANDS PROVINCE

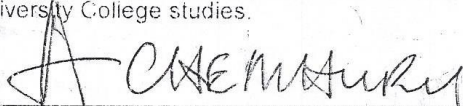
Permission to carry out a Research on:-

"Effectiveness of the teacher performance
evaluation system in Kwekwe District
of Zimbabwe"

In the Midlands Province has been granted on these conditions.

1. That in carrying out this you do not disturb the learning and teaching programmes in schools.
2. That you avail the Ministry of Primary and Secondary Education with a copy of your research findings.
3. That this permission can be withdrawn at anytime by the Provincial Education Director or by any higher officer.

The Education Director wishes you success in your research work and in your University College studies.



For: PROVINCIAL EDUCATION DIRECTOR: MIDLANDS PROVINCE



Appendix 4.1: Consent form for questionnaire

Introduction

My name is Mrs Belinda Rindai Musodza. I am a PhD candidate at the University of Venda and I am conducting a research on the effectiveness of the teacher performance evaluation system in Kwekwe district.

I am requesting you to voluntarily participate in the research by expressing your views on the topic.

I will provide you with full information on what the study is about and how you will participate to enable you to make an informed decision. If you have any questions, you are encouraged to freely ask either by writing your questions to my email (musodzab@gmail.com) or to phone me on the following numbers: +27 73 667 7258 (South Africa) or 0772 938 433 (Zimbabwe). You are free to participate or not to participate in this study. You are encouraged to make your decision on your participation after you have fully understood what the research is all about and if you are happy with what you will do in the study.

What is the purpose of the study?

The main purpose of the study is to gather the views and opinions of teachers, Heads of departments, and School heads on the effectiveness of the effectiveness of the teacher performance evaluation system in the district. The information that is gathered in the study will be analysed to identify the benefits and strengths and at the same time the disadvantages and the weaknesses of the teacher evaluation system and suggest how it can be improved.

How will you participate in the study?

If you agree to participate in the study, you will be asked some questions in a questionnaire where you can express your views. The questionnaire requires less than an hour to complete. You are reminded that the questionnaire is not an examination and therefore as such, you are free to express your views since it is known and accepted that people have different opinions on the subject being studied.

What are your rights as a participant in the study?

Your participation in the study is totally voluntary and therefore you are free to refuse to take part. If you are agree to participate, but later decide against it, then you are allowed to withdraw your participation at any stage and time without having to explain yourself or being asked any question.

What are the risks for your participation in the study?

There is no foreseeable potential of physical harm since your participation will involve only answering questions on the research topic. There is also no possibility of emotional risk if you decide to participate since the study topic is on teacher performance appraisal and has nothing to do with your personal life.

How will your personal privacy and confidentiality be protected?

As mentioned earlier on, no information on your private life will be required. All the views that are gathered in the study will be handled in a confidential manner. First, you will not be required to indicate your name on the questionnaire. Second, the information that is collected will only be used for the study and will only be disclosed to other researchers working on this study.

Has the study got approval from authorities?

Permission to conduct the study in the schools was granted by the University of Venda Research Ethics Committee, the Permanent Secretary of the Ministry of Primary and Secondary Education (Zimbabwe), the Provincial Education Director (Midlands province), the District Inspector (Kwekwe district), and the Head of your school. This does not however take away your right to decide whether or not to participate in the study.

Informed consent declaration form

In terms of the ethical requirements of the University of Venda, I now invite you to complete this form as an indication of your voluntary acceptance to take part in this study.

I..... hereby confirm that I have been fully informed about the purpose, procedures and activities of the study. The rights and risks of my participation have also been fully explained to me, and I have read and fully understood what this study is all about. I was given full opportunity to ask any questions and I understand that I can withdraw my participation from the study at any stage and time without giving any reasons.

I therefore hereby freely give/do not give my consent to participate in this study (Delete inapplicable)

Signature..... Date.....

Researcher's signature..... Date.....

Appendix 4.2 : Consent form for interview

Research title: Effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe.

My name is Belinda R. Musodza and I am a doctoral candidate at the University of Venda. I would like to find out your views about the effectiveness of the teacher performance evaluation system in Kwekwe District of Zimbabwe. The findings from the study will assist in bringing out recommendations that can inform policy and enhance the teacher performance evaluation system which will translate into better learning outcomes for students.

The interview will take about one hour. We don't anticipate that there are any risks associated with your participation, but you have the right to stop the interview or withdraw from the research at any time.

Thank you for agreeing to be interviewed as part of the above research project. Ethical procedures for academic research undertaken from University of Venda require that interviewees explicitly agree to being interviewed and how the information contained in their interview will be used. This consent form is necessary for us to ensure that you understand the purpose of your involvement and that you agree to the conditions of your participation. Would you therefore read the accompanying information sheet and then sign this form to certify that you approve the following:

- The interview will be recorded and a transcript will be produced
- You will be sent the transcript and given the opportunity to correct any factual errors
- The transcript of the interview will be analysed by B.R Musodza as research investigator
- Access to the interview transcript will be limited to B.R. Musodza and academic colleagues and researchers with whom she might collaborate as part of the research process
- Any summary interview content, or direct quotations from the interview, that are made available through academic publication or other academic outlets will be anonymised so that you cannot be identified, and care will be taken to ensure that other information in the interview that could identify yourself is not revealed
- The actual recording will be kept for 6 months after the interview and then destroyed.
- Any variation of the conditions above will only occur with your further explicit approval

Quotation Agreement

I also understand that my words may be quoted directly. With regards to being quoted, please initial next to any of the statements that you agree with:

I wish to review the notes, transcripts, or other data collected during the research pertaining to my participation.

I agree to be quoted directly.

I agree to be quoted directly if my name is not published and a made-up name (pseudonym) is used.

I agree that the researchers may publish documents that contain quotations by me.

All or part of the content of your interview may be used;

- In academic papers, policy papers or news articles
- On our website and in other media that we may produce such as spoken presentations
- On other feedback events
- In an archive of the project as noted above

By signing this form I agree that;

1. I am voluntarily taking part in this project. I understand that I don't have to take part, and I can stop the interview at any time;
2. The transcribed interview or extracts from it may be used as described above;
3. I have read the Information sheet;
4. I don't expect to receive any benefit or payment for my participation;
5. I can request a copy of the transcript of my interview and may make edits I feel necessary to ensure the effectiveness of any agreement made about confidentiality;
6. I have been able to ask any questions I might have, and I understand that I am free to contact the researcher with any questions I may have in the future.

Participant's appointment

Participant's Printed Name.....

Participant Signature..... Date

Researcher Signature..... Date

Appendix 5.1: Questionnaire for the teachers

--	--	--

Title: Effectiveness of the teacher performance evaluation system in Kwekwe district of Zimbabwe

My name is Belinda R. Musodza and I am a doctoral student at the University of Venda. I am interested in finding out your views about the effectiveness of the teacher performance evaluation system in Kwekwe District of Zimbabwe. The findings from the study will assist in bringing out recommendations that can inform policy and enhance the teacher performance evaluation system which will translate into better learning outcomes for learners.

INSTRUCTIONS

There are 2 sections in this questionnaire, sections 1 and 2

In **Section 1**, you will be required to indicate your answer by ticking the appropriate box.

In **Section 2**, you will be required to read a statement and then circle the appropriate number.

SECTION 1. DEMOGRAPHIC INFORMATION

(1). My age is

less than 25 years

☐

26 to 35 years

☐

36 to 45 years

☐

Over 45 years

☐

(2). I am

Female

☐

Male

☐

(3). My highest academic qualification is

Certificate/Diploma in Education

☐

Bachelor of Education

☐

Masters in Education

☐

Other (Please provide name).....

(4). My teaching Experience is

Below 5 years

☐

6 to 10 years

☐

11 to 20 years

☐

Over 20 years

☐

(5). My school is classified as

Urban

☐

Rural

☐

SECTION 2. Indicate your position on the following statements by circling the corresponding number.

Key: 1=Strongly Disagree (SD), 2=Disagree (D), 3= Not sure (NS), 4= Agree (A), 5= Strongly Agree (SA)

	Timing of inception of the performance evaluation system	SD	D	NS	A	SA
6	There was adequate consultation before inception of Results Based Management.	1	2	3	4	5
7	I was trained on Results Based Management System (RBM) before it was introduced.	1	2	3	4	5
8	I was ready for RBM when it was introduced.	1	2	3	4	5
9	I knew all the requirements for RBM.	1	2	3	4	5
10	My school had all the necessary requirements for RBM when it was introduced.	1	2	3	4	5
11	When RBM was introduced, I was not sure of what was expected of me.(reverse coding)	1	2	3	4	5
12	RBM was needed in the education system to appraise teacher performance.	1	2	3	4	5
13	My school head as the manager supported the Introduction of RBM as an appraisal system.	1	2	3	4	5
14	For me, the introduction of RBM was at the right time.	1	2	3	4	5
15	My opinion is that the timing of the inception of the teacher evaluation system has impacted positively on the teacher performance evaluation effectiveness.	1	2	3	4	5

	Design process of teacher performance evaluation system	SD	D	NS	A	SA
16	There was adequate consultation when the teacher performance evaluation system was designed.	1	2	3	4	5
17	The appraisal process at my school is guided by regulations on appraisal in Statutory 1 of 2000.	1	2	3	4	5
18	Regulations on appraisal in Statutory 1 of 2000 are not used at my school. (reverse coding)	1	2	3	4	5
19	The appraisal process is based on clearly stated job descriptions.	1	2	3	4	5
20	RBM is based on clear standards of performance	1	2	3	4	5
21	I know enough about what constitute good teaching as a practitioner.	1	2	3	4	5

	Design process of teacher performance evaluation system	SD	D	NS	A	SA
22	I know the methods that can be used to get accurate performance information.	1	2	3	4	5
23	The RBM appraisal system at my school has rating levels that differentiate well-performing teachers from poor performing teachers.	1	2	3	4	5
24	Teacher performance appraisal at my school is based on classroom observation only. Reverse coding	1	2	3	4	5
25	Learner achievement is used to appraise teacher performance.	1	2	3	4	5
26	The appraiser uses information from different sources to Appraise me.	1	2	3	4	5
27	The teacher appraisal process at my school has clear guidelines.	1	2	3	4	5
28	The appraisal process at my school is well planned	1	2	3	4	5
29	My opinion is that the design of the teacher evaluation system has impacted positively on teacher performance evaluation effectiveness.	1	2	3	4	5

	Implementation process of performance evaluation	SD	D	NS	A	SA
30	At my school, every teacher's performance is appraised.	1	2	3	4	5
31	At my school, performance targets are set and agreed on by both the appraiser and appraisee at the beginning of the appraisal cycle.	1	2	3	4	5
32	At my school, performance is appraised and rated on agreed targets only.	1	2	3	4	5
33	At my school, performance appraisal actually measure the results of performance by the teachers.	1	2	3	4	5
34	There is continuous monitoring of performance during the appraisal cycle at my school.	1	2	3	4	5
35	The appraisal process at my school identifies performance areas that require improvement. Significance	1	2	3	4	5
36	Appraiser and appraisee at my school agree on the strategies to address the identified areas needing improvement. Significance	1	2	3	4	5
37	Training needs that require in-house training are addressed through staff development programs at the school. Significance	1	2	3	4	5
38	Review discussions at my school are based on progress made towards the achievement of performance targets.	1	2	3	4	5
39	I am given opportunity to comment during performance reviews at the school.	1	2	3	4	5
40	I can comment during the appraisal final rating process.	1	2	3	4	5
41	I know the procedure to be taken if I fail to agree with the appraiser on the work plan for the appraisal process.	1	2	3	4	5
42	I know the procedure to be taken if I fail to agree with the appraiser on the final rating.	1	2	3	4	5

	Significance of the appraisal process	SD	D	NS	A	SA
43	The performance appraisal at this school considers all aspects of teaching. Implementation	1	2	3	4	5
44	The person who appraises me at the school is knowledgeable on teacher evaluation. Implementation	1	2	3	4	5
45	My appraiser has adequate time to conduct the appraisal activities. Implementation	1	2	3	4	5
46	The feedback on my appraisal helps me to improve my teaching	1	2	3	4	5
47	The appraisal instrument at my school measures teacher performance accurately. Design	1	2	3	4	5
48	Appraisal results inform decisions on promotion.	1	2	3	4	5
49	Appraisal results inform decisions on regrading.	1	2	3	4	5
50	Appraisal results inform decisions on whether or not one continues to be on probation.	1	2	3	4	5
51	Appraisal results inform decisions on termination.	1	2	3	4	5
52	Outstanding teachers are rewarded through teacher appraisal.	1	2	3	4	5
53	Teachers at my school who perform poorly are mentored by more experienced and well-performing teachers.	1	2	3	4	5
54	Appraised teachers that continue to perform poorly in spite of remediation have their employment terminated.	1	2	3	4	5
55	Appraisal results help to decide what is covered in our school staff development programmes	1	2	3	4	5
56	There is money to support the appraisal process at my school Implementation	1	2	3	4	5
57	We are given time for collaborations as teachers to help us to improve performance.	1	2	3	4	5
58	We are given manpower development support if appraisal results require training that cannot be provided within the school	1	2	3	4	5
59	Teachers at my school are given a chance to say their contributions on how the performance appraisal system can be improved	1	2	3	4	5

	Recommendations	SD	D	NS	A	SA
60	RBM should be removed as a performance appraisal system for teachers.	1	2	3	4	5
61	RBM needs to be improved to suit the Ministry of Education environment.	1	2	3	4	5

THANK YOU FOR YOUR TIME

Appendix 5.2: Interview schedule

Introduction and instructions

I wish to start by thanking you for accepting to be interviewed. Please feel free to say anything bearing in mind that whatever you say here will only be used for the study and that there are no correct or wrong answers.

May you tell me about your professional life from the time you started until now? What have you enjoyed about this journey and what have you hated about it?

1. Timing of introduction of the teacher performance evaluation system

The Results Based Management System has been operational in the Ministry of Primary and secondary education since the year 2005. There were some concerns that were raised concerning the timing of the introduction of the appraisal system.

1.1 What were the concerns raised by you as educators on the timing of the introduction of the RBM system?

1.2 Explain what you found positive about the introduction of RBM.

1.3 What further steps could have been taken to enhance the appropriateness of the timing of inception of RBM?

2. Design process of the performance evaluation system

Design refers to the act of planning the implementation process. It involves deciding on the issues that include: developing an appraisal policy, defining appraisal roles, stating appraisal standards, developing appraisal guidelines, developing appraisal tools, and the reporting instruments, and deciding sources of appraisal information.

2.1 How do you prepare for the appraisal process at your school?

2.2 Explain the way the appraisal system that you are using was developed?

2.3 Statutory Instrument 1 of 2000 by the Public Service Commission contains regulations on appraisal.

2.3.1 Describe what you know about the appraisal regulations in Statutory 1 of 2000?

2.3.2 Elucidate what you find useful in that policy?

2.3.3 Clarify what you think can be added in that policy to enhance its impact on aiding effective implementation of the appraisal process?

2.4 In Results Based Management, you make use of the key result areas to set the performance targets that are used as the basis of the appraisal. Performance targets are set and agreed to by the appraiser and appraisee.

2.4.1 Indicate what you find useful about the usage of performance targets for measuring teacher performance?

2.4.2 What in your view are the weaknesses of using performance targets for measuring teacher performance?

2.5 The Public Service Commission has guidelines on the completion of the appraisal form. Do you use the guidelines when completing the form and do the guidelines help to make the process easier? Please explain your answer.

2.6 Who decides the sources of information used for appraisal at your school?

2.7 What is your comment on the ability of the appraisal form to capture the teacher performance level accurately?

2.8 The appraisal system is said to be results- based, meaning that it is supposed to measure results, would you say that the system actually measures results? Please explain.

2.9 In your opinion, which rating system do you think is more effective, the overall rating of performance or rating the specific teaching areas separately?

2.10 What needs to be added or improved in terms of design of the appraisal system to enhance its effectiveness?

3. Implementation process of the teacher performance evaluation

3.1 How is teacher performance evaluation conducted in your school?

3.2 Describe the strengths and shortcomings of the way performance appraisal is conducted at your school in relation to the regulations on appraisal?

3.3 Clarify the challenges that you find in implementing appraisal process at your school

3.4.1 How would you rate the system in terms of its ability to capture adequately all aspects of teaching?

3.4.2 How would you describe the rating process in terms of objectivity and validity (ability to measure what it is supposed to measure)?

3.5 In what way can the appraisal implementation process be improved?

4. Significance of the performance evaluation system

For a process to be respected it should be significant and have consequences. The significance of the appraisal involves the consequences or the implications of the process.

4.1 What would you say are the consequences of the appraisal process or its significance?

4.2 What measures could be used to enhance the significance of the appraisal process to enhance the effectiveness of the appraisal system?

4.3 In your view, how is teacher performance appraisal beneficial to the following: learners, teachers, Heads of Department, and Heads of schools?

4.4 Do you have any other comment you would like to make in relation to the performance appraisal process at your school?

Thank you for your time.

Appendix 6: Sample of codes and quotations from ATLAS ti.

Project: Musodza PhD Project

Report created by PhD_Student on 10/30/2018

Code Report

(5) codes

Local filters:

Show codes in group RBM development process

● Appraisal system Imposed

10 Quotations:

1:34 . Sorry to say this, the head during one opening meeting only said the..... (2882:3488) - D 1: Participant 1

. Sorry to say this, the head during one opening meeting only said there's now a new system of appraisal. It was imposed on us. When it was imposed on us, we were not given any room for additional or subtractions so it's something which was just imposed (emphasis with hands).P. Grudgingly we accepted it because it's being taken as the policy and we presume that if you do not do it, then you might be forced to resign. So because of the current situation on the ground, someone simply agrees to do it not because of its attachments which are positive but because he/ she will be safeguarding his/ her job

2:33 We were never involved in anything. It was just imposed on us. We were..... (3084:3272) - D 2: Participant 2

We were never involved in anything. It was just imposed on us. We were simply told that there is a new system and we should adhere to the new system. We were never involved as stakeholders.

3:33 I think it's the Ministry of Education, I am not very clear. P: No, I..... (3166:4030) - D 3: Participant 3

I think it's the Ministry of Education, I am not very clear. P: No, I really do not know there, whether teachers were consulted or it was just an imposition that you must just follow this. I am sure that is the case that it was just imposed because if it had been a thing that had involved teachers nationwide then I would definitely know about it and we would not be asking why this thing is not rewarding us. We should have seen its rewards long back so it's an imposition from the ministry. P: Maybe the heads, am not sure because they are the ones that came and tried to educate us on the system. P: I do not think they were involved? P: I think earlier on I mentioned that this thing is not working because teachers were not involved so it must involve teachers so that it is credible and acceptable to the teachers. It should come from the teachers themselves

4:34 I am not sure if there was an involvement of the teachers. If it was t..... (4801:5628) - D 4: Participant 4

I am not sure if there was an involvement of the teachers. If it was there, then it could have been indirect involvement. The ministry is using the top down approach where they just tell you that we are now doing this. They just impose programs on us which is not good at times. It is important to do research like you are doing so that before they can implement they get the input and contributions of the stakeholders on the ground. I don't really think that it was done properly. The system was introduced haphazardly because some bosses up there felt that we were lagging behind as a country. The ministry should have instituted a two way communication and not forced the program down our throat. If there had been two way communication the ministry would have been given feedback on whether the program was feasible or not

6:33 We were not consulted or involved. As I have said before, this thing i..... (1357:1515) - D 6: Participant 6

We were not consulted or involved. As I have said before, this thing is just an animal which was just imposed on us. We do not have any knowledge of what it is

7:36 As far as I know, we were not involved in the design process. Like I s..... (6870:7518) - D 7: Participant 7

As far as I know, we were not involved in the design process. Like I said, we had the KRA system before RBM and out of the blue, we were told that we are now using RBM system, so we were not involved in any way. There are actually some outcomes that are made compulsory for everyone. I may not want to look at that area that I am forced to look at. Remember we are teaching learners with different learning needs and the instrument that we are using does not cater for such learners. I have noticed this because I am a special needs teacher. The outcomes may not really suit the learners that we have with their diverse needs so it is very difficult

8:33 The system was just imposed and we had no input whatsoever (2479:2536) - D 8: Participant 8

The system was just imposed and we had no input whatsoever

9:32 it was just imposed on us. We were just told that we have changed from..... (3648:3779) - D 9: Participant 9

It was just imposed on us. We were just told that we have changed from key result areas to RBM? How it was developed, I am not aware

10:30 No. As I have already alluded to earlier on that with policies normall..... (4232:4792) - D 10: Participant 10

No. As I have already alluded to earlier on that with policies normally, we have a few people at the top who design a program and it is ushered down so the majority even the unions which represent us were not involved. The program was just brought in and whether you like it or not, the program is just forced on people. P: Personally I think if something is forced upon someone, I think it will not be taken with the seriousness it deserves because people will tend not to embrace it since they feel that it was just something that was imposed by the government

11:32 As I have said before, I think to me this was a borrowed idea from the..... (4111:4426) - D 11: Participant 11

As I have said before, I think to me this was a borrowed idea from the industry. It came to us just as policy. We had no input. I think it was done up there. We were simply told how to use it, that is, formulating objectives with supervisor and agreeing on the quarterly reviews. We were simply told and had no input

● **Inappropriate tools developed**

1 Quotations:

2:34 I think it takes someone on the ground to design the tools because if..... (3277:3847) - D 2: Participant 2

I think it takes someone on the ground to design the tools because if someone who is outside tries to bring in their own mechanism, those on the ground might find it very difficult to implement and this can result in a lot of cheating when people try to make it work in spite of its short comings and inappropriateness. Due to lack of time you find teachers and HODs sitting down and just agreeing on the mark to be awarded to the appraisee without even going on the ground to see the work that has been done. Lesson observations will not be done because there is no time

● **No stakeholder involvement**

7 Quotations:

6:34 We were not consulted or involved (1357:1389) - D 6: Participant 6

We were not consulted or involved

7:35 As far as I know, we were not involved in the design process. Like I s..... (6870:7151) - D 7: Participant 7

As far as I know, we were not involved in the design process. Like I said, we had the KRA system before RBM and out of the blue, we were told that we are now using RBM system, so we were not involved in any way. There are actually some outcomes that are made compulsory for everyone

8:32 No we were never involved at all (2445:2476) - D 8: Participant 8

No we were never involved at all

9:33 No, there was no involvement of people on the ground, it wasn't there..... (3785:4051) - D 9: Participant 9

No, there was no involvement of people on the ground, it wasn't there. Because at times you find that the problem with our ministry is that some of these things just come as policy and you will simply be told to go and implement. P: So we are forced just to implement

10:31 People will continue to do it because after its introduction it worked..... (4795:5239) - D 10: Participant 10

People will continue to do it because after its introduction it worked for some few years and only very few people benefited then but otherwise the forms would just pile and pile at the human resource department since there was inadequate staff to work on the forms. I do not think from the school level you can get a negative rating, most of them were quite ok, i.e. they were illegible for the reward but the increment was nowhere to be found

11:32 As I have said before, I think to me this was a borrowed idea from the..... (4111:4426) - D 11: Participant 11

As I have said before, I think to me this was a borrowed idea from the industry. It came to us just as policy. We had no input. I think it was done up there. We were simply told how to use it, that is, formulating objectives with supervisor and agreeing on the quarterly reviews. We were simply told and had no input

12:36 Like I mentioned earlier on, this system was mistimed in that it was n..... (5533:5664) - D 12: Participant 12

Like I mentioned earlier on, this system was mistimed in that it was not involving all the stakeholders in crafting that instrument.

● **Not aware of appraisal**

3 Quotations:

5:30 I don't know (laughs) I simply have to use that instrument but I don't..... (2726:2867) - D 5: Participant 5

I don't know (laughs) I simply have to use that instrument but I don't know how it was developed. P: I am not aware like I said before (laughs

8:31 No am not aware of how the system was developed (2393:2439) - D 8: Participant 8

No am not aware of how the system was developed

9:31 I am not aware, it was just imposed on us (3632:3672) - D 9: Participant 9

I am not aware, it was just imposed on us

● **Stakeholder input later sought**

1 Quotations:

12:37 It has however been improved from the inputs that we gave to the minis..... (5665:6630) - D 12: Participant 12

It has however been improved from the inputs that we gave to the ministry after its inception. P: I think in 2009 thereabout, they created a space on the form that instead of just having a supervisor and supervisee, there was now a reviewer. The reviewer is an inspector of the area who is my immediate superior. Above him is the District Schools Inspector (DSI). The suggested that information should be passed from the supervisor at school level to the Area Inspector who also wrote his/her comments, the DSI who also wrote his/ her comments and then it was sent to the directors where the instrument was created. This is how the information was conveyed from school level to the relevant authorities. P: The comment was open. They did not restrict us. It could be on the problems faced or on how the instrument could be improved and any other suggestion pertaining to the appraisal process. P: Yes but we are not sure whether the inputs were incorporated or not