A STUDY ON THE EFFECTS OF WITHOLDING MARKS AND PROVIDING WRITTEN FEEDBACK FOR GRADE 9 LEARNERS WHO PERFORM POORLY IN ENGLISH HOME LANGUAGE ASSESSMENTS

by

Gwen Barry

817004

A research report submitted to the Faculty of Humanities, School of Education at the University of the Witwatersrand in partial fulfilment of the requirements for the degree of Master of Education.

gclbarry@gmail.com

Contact Number: 084 453 5212

Supervisor: Dr Carola Steinberg

TABLE OF CONTENTS

DECLARATION

I declare that this Research Report is my own, authentic and unaided work. Where external sources have been used they have been properly referenced and acknowledged. This report is being submitted for the degree of Master of Education at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

_____

Gwen Barry

817004

15 February 2017

Johannesburg

ABSTRACT

Assessment in education is an extremely important component for both teachers and learners. In a high-stakes environment where marks are all-important both teachers and learners are under immense pressure for all learners to perform well in summative assessment.

Teachers want to maximise real learning as well as ensure that their learners achieve good marks. Providing well-structured feedback on assessment tasks is a very effective way of doing this. The salience of this feedback can be further enhanced for weak learners by withholding marks and providing feedback only.

This study used an action research model with both quantitative and qualitative data. The conclusions drawn were that when marks are withheld weak learners, in particular, remember and apply comments better. Withholding marks and providing task-oriented comments only results in weak learners improving their work at a faster rate than learners who receive marks and comments together.

DEDICATION

To my mother, friend and mentor

Prof. Jane Spowart

ACKNOWLEDGEMENTS

Thank you to my supervisor, Dr Carola Steinberg, for her patience, insight, guidance and wise counsel in all things.

Thank you to each of the learners in my grade 9 classes from 2016 for their enthusiasm and the pleasure of all that I have learned during my time with them.

Thank you to Dr A. Ceresto for her kind permission to carry out the research at her school.

Thank you to Mrs Geraldine Connell for her language editing skills.

Thank you to all of my friends and family for their unwavering support.

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND CONCEPTS

CAPS- Curriculum and Assessment Policy Statement

FET- Further Education and Training

GDE- Gauteng Department of Education

UMALUSI- Council for Quality Assurance in General and Further Education and Training

CHAPTER 1: OVERVIEW OF THE STUDY

1.1.INTRODUCTION

Assessment is an integral part of schooling. However, in an environment of high-stakes testing and high accountability, assessment has come to represent more than just the extent to which learners achieve learning goals (Madaus and O'Dwyer 1999, p. 686). Assessment results determine access to resources and opportunities in broader society (Gipps, 1999 p. 360) and are, erroneously, understood as an indicator of intelligence (Elwood and Murphy 2015, p. 184).

Tremendous pressure to perform well in summative assessment is placed on learners and teachers alike and this pressure can have detrimental effects on real learning. Adapting to the demands of high-stakes testing can put the focus on performance in assessment rather than on the development of overall understanding. For these reasons, it is important for teachers to find a way of using assessment tasks in such a way that they enhance learners' capacity and understanding. The most effective way to do this may be by withholding marks and using carefully constructed task-involving feedback to enhance interest and understanding.

1.1.1. BACKGROUND TO THE STUDY

There is tremendous pressure on learners and teachers alike to produce good assessment results in the high school environment. English Home Language is a core subject that Grade 9 learners are required to pass with 50% in order to progress to the next grade. As a teacher, I spend a considerable amount of time trying to assist and support academically weak learners in an effort to help them to pass and I am always investigating new strategies to do this.

When referring to academically weak learners, for the purposes of this study, I am referring to learners who do not attain the 50% minimum required to pass English Home Language. These learners are underachieving because they have difficulty constructing good sentences and paragraphs, learning and using nuanced vocabulary, and responding adequately to the demands of a task in general.

I have noticed that weak learners are particularly distressed by their marks and that they really struggle to apply corrections to their writing. In the new Curriculum and Assessment Policy (CAPS) from the Gauteng Department of Education (GDE), writing comprises a large

part of the overall result for Grade 9 English, and writing as a skill influences achievement in almost every other subject. When I read Butler's 1988 study on the differential effects of withholding marks and providing remedial comments, I was interested in trying it in a real-life high-stakes assessment setting to improve my practice as a teacher.

1.1.2. PURPOSE STATEMENT

The purpose of this research is to establish whether the practice of withholding the marks of weaker learners and providing, instead, detailed comments on the quality of their work and ways in which they may improve it, will in fact increase their engagement in their work and thereby improve their marks. This study aims to replicate the practices of Butler's 1989 study, in which an experimental group is provided with comments only and not marks, but in the context of a functional school assessment programme.

This study intends to explore whether withholding marks and providing comments alone to a class of learners with lower academic ability, on writing assessments will improve their interest in the comments provided and thus help them implement the instructions that will in turn improve their marks.

1.2.RESEARCH OBJECTIVES

RESEARCH QUESTION

Will the practice of withholding the marks of at risk Grade 9 learners and providing detailed comments encourage them to make an effort to improve the quality of their work, apply the guidelines suggested by the teacher and thus improve their marks?

Empirical sub-questions:

1. Does withholding marks induce learners to take more interest in and remember a remedial comment?

2. Will learners be encouraged to apply the guidelines of that comment to future assessment tasks and thereby improve their marks?

1.3.RESEARCH HYPOTHESIS

The primary hypothesis in this study was Butler's original hypothesis that intrinsic motivation, and probably self-regulation, will be differentially affected by task-involving and ego-involving evaluation (1988, p.1). Butler (1988, p.2) identified remedial comments as task-involving evaluation and marks as constituting ego-involving evaluation. Brookhart (2008, p.22) and Rubie-Davies, Stephens and Watson (2015, p.12) point out that since learners engage in their own form of norm referencing, by comparing marks, marks can have a negative effect on the self-esteem and performance beliefs of weaker learners. This negative emotional effect can be so pronounced that learners ignore all other types of feedback relating to the task and therefore continue to receive poor marks (Emmer and Stough 1998, p350).

Butler's study found that learners were more motivated to apply task-related feedback in the form of remedial comments when ego-involving feedback (in the form of marks) was absent (1988, p13).

My hypothesis, based on the findings of Butler's study, was that if marks for writing assessments were withheld from weak learners, these learners would pay more attention to the remedial comments provided. By paying more attention to the comments, learners would remember them better and apply them to their work more consciously. All of this would culminate in weaker learners being able to improve their marks.

1.4. LIMITATIONS, ASSUMPTIONS AND SIGNIFICANCE OF THE STUDY

1.4.1. LIMITATIONS

The primary limitation of this study is the small size of the sample, which makes all of the quantitative findings statistically insignificant. Furthermore, the gains made by the experimental group, when measured against those of the control group, were not big enough to be statistically relevant.

1.4.2. ASSUMPTIONS

Going into the study, I assumed that learners would all have an understanding of how to write an English creative essay. I have also assumed that the essay would be administered under test conditions in each class.

1.4.3. SIGNIFICANCE OF THE STUDY

The primary significance of this study is on a qualitative and individual level. Firstly, the quantitative data from the experimental group did not show gains large enough to be significant; the significance of the gains was, however, evident on an individual level. The small, but increasing, gains that individual learners experienced contributed to a growing sense of confidence in their own ability as well as the awareness that taking the comments into consideration did indeed help them to improve.

Secondly, the study was important for me on a qualitative level as a researcher and teacher. It was enriching to go through the process of collecting the data and analysing the results. This process gave me a real insight into the effects that interventions can have on the learners and their performance in tasks. The realisation that a study like this needs to be performed on a much larger scale for it to be statistically relevant and affect any kind of change in the school or larger educational community was a bit disappointing, but it has made me determined to carry out larger projects in the future.

The study also gave me a real insight into the problems that many of my learners were experiencing with writing and helped me to come up with strategies to assist them.

1.5. RESEARCH METHODOLOGY

Butler's (1988) study was designed in order to confirm four experimental hypotheses generated around the various effects of giving task-involving and ego-involving feedback (p. 3). The experimental approach to research design concerns itself directly with causality, "if x, then y (Cohen et. al. 2011, p. 312)". The study carried out by Butler was an experiment in its most classic form, conducted in a contrived, artificial environment (Cohen et.al 2011, p. 315).

The results of Butler's study showed a clear causality between the provision of task-involving feedback, in the form of comments, and learner motivation and performance (Butler 1988, p.

13) however, as Black and William (2006) point out; it lacks ecological validity because it was conducted outside normal classroom conditions (p. 14).

My research focus and questions were inspired by Butler's study, but given the research context of teaching in a school and needing to make adjustments to my own practice, her methodology of a controlled experiment would not be appropriate for this study. So this study uses an action research approach collecting quantitative and qualitative data, which offers both the ecological validity that Butler's Study lacked (Black and William 2006, p. 14) and the opportunity to test Butler's hypothesis and observe the results.

1.6.OVERVIEW OF THE CHAPTERS

Chapter 1: Overview of the study: This chapter introduces the background and concerns of the study. It also presents the problem and purpose of the study, provides the research questions and offers this overview.

Chapter 2: Literature Review: Secondary information was collected from books and academic journals to support the rationale of the study and help to explain the findings.

Chapter 3: Research design and methodology: This chapter explains the use of a mixed methods approach within the paradigm of action research. Data was collected through the recording of marks and comments, observations, questionnaires and a focus group.

Chapter 4: Data presentation: The quantitative and qualitative data collected during the pilot study and three cycles of action research are presented and explained using graphs and a table.

Chapter 5: Analysis of data: Six claims are made that link the findings of the study to the original research question.

Chapter 6: Recommendations and Conclusion: This chapter weighs the claims against each other, draws conclusions and makes recommendations. It also describes two subsequent projects that developed out of the findings of this study.

1.7. SUMMARY

Assessment is an integral part of the current education system in South Africa and it carries very high stakes that determine the future of many learners and teachers. There is no alternative to high-stakes assessment at this point, therefore it is important for teachers to operate within the system yet find ways to use assessments to create real learning opportunities and assist their learners in improving their performance.

Withholding marks so that the learners focus more on the demands of the task and seek additional feedback may be a powerful way for teachers to help learners do the learning that will enable them to perform well in assessment tasks.

2. CHAPTER 2: LITERATURE REVIEW

Assessment and schooling are inseparable. Assessment provides information that is used as a means of access to the institutions of the adult world. The pressures that assessment generates are felt by learners and teachers alike, therefore it is important that it is done in a way that encourages real learning. Feedback from assessment is the key to making the process a learning experience, if it is done correctly. Well-crafted feedback enhances a learner's performance by tapping in to their intrinsic motivation.

2.1.THE ROLE OF ASSESSMENT IN SCHOOLS

Assessment is an integral part of schooling. Madaus and O'Dwyer (1999) use the constructivist perspective to express the two prominent conceptual ideas behind assessment in the educational context. Firstly, assessment can be used to determine the extent of the learning that has taken place and is thus summative in nature (p. 689). This kind of assessment has traditionally been used to establish whether students had acquired minimal proficiency and appropriate skills for promotion (Graue and Johnson 2011, p. 1827). Secondly, they point out that the process of assessment provides opportunities for the student and the teacher to learn, and is therefore formative in nature (p. 689).

The trend in the focus of assessment from 1983 onwards supports this idea of assessment as helpful in orienting the teaching and learning process; assessment began to be used as a means to "push for excellence" and measure outcomes that represented learning (Graue and Johnson 2011, p. 1830). In short, assessment can be used as a form of research in which teachers can find out what learners "know, understand and can do" (Conley et al, 2010, p. 103) and good assessment practises can foster deep and meaningful learning (p. 111).

Assessment is also a means of controlling the access that a growing population has to certain social resources. Gipps (1999) states that assessment is used for certification and selection so that the nature and rewards of mass education can be controlled (p. 360). Foucault (in Scott 2008) takes this notion a step further and suggests that it is through "the examination" and the access that it either grants or denies that the population polices itself. The implication of this is that people are denied access to social and educational resources because they do not meet the achievement criteria of a final summative assessment and they accept that they are not worthy of access to these resources.

Assessment also has its roots in psychometric testing (Madaus and O'Dwyer 1999, p. 693) which links the results of summative assessment to levels of intelligence from a socio-

cultural perspective. Elwood and Murphy (2015) point out that the social world is largely responsible for the idea that the practice and results of summative assessments are of "scientific value" even though the scores obtained by learners are not objective (p. 189). When it comes to psychometric testing in adolescents, Gumbiner (2003) points out that test performance is influenced by many variables – with parenting, not individual intelligence, being the most significant influence of both IQ test scores and academic achievement (p. 39). In the school system, academic tests have been criticised for testing only the learner's ability to write and for disadvantaging many learners because of language barriers (Bloch 2009, p. 66-7).

However, the perception remains amongst parents, universities, potential employers and broader society that assessment results are the only quantifiable means of determining a person's intelligence and therefore making a judgement about accepting or denying access to further education or employment. It is for these reasons that the advice of Shepard (2000) is important: teachers and learners need to be taught how to do assessment well (p. 10).

Due to the predominant conception of assessment as a means of certification and the misconception that the results of assessment are an indicator of intelligence, the formative benefits of assessment can be overlooked. Therefore teachers need to use assessment also to focus on the nature of learning as a process, rather than only as a product and encourage learners to develop a growing capacity, rather than just acquire commodities that aid performance (Swaffield 2011, p. 438).

2.1.1. CERTIFICATION AND CONTROL IN THE SOUTH AFRICAN CONTEXT

The focal point of assessment and certification in South Africa is the National Senior Certificate that marks the end of a learner's school career in Grade 12. However, there is another possible exit level in the system at the end of the Senior Phase in Grade 9.

A pass at the end of Grade 9 enables a learner to either continue with formal schooling in the Further Education and Training (FET) phase at school or to enrol in a FET college. If a learner fails Grade 9 they can only repeat the year once, after which they obtain a condoned pass, which means that they are allowed access to Grade 10 or they exit the school system, but their marks reflect that they have not achieved the minimum requirements. Therefore, the marks that Grade 9 learners achieve control their access to higher levels of education.

Grade 9 learners are required to pass their Home Language in order to be promoted to Grade 10. In 2014, the Gauteng Department of Education increased the mark required to pass from 40% to 50% (Department of Basic Education 2013, p. 23). This has resulted in many learners failing English Home Language who would otherwise have passed under the old requirements. In the school where I teach, which is considered to be a good government school in the province, the number of learners failing English Home Language has increased from 2.2% in 2013 to 3.4% in 2014; an increase of 50%.

2.2.THE PRESSURES AND EFFECTS OF HIGH-STAKES ASSESSMENT

As Grade 9 is an exit level year, the Grade 9 assessments, particularly those for English Home Language, can be considered to carry high stakes. With high stakes, come the pressures of accountability and teachers are under pressure to maintain and improve the pass rate without sacrificing the standards of instruction and assessment. This tension is very similar to what teachers around the world face; teachers need to keep up with and prepare learners for the demands of high stakes accountability testing and develop rich learning and understanding through the combined use of formative and summative assessment (Black 2015, p. 165).

Furthermore, in a system with a heightened sense of accountability, like we have in South Africa, schools are expected to administer provincial exams and are then subject to cluster moderation. The public nature of assessment increases the pressure on teachers and learners to perform in assessment tasks and increases teachers' anxiety about their learners' performance (Steinberg 2013, p.5.). Quality assuring bodies like the Council on Higher Education and the Higher Education Quality Committee as well as independent organisations like UMALUSI also have an interest in assessment so that they can monitor how the educational system in South Africa is working and assure that assessment meets with their standards (Conley et. al 2010, p. 101).

Bloch (2009) highlights the particular pressures of the South African school system when he says that test scores at any level are still recognised as the best indication of how the schooling system has performed (p. 60). He takes this idea further by stating that "South Africa is the worst performer in maths and literacy, probably in the world. (p. 61)". What is extremely valuable about Bloch's insight is not only its comment on how South African learners perform in tests, but the significance of this in light of the objections about testing itself. He acknowledges that tests are criticised for their validity in light of cultural bias and

what is actually being tested. Poverty, a lack of resources and language barriers are generally highlighted as mitigating factors when learners perform poorly in standardised tests. His point is that learners from other countries, who face the same conditions, still out-perform South African learners (p. 66-7). There is no alternative to assessment in schooling at this stage, so it is up to teachers to work with it and do their best to help it benefit the learners that they teach.

2.2.1. The pressure to pass

The learners who achieve results just below the pass requirement are dubbed "bubble kids" by Graue and Johnson (2011) and they point out that schools can become over-focused on these learners, targeting most instructional interventions at them (p. 1832). Furthermore, because accountability is intimately linked to system wide assessment, and assessment is understood as the yardstick by which learning is measured, teachers are implicated when learners fail (Steinberg 2008, p. 55).

When assessment is viewed as a measure of what learners have gained from the school system in terms of knowledge and skills, they begin to inform decisions on accountability and the ability of teachers to meet certain standards (Conley et.al 2010, p. 105). This places a tremendous amount of pressure on teachers because, as pointed out earlier, there are many reasons why learners may not perform well in assessments - many of which are beyond a teacher's control. Teachers also tend to dislike the summative aspect of assessment because it interferes with the way in which they perceive their real roles as nurturers of knowledge (Brookhart 2004, p. 443).

In extreme cases, external programs can be imposed on teachers to such an extent that they become de-skilled and de-professionalised as their own thoughtful classroom practice is set aside and they begin to teach to the test, thus narrowing the curriculum (Shepard 2000, p. 9). The pressures of accountability can result in a focus on performance and competition amongst the learners with the emphasis being placed on end results, rather than the process of learning (Hugo 2013, p. 111). Adolescents are focused on many developmental tasks in this phase of their lives and good academic performance is a significant factor in peer, self and social acceptance, making the social pressures for adolescents to perform well immense (Gumbiner 2003, p. 45).

Summative results can certainly improve under this pressure, but this improvement does not necessarily reflect learning gains (Shepard 2000, p. 9). External pressure to perform causes

teachers to emphasise grades and real learning gains are often ignored, which causes learners to focus on just getting through tasks. Learners then resist engaging in cognitive behaviours that are unusual and thus feel risky, yet could enhance their learning (Black and William 2006, p. 18).

2.3. ASSESSMENT AND THE TEACHER

The teacher plays a vital role in assessment because it is his or her judgement on the abilities of the learner that translates into marks and qualitative feedback about ability. In the most extreme cases teachers actually conceptualise assessment as something completely separate from instruction (Shepard 2000, p. 5).

According to Urhahne and Zhu in Rubie-Davies et.al (2015) a teacher's ability to judge the performance of students is not only dependant on his or her ability to determine the quality of the work but also on how well they know their learners and their individual characteristics (p. 305). In the Chinese study, Cheng and Sun (2008) found that only a small number of teachers assigned a grade in response to the student's actual academic achievement (p. 336). Teachers used an assigned grade for a variety of different purposes from encouragement, to reminders (p. 337). Brookhart (as cited in Cheng and Sun) comments that, "achievement is a part of the construct but not the whole of it (p. 339)."

The central theme that was mentioned by teachers in this study is that they were extremely concerned with fairness when awarding marks. This, coupled with a desire to ensure benefit for their students, highlights the dual roles that teachers play, those of coach and judge (p. 339). Thus grading is never a completely value-free process.

The way in which teachers assess learners can have a profound effect on a learner's motivation. Urhahne and Zhu in Rubie-Davies et.al (2015) found that teachers tend to be too optimistic when it comes to judging the skills of their learners. It is easy to imagine that this optimism can have negative effects for both the teacher and the learners when compared with a more accurate measure of what the level of the learner's skill really is. For example, when marking is moderated by a head of department or marked by an external marker.

However, learners do actually begin to develop in line with teacher expectations. Learners whose abilities are over-estimated by teachers and are exposed to higher teacher expectations tend to perform better, learners whose abilities are underestimated by teachers and whose

teachers have low expectations of them tend to under-perform. This is known as the Pygmalion effect (p. 311).

Before adopting a practice of simply, blithely communicating heightened expectations to learners it is important to note that Urhahne and Zhu caution that not all learners are susceptible to this self-fulfilling prophesy. If learners attribute poor performance to their own ability and success to mere chance, they tend to perform badly regardless of teacher expectations (p. 311). Therefore, it is vital that teachers communicate accurate and appropriate strategies for improvement that learners can engage with.

2.4.ASSESSMENT AND THE LEARNER

Macfarlane-Dick and Nicol (2006) make use of a conceptual model to describe learner self-regulation and internal feedback during the assessment process. This model is a seven-step process that begins with the assessment task itself. Engaging with the task requires the learner to draw on prior knowledge and beliefs about their ability to address the requirements of this task, the learner then may formulate their own goals for the task.

The goals formed by the learner are an important point because they can either, overlap with those of the teacher or can be significantly altered by teacher expectations, as demonstrated by studies investigating the Pygmalion effect mentioned earlier. After these processes, the learner begins to generate internal and external outcomes. Internal outcomes refer to the internal cognitive and motivational states that the learner experiences during the task and external outcomes refer to the actual product of the assessment, for example an essay (p. 202).

The process of feedback for learners begins almost immediately upon engagement with the task. This is what Macfarlane-Dick and Nicol refer to as internal feedback. Learners already have some idea of how competent they are at meeting what they understand the requirements of the assessment to be as soon as they begin the task. They regulate themselves by adjusting their strategies, referring to prior knowledge, setting goals and motivating themselves accordingly (p. 202).

Once the task is completed, the learner can then receive external feedback from either teachers or peers. This kind of feedback can either: support the learner's initial understanding of the task and strengthen their belief in their competence or it can be in conflict with how they approached the task. A conflicting view can either encourage learning or decrease

motivation to such a degree that a learner's beliefs about their performance are negatively affected. When a learner engages with external feedback it ultimately influences their internal feedback and self-regulation during subsequent assessment tasks (p. 202). This highlights how important it is for a teacher to be able to provide effective feedback.

2.5.HOW TEACHERS CAN COMBINE HIGH-STAKES TESTING AND REAL LEARNING

Palmer and Stevenson (1994) maintain that effective teachers constantly assess their learners (p. 154). This assessment is not necessarily always formal and it is driven by the need that a teacher has to access information that will help learner learn (p. 154) by monitoring how they interact with assessment tasks (p. 151).

In order for teachers to use high stakes testing and standards-based accountability effectively, as opposed to destructively, Graue and Johnson (2011) propose that two conditions need to be met. Firstly, there needs to be an alignment between the curriculum and the assessments (p. 1830). Conley et.al take this a step further and call it constructive alignment (p. 109). Constructive alignment is when the teacher takes care to make the learning outcomes and levels of skill and understanding explicit during the teaching process. Tasks and activities in class are then designed to give learners the opportunity to attain the intended learning outcomes and then formal assessments are designed to measure these same outcomes (p. 110).

Secondly, teachers need to focus on capacity building. Capacity building is important to learners and teachers alike; teachers need to have the knowledge, skills and resources to implement an aligned curriculum and it is through the aligned curriculum that learner will be able to build their capacity for knowledge. (p. 1831).

Black (2015) asserts that the main aim of schooling is "learning how to learn" (p. 171), not performance for assessment purposes. Teachers perceive the nurturing of growth and understanding in their students as central to their job and often become frustrated with the summative nature of assessment when they perceive it to be counter-acting capacity building in their learners (Brookhart 2004, 443). Therefore aligning a rigorous high-stakes testing program with the ideal of capacity building for teachers and learners will produce the best results.

2.5.1. HOW SUMMATIVE ASSESSMENT CAN BE USED FORMATIVELY

Stiggins (in Brookhart 2004) notes that assessment practices often emphasise a summative judgement and that current assessment practices reinforce the idea that failure is bad (p. 439). The idea that failure is bad is compounded by the socio-cultural perspective, mentioned earlier, that the results of summative assessments are linked in some way to the measure of a learner's intelligence (Elwood and Murphy 2015, p. 189). Furthermore it has been, and still is, assumed that assessment itself would motivate learners by intimidation (Brookhart 2004, p. 439). Therefore, it is vital that assessment be used correctly as a capacity building tool rather than an external motivator linked to a learner's idea of their identity and position in society.

In order for teachers to play an active and mindful role in building the learning capacity of their learners they need the skills to identify the areas in which that learning needs to take place. Palmer and Stevenson (1994) make the link between assessment and diagnosis (p. 153). By the term diagnosis, they mean constructing an idea of what the learner's existing level of knowledge is, determining any difficulties that the learner may be facing and also identifying future learning needs (p.151).

Hugo (2013) states that summative assessment is most useful as a "feed forward mechanism" as it provides a basis from which teachers and learners can see what kind of work is required for future improvement (p. 111). Taras (in Jiang 2015) concurs with this by stating that formative assessment is actually only possible after some kind of summative judgement has been made on learners' capabilities (p. 180) therefore the results of a summative assessment provide a suitable starting point to address the gaps in learners' understandings and begin to build their capacity.

Summative assessment can also be used formatively, particularly when thorough feedback is provided after tests and interventions are specifically structured to help learners act on that feedback (Jiang 2015, p. 182). It is in this way that the curriculum and the assessment tasks can be aligned in the classroom and build the capacity of the learner at the same time. Black (2015) also points out the metacognitive value of summative assessment as a review of a "learning episode" (p. 170). Weaver (2006) cautions that feedback on summative assessments that is designed with the aim of developing learning must be part of a consistent assessment program where assessment tasks are relatively similar (p. 382) otherwise learners do not see

the relevance of it and are not motivated to use it in future assessments. This is exactly what Conley et.al suggest in their description of constructive alignment (2010, p. 109).

Swaffield (2011) encapsulates the process and value of summative assessment best in her exploration of the etymology of assessment. Assessment has a Latin root in the word "*assidere*" meaning "to sit beside (p. 434). Through this definition she identifies assessment as a support for learning rather than a test of ability. This underpins the attainment of what Black (2015) proposes as the main aim of schooling: "Learning how to learn" (p. 171). When teachers 'sit beside' a learner and take a close interest in what they think and do, they can have a better understanding of how to shape further learning and capacity building through appropriate feedback (Swaffield 2011, p. 435).

2.6.THE IMPORTANCE OF FEEDBACK

It is also important to remember that assessment is not only the administration of tests and collation of results, but also the ability to interpret the results and make them meaningful in terms of the learner's capabilities through giving feedback (Jiang 2015, p. 177). It is through the correct use of good formative assessment and feedback that the standards of learning can be raised (Black 2003, p. 21).

Feedback is ultimately linked to the interpretation of results. However, it is important that the feedback is thoughtfully structured in order for it to be of maximum benefit to the learning process. Feedback should demonstrate to learners the link between what they did in the assessment and the results that they got (Brookhart 2008, p. 21). Feedback can be general or private depending on the context but it must encourage the view that mistakes are part of learning (Black and William 2006, p. 24).

Black and William (2006) maintain that the way in which feedback relates to the task is crucial, (p. 51). This kind of task-oriented feedback has to be carefully constructed; it cannot be overly prescriptive or too comprehensive, it is most effective when there is room for the students to apply their own knowledge to mitigate it (p. 52).

Feedback should focus on the nature and importance of the effort required by the learner; otherwise, it can have a negative impact on learner's beliefs about their ability to learn (Black 2003, p. 18). Furthermore, it should be explicit about the gap between the learner's present performance and the desired standard and provide guidelines about the action that can be taken by the learner to close that gap (Black and William 1998, p. 20).

Sadler (in Macfarlane-Dick and Nicol 2006) describes this in a three-point process. Firstly learners must know what good performance is, secondly how their performance relates to good performance and lastly, how they can close the gap between the two (p. 204).

Feedback can have negative effects when it is directed at the self in the form of rewards or punishment (Black and William 2006, p. 49). Even praise can be harmful in certain instances because it is an external regulator of motivation and interest as opposed to the learner being interested in the task itself (p. 50). Brookhart (2008) takes this a step further and links self-referenced feedback to what she calls "academic fatalism (p. 21). Academic fatalism arises from a situation where students believe that their performance is a direct result of a fixed level of intelligence that they have. Therefore feedback on assessments that is directed at the self, compounds this belief, especially in academically weak learners.

Black and William (2006) attribute the feeling of academic fatalism amongst weaker students to the fact that grade associated performance is highlighted by the school system and students working towards performance goals tend to resort to maladaptive practices (p. 22). Therefore, a low mark accompanied by feedback that provides the reason why the mark is low, can have a compounded negative effect.

Most teachers take it for granted that learners will apply feedback and correct their own errors (Shepard 2000, p. 11) however, this only happens when learners understand the value of developing their capacity. Students sometimes do not recognise the helpful nature of feedback. Weaker students especially can feel that it carries an enhanced penalty for failure because it acts as further evidence of their low ability (Black and William 2006, p. 22). Furthermore, teachers tend to take it for granted that their feedback is clear and easily translated into action on the part of the learner. This is not always the case; learners often require further clarification and interpretation in order to begin to apply feedback (Macfarlane-Dick and Nicol 2006, p. 200-1).

Feedback needs to explicitly support autonomous learning (Hargreaves 2013, p. 230). It is a prime determiner of the processes inherent in self-regulated learning (Butler and Winne 1995, p. 245). Once learners adopt the processes of self-regulation they actively seek feedback from external sources to help them to monitor their own states of engagement (p. 246). This makes perfect sense when applied to Brookhart's (2008) claim that "feedback is most powerful when it corrects misconceptions (p. 20)".

Butler (1988 in Black 2003, p. 11) established that learning can be advanced through formative comments and in this way students can learn through improving their work (Swaffield 2011, p. 443). Brookhart (2008) describes this as the "double-barrelled effect of formative feedback" because formative comments have a positive effect on both performance in assessments and the building of the learner's personal capacity (p. 8).

In order for feedback to correct learners' misconceptions and develop their capacity, feedback needs to be both "backward-reaching" and "forward reaching". It needs to provide information about the level of the work that the learner has presented and it needs to address the factors necessary to improve future performances (Butler and Winne 1995, p. 247). This process-driven feedback gives learners information on how they approached the task and helps them understand how their performance matched up to the learning goals of the assessment (p. 246).

2.6.1. CRAFTING SALIENT FEEDBACK

a. Timing

The most important criterion that feedback needs to meet is salience. In order to achieve this, it is very important for it to be either immediate or only slightly delayed after a task has been completed (Brookhart 2008, p. 10). It is important that learners are still mindful of the learning goals and feel that they are still worth striving for (p. 11). This makes sense as task related motivation can only work while mastering the task is relevant to the learner.

Some learners may also require clarification of what the feedback actually means. In a study by Weaver (2006), she discovered that 2nd and 3rd year university students felt they needed guidance in understanding and using the feedback they had received in written comments on assessment tasks (p. 384).

If this is the case for university students, high school learners must have similar, if not more acute, needs for clarification on written feedback. Therefore it is necessary for the teacher to be present to explain the feedback in person otherwise it can be misinterpreted, causing further confusion (Weaver 2006, p. 389). Of the university students that Weaver surveyed only 26% reported having had help with understanding and applying feedback in high school (p. 385).

b. Quantity

The amount of feedback provided also needs to be considered. Brookhart (2008) uses what she calls "The Goldilocks principle (p. 13)" to explain this, "not too much, not too little, just right". Teachers need to be aware of both the topic and the typical learning progressions involved in mastery, and balance the amount of feedback that they give against the capacity of the individual learner to attend to and apply comments. Thus, they need to tailor the communication of requirements using their knowledge of their individual students and how these students have interacted with that particular task (p. 12).

An approximation of a good amount of feedback would be two or three main points that give information on what the student may benefit from working on (p. 13). However, this is where the knowledge of the individual learner's capacity comes in, some learners can work on two or three comments simultaneously, others prefer to master one step at a time (p. 12). Hargreaves (2013) concurs with this point as her study on how children experience feedback revealed that learners feel they benefited from feedback that was substantial but not burdensome (p. 236).

Conley et.al (2010) propose that feedback "should answer three major questions asked by a learner; where am I going? How well am I doing? Where to next? (p. 115). By addressing these three points, teachers can give learners enough information for self-regulation and further learning opportunities.

c. Type

Teachers tend to tailor feedback to the needs of individual students as recommended by Brookhart (2008, p. 17). Lower achieving learners can tend to resist feedback as they perceive it as repetitive and overly directive (Hargreaves 2013, p. 233). However, as lower achieving learners do tend to make more mechanical mistakes, teachers can easily fall into simply correcting these kinds of errors, which does not provide as much assistance to the student as feedback about the process of constructing a text (Brookhart 2008, p. 15). In fact, feedback that is content driven rather than process driven is actually unlikely to help learners in future assessments (Weaver 2006, p.381).

It is feedback on learning processes that has been proven to have the most impact on learning in the long term (Hargreaves 2013, p. 239). However, it is very important that the feedback that is given is directly linked to the assessment criteria otherwise it is not helpful at all

(Weaver 2006, p. 389). Furthermore, very vague and general feedback is of no use because it does not provide specific guidelines for improvement and students can even feel short-changed (p. 390).

Merely noticing what is wrong without giving suggestions as to how to make it right is not beneficial to learning (Brookhart 2008, p. 26). Providing feedback with too narrow a scope, or that is perceived as judgemental can cause learners to become hurt or angry leading to them being unable to use the feedback to enable them to learn how to improve their work (Weaver 2006, p. 381). Learners with low self-esteem are particularly vulnerable to any feedback that they perceive as judgemental and tend to feel defeated by it rather than motivated (p. 382). Weaver's (2006) study found that even learners with medium to high self-esteem can find overly critical and negative feedback damaging to their motivation and confidence (p. 390).

2.6.2. FEEDBACK AND LEARNING

If feedback is to promote and support learning it is important that it be constructed in such a way that it engages with how learning happens in the first place. Palmer and Stevenson (1994) identify implicit and explicit learning as the two ways in which people learn in any situation. Implicit learning is largely unconscious whereas explicit learning requires conscious and deliberate effort (p. 1).

Schooling focuses on explicit learning, but it is important to note that the effects of implicit learning are always present. The three ways in which explicit learning is described by Palmer and Stevenson are memorisation, problem solving and understanding (p.4). Effective feedback would need to be constructed in such a way that it supports these different facets of explicit learning.

Explicit learning is demonstrated in a learner's performance in assessment tasks. We can link this to the external outcome of an assessment described by Macfarlane-Dick and Nicol (2006, p. 202). It is in the external outcome of an assessment that the impact of both explicit and implicit learning is revealed.

Explicit knowledge becomes implicit when it is practised enough that it becomes automatic (Palmer and Stevenson 1994, p. 3). The learners may not even be consciously aware that they are demonstrating or applying it to their work; "it is easy to use but hard to describe, because we are rarely consciously aware of it (p. 3)".

2.6.3. FEEDBACK AND LEARNER MOTIVATION

Learners use any kind of feedback that they get at school from a very young age to develop ability-related beliefs and values (Rubie-Davies, Stephens and Watson, 2015, p. 13). These are distinct for different subject areas and seem to decline as the learner gets older and they either become more realistic or more pessimistic about their abilities (p.14). Learners also tend to compare themselves to their peers as they get older and learners who perform poorly become aware that they lag behind their peers. This awareness often leads to diminished ability-related beliefs and decreased motivation (p. 14) much like the "academic fatalism" described by Brookhart (2008, p. 21). According to Rubie-Davies, et.al (2015) interventions aimed at improving learner motivation are an important new direction in the drive to improve learner achievement (p.15).

Feedback as an important catalyst for self-regulated learning as it helps learners to monitor the level of their progress on assessment tasks (Butler and Winne 1995, p. 246). Self-regulated learning involves the learner being able to control their thinking, motivation and behaviour during learning (Macfarlane-Dick and Nicol 2006, p. 199). Self-regulated learning also extends into the types of goals that learners set themselves, how much effort they are willing to exert and their reactions to external feedback (p. 199). Information about progress or competence is an important factor in keeping learners motivated to engage with the demands of assessment tasks (Butler 1988, p. 2). This information allows the learner to generate internal feedback on their progress towards achieving their own goals (Macfarlane-Dick and Nicol 2006, p. 200).

Butler (1988) conducted a study in which she tested the hypothesis that intrinsic motivation, and probably self-regulation, will be differentially affected by task-involving and ego-involving evaluation (p. 1). Task-involving and ego-involving motivation are both characterised as forms of intrinsic motivation (p. 2).

a. Ego-involved motivation

Ego-involved motivation addresses the concerns of self-worth in an evaluative way. This makes the learner concerned with demonstrating high ability or masking low ability in an effort to preserve self-worth. Butler (1988) identifies grades or marks as having an ego enhancing or diminishing effect on students (p. 2). This same concept is described in Rubie-Davies et.al (2015) as performance approach goals, in which learners focus on maximising

favourable evaluations, and performance avoidance goals in which negative evaluations are minimised (p. 11).

Due to the high competition that accompanies a high-stakes testing environment, students tend to engage in their own form of norm referencing. Norm referencing is when the performance of one student is compared to that of other students (Brookhart 2008, p. 22). This can be particularly dangerous to the motivation of learners who do not perform well in assessments (p. 23). Rubie-Davies et.al (2015) support this by pointing out that as learners get older, they are more likely to compare themselves to their peers. This causes not only a drop in motivation, but even poorer performance too (p. 12).

A learner who has achieved a disappointing mark in an assessment task can have such a negative emotional reaction to that mark that they are unable to process any other feedback information in the form of remedial comments (Emmer and Stough 1998, p. 350). Because of the learners' emotional investment in the mark that they achieve, feedback in the form of comments is often ignored. Black and William (2006) comment that general reviews of the field of assessment stress that feedback that comments directly on the learners themselves impacts on their self-esteem and can have a negative effect on performance (p. 23).

b. Task-involved motivation

Task-involving motivation is directed at improving the individual's mastery of task demands and improving previous performances (Butler 1988, p. 2). Bandura (in Butler 1988, p.2) asserts that people actively seek information about their own competence and that this is vital in maintaining self-regulated motivation. Brookhart (2008) aligns this kind of feedback with criterion referencing in which student work is compared with the specific requirements of the learning target. This kind of feedback helps the student to regulate their own performance based on how close they are to the target requirements (p. 23), to "get smarter though effort (Swaffield 2011, p. 438)". This allows learners to develop what Rubie-Davies et.al (2015) refer to as "task-approach goals (p. 12)" based on competence in relation to the characteristics of the task. These goals are positively related to students' intrinsic motivation.

Butler took this a step further in terms of its implications in assessment settings and predicted that when learners are motivated to improve their mastery of a task, their performance of that task will be enhanced (p. 2). The assessment itself is not enough to motivate learners; they need to be able to experience real success in assessment tasks (Brookhart 2004, p. 439).

When feedback is explicit about how mastery can be achieved, learners will be more motivated to perform well in an assessment task. Constructive feedback improves learner performance far more than the simple knowledge of results (p. 440). Rubie-Davies et.al (2015) actually differentiate between mastery and performance goals saying that mastery goals link to the intrinsic interests and motivations of learners and encourages them to develop and use deep cognitive strategies when processing information. This process does not always have a corresponding positive effect on performance however, because in some cases learners can become so focused on developing their own interests within the material that they actually ignore the assessment requirements (p.12).

Brookhart (2004) states that for feedback to be effective, it needs to be given under conditions that encourage learners' mindful reception of the content (p. 440). Feedback about the process of meeting assessment goals is particularly valuable as it makes the learner aware of the connections between what they did and the extent to which that helped them achieve the goals of the assessment (Brookhart 2008, p. 21). It is possible that Butler's study has hit on the perfect conditions to promote the mindful reception of comments; namely, the absence of marks.

2.6.4. THE DIFFERENT EFFECTS OF EGO-INVOLVING AND TASK-INVOLVING FEEDBACK ON LEARNER MOTIVATION AND LEARNING

Weaver (2006) reported that many teachers believe that, although much time and effort is put into structuring task-involving feedback, learners pay scant attention to it. She cites two studies that may have contributed to this belief; one by Wojtas in 1998 which showed that learners are only interested in their marks and actually ignore feedback and one by Fritz in 2000 which discovered that task-involving feedback was not effective in improving learning at all (p. 380).

However, there are indications that this is not the full picture of learner perceptions of task-involving feedback. Firstly learners may need advice on how to understand and use task-involving feedback (p. 379) secondly they need to perceive the feedback as intended to support their learning, rather than acting as a further judgement on their performance (p. 381).

Butler's statistical study demonstrated that self-regulated, task-involved motivation in learners was highest when marks are withheld from learners and only comments are provided. This increased learners' interest in the task and the application of the comments

and thereby improved their performance (p. 13). The combination of ego-involving feedback (marks) and task-involving feedback (comments) merely induced an ego-involving orientation in the learners, through which one can deduce that emotional salience of the marks distracted learners from interest in the task itself.

2.7.WRITING AS CENTRAL TO LEARNING AND ASSESSMENT PERFORMANCE

Writing tasks in English Home Language account for a large percentage of a learner's term mark. Essays are marked out of a possible 50 marks, which is the highest total for any of the assessment tasks, making a learner's ability to construct their own texts the most significant indicator of their performance. Palmer and Stevenson point out that, the English curriculum is the only place in school where essay writing is explicitly taught as a skill in its own right (1994, p. 122). Furthermore, it is through writing that learning through understanding takes place (p. 122). All of these factors make writing central to learning and assessment performance in school.

Essay writing is a process that links very well to Palmer and Stevenson's explanation of how explicit learning may be demonstrated through memorisation, problem-solving and understanding (1994, p. 4). Essays are assessed using the criteria of content, language and structure (Appendix A). On the most basic level a learner can use memorisation to remember the guidelines of essay writing; the focus of the essay, the structure and even new and exciting words that they may want to try out. Writing also encourages problem-solving (Palmer and Stevenson 1994, p. 125), a learner has to be able to convey an idea or description using the structure and word limits of the task: "it is a mechanical problem solving strategy in which the goal is to put one's thoughts onto paper (p. 116)".

Finally writing not only encourages learning through understanding (p. 125), it also requires the understanding of the "substantive goals" or information that is to be conveyed and the "rhetorical goals" or way in which the information is communicated. Rhetorical goals involve the understanding of the subtle differences between the meanings of words, sentence structure and planning for coherence (p. 113) and thus require the learner to draw on memorisation, problem solving and understanding in one task.

The assessment of and provision of feedback on the writing process may be one of the most effective ways in which a teacher can align the goals of explicit learning with the assessment process.

2.8.SUMMARY

Assessment is an integral part of schooling. However, it has come to represent more than just the extent to which learners achieve learning goals; it is a means of certification, social control and, very damagingly, a socially accepted measure of a learner's intelligence.

High stakes testing in an environment where high levels of accountability are present can improve learners' assessment results. However, there are many maladaptive by-products of learner achievement in this situation. Teachers teach to the test and become de-skilled and demotivated, the curriculum becomes narrowed and learning interventions tend to target only weak learners. Therefore marks may improve, but real learning, i.e. the building of each learner's capacity for knowledge, is lost.

Assessment does have many formative benefits that are overshadowed by increased accountability and high stakes testing. The most important of these is that both teachers and learners alike learn from assessment if it is conducted mindfully with good quality feedback. The combination of aligning the curriculum with the assessments and the focus on capacity building as central to schooling will help learners to perform and to learn in an environment where high stakes testing and high levels of accountability are present.

Feedback is the ideal mechanism for aligning the curriculum to assessment and building the capacity of learners and teachers alike. It is through feedback that summative assessment can become a formative tool in the classroom. Feedback is also a very powerful motivator because it is also a feed-forward mechanism, meaning that not only does it comment on past performances, it can also correct misconceptions and influence future assessments.

Feedback is a powerful motivator; yet, different types of feedback have different effects. Marks are also a form of feedback, which Butler (1988) identified as ego-involving. Marks have a negative impact on low achieving learners because they confirm their beliefs about their low ability. Task-involving feedback refers to the comments on how to achieve mastery of the task. These comments are most powerful in the absence of ego-involved motivation because they provide information about the task that increases interest and self- regulated motivation.

Providing the most effective method of feedback on the writing process may be one of the best ways in which a teacher can have a positive effect on the learning process.

2.9. CONCLUSION

Maybe do a short summary here of the ground you have covered in the Literature Review: the main insights you have gained and which ideas you want to take forward into your research analysis.

High stakes assessment is an integral part of the school system. Teachers and learners need to know how to work with assessment in order to achieve the results that satisfy accountability pressures and nurture capacity building and real learning gains at the same time.

It may be that by withholding marks, the ego-involving element of motivation falls away and that will leave the space for the concern with mastery to come to the fore.

Once weak learners are unencumbered by the negative effects of ego-involving feedback, they can focus on the steps that they can take to improve their understanding of the assessment task and the steps that they need to take in order to improve their performance.

CHAPTER 3: RESEARCH METHODOLOGY

3.1. Introduction

This chapter contains a description of the methodology, tools and process used in this study. This study is based on an experimental study already conducted by Butler in 1988 *Enhancing and Undermining Intrinsic Motivation: The Effect of Task-Involving and Ego-Involving Evaluation on Interest and Performance*. The aim of my study is to investigate the effects of withholding marks and providing written feedback to academically weak learners in a real life school setting using formal assessments.

3.2. From an experimental approach to action research

Butler's 1988 study was designed in order to confirm four experimental hypotheses generated around the various effects of giving task-involving and ego-involving feedback (p. 3). The experimental approach to research design concerns itself directly with causality, "if *x*, then *y* (Cohen et. al. 2011, p. 312)". The study carried out by Butler was an experiment in its most classic form, conducted in a contrived, artificial environment (Cohen et.al 2011, p. 315). This controlled environment was a good choice for an initial investigation into the effects of task-involving and ego-involving feedback and is certainly a way of ensuring the internal validity of an experimental study, as Leedy and Ormond, 2013, p, 102 point out. The results of the study show a clear causality between the provision of task-involving feedback, in the form of comments, and learner motivation and performance (Butler 1988, p. 13).

However, as Black and William (2006) point out, Butler's study lacks ecological validity because it was conducted outside normal classroom conditions (p. 14). Therefore, it is important to establish the ecological or external validity of Butler's hypothesis in a real-life setting so that it can be established whether or not her findings would apply to situations beyond the controlled conditions that she conducted her study within (Leedy et.al 2013, p. 103).

3.2.1. Providing the ecological validity lacking in Butler's study

In order to give Butler's study the ecological validity that it lacks, it needs to be performed in the real world; in a functional classroom setting using regular assessments. Cohen et.al (2011) refer to this as a quasi-experiment (p. 315).

Leedy et.al. (2013, p. 103) propose conducting a study in a real-life setting or the replication of a study in a different context as two ways of enhancing the ecological or external validity of a study. In that case, the study would still have the experimental element of testing the hypothesis that if marks are removed, learners will focus more on feedback and this will improve their performance. Yet what would also be established is whether or not the hypothesis remains true in the context of a real-life classroom.

In this study, as the researcher I will be altering my own grading and feedback practice in order to test this hypothesis, instead of changes to normal classroom grading practices being administered by an outsider. With the teacher occupying the dual positions of participant and researcher, this study then becomes action research (Leedy et.al 2013, p. 100).

3.3. The research method: Action research

In an educational setting, action research is a popular tool amongst researchers and teachers alike because it lends itself to situations that involve people, tasks, procedures and the need for solutions (Cohen et. al.2011, p. 344). Koshy (2008) defines action research as a rigorous enquiry undertaken to constantly refine practice (p. 1-2). It is based on enquiries in practical contexts and has the potential to create new knowledge and propel professional development (p. 3). McIntosh (2010) highlights the practical functions of reflexive thinking and action research in enhancing both personal and organisational functioning (p. 31).

Action research is also defined as "a small-scale intervention in the functioning of the real world and a close examination of the effects of such an intervention (Cohen et.al 2011, p. 345)". For the purposes of this particular study, the latter definition is perhaps the most accurate. Furthermore, Cohen et.al point out that action research seeks to "bridge the gap between research and practice" (p. 345). This study aims to do exactly that by establishing whether a hypothesis that has been proven under experimental conditions can hold true in a real-life classroom setting.

The result of an action research project can contribute to theories of education because it is accessible to other teachers (Cohen et.al 2011, p. 345). This is what makes the replication of Butler's study in a real-life classroom important: Should Butler's (1988) findings, that learner interest and performance on assessment tasks was highest after comments only and that receiving marks only or marks plus comments had an undermining effect on learners' interest and performance (p. 9) be confirmed in a real-life setting, teachers in almost any context

could withhold marks as a legitimate method of enhancing learner interest and assessment performance.

3.3.1. Action research in adapting Butler's 1988 study to a school context

By conducting an adaptation of Butler's study in a regular classroom situation, using the CAPS prescribed assessments (CAPS Document p. 124), the hypothesis that withholding marks and providing comments only will improve learner interest, motivation and performance can be given ecological validity through action research.

The study constitutes action research because it involves an enquiry into the effectiveness of a change of practice on the part of the teacher. The desired outcome of this enquiry is continuing professional development and the observation of evidence-based outcomes that will hopefully contribute to the practice of assessment as a whole (Koshy 2008, p. 2). The study is also aligned with action research because it aims to use an experimental hypothesis and it is applied in a practical context (p. 3).

Action research involves self-reflection and reflexive practice. The research is not typically "performed on other people" (Cohen et. al 2011, p. 345) as Butler's study was. A replication of Butler's study in the context of a normal classroom may appear to be a quasi-experiment carried out on the learners who participate in the assessments, as it tests a hypothesis in a real world setting. This would be the case without an awareness of the nature of effective assessment. Yet it is important to remember Swaffield's (2011) etymology and definition of assessment as "to sit beside" (p. 343). Assessment should not be something that is done to learners; the process of 'sitting beside' should also affect the teacher's current understanding and future practice.

Therefore, this study does not simply test a hypothesis in a classroom setting; the change required in the teacher's approach to assessment and feedback will be an enquiry in which new knowledge can be made available, classroom practice will be refined, and the changes will be understood and evaluated. It is in this way that Butler's experiment translates into action research in an ecologically valid context.

3.3.2. Advantages of action research

Action Research has the advantage of involving the researcher as a participant in the study as it can provide information about a specific context (Koshy 2008, p. 21). Because action research lends itself to a quasi-experimental paradigm by observing and documenting the outcomes of an intervention in the setting of a real classroom, it can provide opportunities for a tried and tested theory to emerge (p. 21). It also has a practical outcome because it aims to improve a particular educational concern and can be significant for future practice and research (Govender 2013, p. 178) - it improves education by changing it (Cohen et.al 2011, p. 347).

The biggest benefit to be gleaned from this is that it makes the findings immediately relevant and applicable to other practitioners in similar situations (Cohen et.al 2011, p. 29). Whitehead (2009) is a particular proponent of action research for this same reason, because it makes experience public to others who experience similar situations (p. 86).

3.3.3. Limitations of action research

"The undertaking of 'action' is far more complex than first considered (McIntosh 2010, p. 32)". There are two main limitations of action research: ethical concerns and validity of findings. The ethical concerns are mainly to do with the protection of participants' identity and the transparency with which the research is conducted (Koshy 2008, p. 84). In this study the participants agreed to be known by their initials. I asked the learners if they would accept me using their initials to identify them in the study because in that way they retained a bit of their identity while also remaining anonymous. As the teacher, I explained the exact process and purpose of the research in person as well as in a letter (Appendix B and C).

A practical concern for this study is the validation of the information for accuracy and relevance. Firstly, the study may have too limited a scope and participant base to be relevant, and secondly, as a participant in the research, the teacher may not be completely unbiased in her assessment of learners' progress. To counteract the first limitation, I focused on the action research element of the study as a small-scale intervention that would inform my own personal practice.

To counteract the second limitation, I asked colleagues to moderate or verify the changing nature of the work that the students produce, as suggested by Koshy (2008, p. 105).

3.4. The research tools

The tools for a study that is both quasi-experimental and action research must be carefully selected. As the main aim of the study is to investigate whether Butler's study can be replicated in a normal school environment, it is important to use tools that replicate conditions that are normally experienced by learners as much as possible.

3.4.1. School based assessments

I chose to use school-based assessments to collect data in the study in order to increase the ecological validity of Butler's hypothesis. As the researcher, Cohen (2011) provides me with the justification for the use school-based assessments as tools for structured observations. According to Cohen et.al 2011, a structural observation can allow the researcher to generate numerical data from observations and this numerical data is particularly useful in making comparisons between different situations, such as a control and experimental group (p. 459). Formal school assessments are designed to generate marks and this is the numerical data that I can use for this study.

The CAPS assessment requirements for English Home Language (CAPS Document p. 24) specify that a creative essay of 250-300 words is to be written every term. An essay task was marked according to a rubric (Appendix A) and this rubric functioned as a kind of rating scale (p. 462) that helped me as the teacher/ researcher to record the degree to which the learners' writing fulfils the criteria that I wanted to observe. Cohen et.al (2011, p. 456) state that because of its contextual sensitivity, observational data can demonstrate high ecological validity making it a suitable tool for a study that aims to demonstrate ecological validity.

The essay tasks and the marks that the learners achieve in them can be subjected to norm-referenced and criterion-referenced interpretation: the method of awarding marks when assessing essay writing is a criterion-referenced one. Learners write the essay with the aim of fulfilling a specific set of criteria (content, language and structure as indicated on the rubric, Appendix A). It is in this way that their efforts can be recorded as a criterion-referenced quantity. In the norm-referenced interpretation, the marks of the learners from the control group will be compared to those of the experimental group so that the differences in their achievement, under the conditions of the study, become apparent (Cohen 2011, p. 478).

The formal essays that I have used were part of the CAPS curriculum requirements for Grade 9 (CAPS Document p. 124). Each essay was written in controlled test conditions and a

different choice of topics is given each term. I marked the essays using the Gauteng Department of Education Rubric (Appendix A).

3.4.2. A Pilot Study

Cohen et.al 2011 recommend the use of a "reconnaissance of the circumstances of the field (p. 352)" before a general plan of action is made. This fits in with the action research model, proposed by Kemmis and McTaggart (in Koshy 2008, p. 5) that involves planning a change, acting and observing, reflecting on the process, reflecting and re-planning and then repeating the whole process as many times as necessary (p. 4). This initial cycle of planning, acting and observing serves as a way for the researcher to gather information for the general plan of action.

3.4.3. The Questionnaire

Questionnaires are useful in collecting both quantitative and qualitative data and can be designed around the research question and the empirical sub questions allowing direct data to be gathered (Cohen et. al 2011. p. 380). The questionnaire in this study (Appendix D) included both closed (subject to quantitative analysis) and open questions (for qualitative analysis). When constructing the questionnaire as the researcher, I focused on the empirical sub-questions:

1. Does withholding marks induce learners to take more interest in and better remember a remedial comment?
2. Will learners be encouraged to apply the guidelines of that comment to future assessment tasks and thereby improve their marks?

The questionnaire was constructed with a closed question and an open question for each of the empirical sub-questions, four questions in total.

Cohen et.al (2011, p. 383) identify that there is a significant drawback when administering open questions to children as they may not hold an opinion relevant to the inquiry but may feel obliged to respond anyway. Furthermore, there is the possibility that the respondent may not have the vocabulary with which to respond to the items. The questionnaire was designed with this in mind and I clarified each of the questions before the learners wrote down their responses.

3.4.4. Focus groups

Cohen et.al (2011) describe focus groups as an additional tool that can help triangulate data (p. 436). For this study, a focus group was useful in order to gather additional information that would triangulate the results of the quantitative part of the study and the qualitative data gathered in the questionnaire. A focus group is useful in that it can yield a large amount of data in a short space of time and participants often spur each other on in the discussion (p. 436-7).

There is also the possibility that only the most confident participants will verbalise their opinion and this will have a negative effect on the overall reliability of the data (Cohen et. al 2011, p. 437) however, as Litchman (2006, p 162) points out, in the qualitative paradigm all forms of data can be considered legitimate.

3.5. The research process

3.5.1. The preparation for the study

The school where I conducted my research is an ex-model C high school in a middle-class suburb in Johannesburg. In this school, the learners are loosely streamed according to academic ability. There are nine classes in each grade and they are divided into three academically strong classes, three classes of average ability and three academically weak classes. The timetable is done by a computerised system called “SMART” and the allocation of specific learners to specific classes and specific classes to specific teachers is largely random but can be influenced by the programme administrator. Letters of the alphabet from A to J (the letter I is skipped) are assigned randomly to each of the nine classes and each class has between 20 and 28 learners.

As the researche, I requested two academically weak classes in grade 9 for the study. These classes became participants in the study as well as being the classes that I would teach English to for the year. I wanted the process of choosing them to be as unbiased and objective as possible so I did not look at class lists or speak to other teachers about the learners. I was offered the options of grade 9A, 9D or 9G and chose A and D because they both had 24 learners in the class. I also assigned the control and experimental groups before meeting the learners and decided that 9A would be the control group, as A comes first in the alphabet, and 9D the experimental group.

As part of the planning phase in the Action Research model, proposed by Kemmis and McTaggart (in Koshy 2008, p. 5), before meeting each class it was important to decide how to formulate the way in which I was going to present the research project as both teacher and researcher at the same time.

I expected the discussion with the control group to go smoothly because they were only going to be subjected to routine classroom practice; receiving marks and comments on all assessments, as is expected in all subjects from all teachers. My concern with the experimental group was that they would object to having their marks withheld because this is not standard practice at the school. The school has a policy that marks for assessments are to be available one week after they have been written, so learners are used to receiving very prompt feedback. Furthermore, there is a considerable focus on marks from learners and parents alike because the school has a reputation for producing good academic results. Therefore, I also anticipated objections from the parents.

In the information letter to both the control (Appendix B) and experimental groups (Appendix C) and the parents of both groups (Appendix E and F) I deliberately omitted any mention of the academic ability of the learners. Whilst this omission could be seen as compromising to the transparency of the study as an ethical concern (Koshy 2008, p. 84), the drawbacks of making it known to the learners that they are considered to be academically weak would have had a far more detrimental effect on the study than making a small ethical compromise. Firstly, it would be damaging to the self-esteem and self-image of the learners to be labelled as "academically weak" on paper and it would have a very negative impact on my relationship with them as their teacher. Secondly, parents would not appreciate the label and it could have led to broader problems with the school's way of organising the timetable. Instead, the letter focused on how comments are important in consolidating the skills and information learned in Grade 8 in preparation for the senior grades, making Grade 9 the ideal year to focus on them as much as possible.

In the letter to the parents of experimental group (Appendix F), I added that the research literature suggested that when a comment is provided without a mark, it appears to be more effective. The decision was to make this statement as oblique and general as possible, because on the one hand I wanted to demonstrate that the learners would not be disadvantaged by not seeing their marks immediately, but I also did not want to over-state the positive effects of this as I could not guarantee exactly what they would be.

When I met each class I explained the research project to them by saying that I was investigating the effect that comments have on helping learners improve their work. I did not anticipate any problems with the first part of the explanation but was very surprised by the way in which the learners reacted.

I did not realise that my perception of this was vastly different to that of the learners; in my own mind, this was a project that I felt was interesting and enjoyable. As a researcher and as a teacher I was curious about the study and its effects. I was approaching it with enthusiasm and almost saw it as fun. However, the ethical protocol demands informed consent and the minute that I mentioned the word "research" there was an immediately fearful reaction from all of the learners - even from the control group from whom I expected no resistance at all.

For the learners the word 'research' had negative connotations and I became afraid that this would create an anxiety that would prevent them from participating. I was not really prepared for this reaction so I took time to find out what their concerns were.

The learners from both groups were primarily worried about people other than the teachers at the school scrutinising their work and judging them harshly for its quality, "will other people mark our work?", "What if they laugh at our mistakes?". I did my best to reassure them that if their work was going to be seen by anyone other than me, then their identities would be completely protected.

3.5.2. The ethical considerations

It is possible that withholding marks could be detrimental to learners because they will not be able to evaluate their performance on a quantitative level, like they are used to doing. However, Black and William (2006) state that they have not come across any report of negative effects following an enhancement of formative practice (p. 17).

There is normally a drop in the learners' marks in the first term of Grade 9 (Appendix G) as they get used to the new requirements of a more senior grade, as discussed above. I took the following steps in order to minimise ethical concerns:

1. I conducted the study using only one type of task so that the rest of the assessment program could compensate for any negative effects on the learner's marks.

2. I made sure both classes knew they are essential to the study.

3. Both groups were at an advantage because of the elaborated comments they received.

4. The groups did not know that they were considered academically weak.

5. The groups were told that the aim of the study was to improve their learning, not necessarily their performance, as I intended to build their personal learning capacity.

6. I also hoped to eliminate any sense of competition between the two groups, as learners might have resorted to norm-referencing using their marks or suspected marks.

3.5.3. The pilot study

Before starting the official research I decided to conduct **a pilot study** with both of the classes. When the study was first conceived, I had envisioned using all of the CAPS assessments. This would have included comprehension tests, literature tests and writing pieces. As I spoke about the study more, it became clear to me that it would be difficult to withhold marks from learners in a comprehension test or literature test where there are too many separate items that require separate marks.

Providing comments without marks for comprehension and literature tests was causing me a lot of concern as a teacher: Firstly because it would be difficult to indicate how much of the answer was correct, according to the memorandum, without making it obvious how many marks had been awarded. Secondly, the English department marks comprehension and literature tests as a group. This means that each teacher marks a section of the test across all of the classes and does so under pressure to meet the official one-week deadline. The latter concern was more significant for me because the volume of marking is large and it would be inconsiderate to ask other teachers to alter their marking practices. Furthermore, involving other markers affects the consistency of how the comments are used. However, I was curious about whether comprehension tests could be used and I wanted to have a basic idea of how the learners of both groups would react to the study.

The attraction of using essay tasks is that providing comments on longer pieces of writing without revealing marks is relatively easy because essays are marked according to a rubric. There are no points at which it is necessary to make it clear whether the response is correct or incorrect.

For **the pilot study** I first taught a lesson on basic comprehension skills (Appendix H) to both classes. I focused on the steps that one would go through when reading and answering a comprehension, factors to be aware of (like looking at the mark allocation and answering in full sentences) and then skills for deciphering words that one may be unfamiliar with. The

learners were then given a short comprehension test (Appendix I) to complete. Both classes started this in class and had to complete it for homework.

When marking 9A (the control group) I marked according to a memorandum (Appendix I) and assigned marks in the margins. For incorrect answers I made a few general comments, for example, “look up the word elapsed” or, “your answer only states one fact and not two”. I then gave a total mark at the bottom of the page and wrote down the numbers of the questions that I wanted them to do again. When I handed the books back to the learners I told them that they had one week to correct the answers and to leave their books on the shelf for me to mark. I reminded the group of the arrangement on the day that I had indicated.

In the experimental group, 9D, I provided the same type of general comments that were developed for 9A but no ticks, crosses or marks were provided. The learners were given their books and told to re-do the entire comprehension and hand it in in a week’s time. They were reminded to hand their work in on the day that had been agreed upon.

From the pilot study I learned that it would indeed be very difficult to withhold marks on comprehension and literature tests. There were two main reasons for this: there were too many separate items that would each need comments and the learners would be able to work out their marks to some extent based on my comments.

3.5.4. The research study

A typical model of action research is the one proposed by Kemmis and McTaggart (in Koshy 2008, p. 5). Briefly, it involves planning a change, acting and observing, reflecting on the process, reflecting and re-planning and then repeating the whole process as many times as necessary (p. 4).

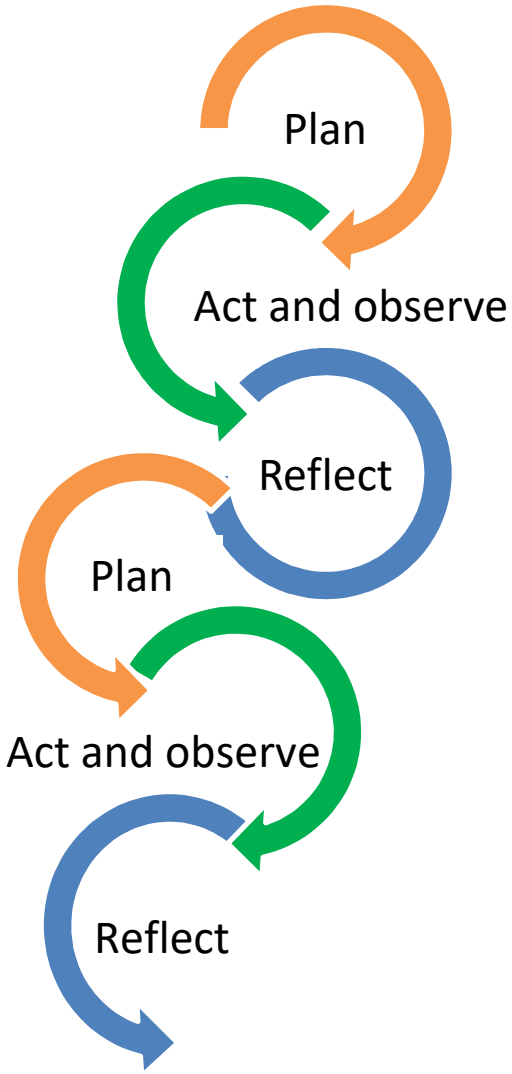


Figure 3.1. Diagrammatic representation of Kemmis and McTaggart's Spiral Model of Action Research

Table 3.1. Description of the process of the study

	First Cycle	Second Cycle	Third Cycle
Date	28 January 2016	3 May 2016	1 August 2016
Planning	Lesson on essay basics (Appendix J)	Revision lesson on basics of essay writing (Appendix K)	Agreeing on colour code to accompany comments
Action and Observation			
Date	4 February 2016	4 May 2016	3 August 2016
Task	First formal essay (Appendix N)	Second formal essay (Appendix O)	Practise essay (Appendix P)
Date	4 March 2016	6 May 2016	11 August 2016
Task	Reworking of first formal essay	Questionnaire (Appendix D) 1 June 2016 Focus group (experimental group only)	Third formal essay (Appendix Q)
Organising the data for analysis	1. Mark control group and experimental group first formal essay essays. 2. Set up Excel	1. Mark control group and experimental group second formal essays. 2. Record marks and work out	1. Mark control group and experimental group practise essays. 2. Record marks and comments

	spreadsheets to record marks (Appendix L) and comments (Appendix M). 3. Record marks and comments. 4. Compare the quantitative and qualitative results of the control group against those of the experimental group. 5. Return marked essays to learners (control group with marks, experimental group without marks). 6. Mark re-worked essays. 7. Record marks and work out class averages on Excel spreadsheet. 8. Record	class averages on Excel spreadsheet. 3. Record comments and determine how many have been successfully addressed. 4. Return marked essays to learners (control group with marks, experimental group without marks). 5. Record the number of "yes/no" responses to closed questions (1&3) on the questionnaire. 6. Decide on categories for responses to open questions (2&4). 7. Transcribe Focus Group	on Excel spreadsheet. 3. Return marked practise essays to learners (control group with marks, experimental group without marks). 4. Mark third formal essay from control group and experimental group. 5. Record marks and work out averages on Excel spreadsheet. 6. Record comments and determine how many have been successfully addressed. 7. Compare marks and averages of practise essay

	comments and note which comments have been addressed.	recordings.	with third formal essay.
Reflection	Reflect on data and decide on approach for the second cycle of action research.	Reflect on quantitative data and decide how to use the suggestions from the focus group.	Reflect on entire process and how to represent all of the data.

3.5.4.1. A detailed description of the action research cycles

3.5.4.1.i. The first cycle of Action Research

a. Planning

When approaching the actual **research study,** I decided to only use essay tasks for a number of reasons: Firstly, all the Grade 9 learners are, in theory, familiar with the structure and purpose of a creative essay as they were assessed on this in Grade 8. Secondly, essays are written under test conditions and the topics are only seen in the time allocated for the test. This means that learners are not able to prepare beforehand and therefore marks for essays are not affected by the learner's ability to remember or understand any content taught in class or the availability of help at home or additional resources like internet, study programs or study guides.

This, in turn, helped me narrow down the criteria that my comments would address because they would be based on the skills of writing. This would help me to align the comments that I had designed to match the way in which Brookhart (2008) suggests that criterion-referenced feedback be formulated. Brookhart suggested that criterion-referenced feedback should help the learner to regulate their own performance based on how close they are to the target requirements (p. 23). Furthermore, I was the only one marking the essays for my classes and a degree of standardisation could be assured.

b. Action and Observation

The first essay task (Appendix N) was written under test conditions on the 4th of February 2016. In the lessons leading up to the task both classes were given a revision lesson on the

basics of essay writing (Appendix J). The main focuses of the lesson were the structure of an essay and of paragraphs, the use of strong verbs and interesting vocabulary and the balance between a good narrative and vivid descriptions. The lesson was designed to address the three assessment criteria for writing on the GDE rubric, namely structure, language/vocabulary and content. Both classes had exactly the same lesson and completed exactly the same practice examples. 9D was somewhat slower and the last task on the use of strong verbs in vivid descriptions had to be completed at home.

When it came to marking the essays for the control group 9A, I corrected the text of each essay, gave each learner a comment on structure, vocabulary and the accuracy of the content, describing the ways in which the learners could improve their work to meet the assessment criteria as Brookhart (2008, p. 28) recommends. I then assigned a mark using the Gauteng Department of Education rubric for essay writing (Appendix A). which uses the same criteria of structure, vocabulary and content that I designed my comments around.

For 9D (the experimental group) I followed exactly the same marking process except that I recorded the mark that was assigned to each learner on a mark sheet. This mark was not indicated on the essay that learners received back.

Whilst marking, I recorded the comments that I was giving in each category (structure, vocabulary and accuracy) and assigned them a code (Appendix R), for example, "S3: Make sure you use a new paragraph for a new idea". This code would be easy to record on a spreadsheet and the comments contained the same guidelines for each learner for whom they were appropriate.

I decided to use the word "accuracy" when I gave written feedback on the content of the essays. The reason for this was because I noticed that the word "content" did not alert the learners to two important aspects of writing an essay; firstly that, as a narrative or description, the essay needed to give a clear indication of a situation and series of events that followed a logical order. Secondly, the narrative must be an exploration of the topic that the reader can relate to. As a teacher, I felt that by focusing on the word "accuracy" would give the learners a better idea of what the "content" of an essay should address.

As a researcher, I do not feel that alerting learners to the focus on accuracy changes the criteria according to which the essays were marked. I felt that by focusing on the accuracy of the content in my feedback it was a way of alerting the learners to the importance of tailoring their writing for a specific audience or making their point of view clear so that it could be

understood by the reader. If, for example, a learner describes an incident that is culturally appropriate for her, she should be aware that a reader from another culture may require some background or contextual information and that this should be woven into the narrative.

When the essays were given back to the learners in both classes, I explained the reasons (stated above) why I wanted to focus more on the accuracy of the content. I also gave verbal feedback that I limited to a comment on the learners' poor use of vocabulary - which I felt was something that applied to all of the learners in general.

The learners in both classes were then told that they needed to re-work their original essays using all of the comments that they had been given. In class, I spent time with each individual learner making sure that they understood what I meant by the comment and how they could use that comment to help them re-write. Both classes were able to make a start on their essays but had to complete them at home and hand in the following day.

c. Organising the data for analysis

I set up an excel spreadsheet for the control group and experimental group respectively (Appendix L). On these spreadsheets, I recorded the marks for the first formal essay, the re-worked essay and the comments that each learner had received for the first formal essay and the re-worked essay respectively.

I calculated the average mark for each task for each group as part of the quantitative data presentation. For the qualitative data, I checked the comments that each learner had received for the first formal essay against the comments that they received for the re-worked essay. All of this data has been transferred onto tables for the data presentation chapter.

d. Reflecting as part of the research process

As a researcher and a teacher, I had quite a naïve expectation that the intervention of withholding marks was going to produce significant results from the beginning. My reflection led me to understanding the intervention as a process rather than an immediate cure.

3.5.4.1.ii The second cycle of Action Research

a. Planning

For **the second formal essay task** on the 4th of May 2016, once again, both classes had a revision lesson on the basics of essay writing. The focus of this lesson was mainly on the dos and don'ts of essays writing and on being aware of what the topic requires as a response. The

learners were also told that it may be useful to read over their previous two essays in preparation for the task. What I hoped to determine was whether the learners of both groups would remember and apply the comments from the previous essay.

b. Action and Observation

During the marking process for the control group I assigned comments to each learner using the standard set of comments that I had developed during the **first essay task** (Appendix R), corrected the text of the essay and assigned a mark using the GDE essay marking rubric (Appendix A). The same process was applied to the experimental group, except that I recorded the marks separately and gave the essays back with only corrections and comments.

Before giving the marked essays back to the learners, I gave out the questionnaire (Appendix D) to both groups asking them to remember their comments from the **first essay task** and re-worked essay and then describe if and how the comments had helped them with the **second essay task** they had just written.

c. Conducting the focus group

After plotting the results of the first and second formal essays for both groups, I decided that I would like to investigate the attitude of the experimental group, 9D, to the process of receiving comments without the marks. I felt that I needed more data to understand the effects of the intervention. I decided to do a focus group with 9D and simply ask for learners who felt that they wanted to talk about the process to come and join the discussion. Before the recording, I asked the learners for their permission to record them if they wished to speak and stressed that participation was voluntary. As it turned out, the whole class wanted to hear the discussion even if they did not necessarily participate.

Again, I did not realise that the possibility of scrutiny from outsiders would have a dampening effect on the spontaneity of the discussion - just like the word “research” had frightened them at the beginning of the year. I had imagined that keeping the focus group in a classroom setting, as a type of class discussion, would minimise the fear of scrutiny that the learners had.

The learners were very afraid of “making mistakes” in the focus group discussion and of what they would sound like in the recording. I explained that they were simply giving their opinions and therefore there was no right or wrong way of answering the questions.

Furthermore, the conversation would be transcribed so it would only be their words on paper and not their voices that were exposed to a wider audience.

When the focus group took place it ended up being more of a group interview (Cohen et.al 2011, p 433) than a focus group. The reasons for this were twofold; firstly, as the researcher I was concerned about the quality of the recording as it was done on a cellular phone. Therefore, in practice, the recording was done with the researcher trying to keep the phone as close to whichever participant was speaking and this gave the interaction more of an interview quality. Secondly, the classroom situation is such that learners engage in classroom discussions in the style of an interview. The question was posed, learners raised their hands indicating their wish to respond and I allocated turns. Due to all of these reasons I believe that the transcript (Chapter 4, p. 68-70) sits somewhere between a focus group and a group interview.

d. Organising the data for analysis

I recorded the marks for the second formal essay on the Excel spreadsheets that I had set up for both groups. I also recorded the comments that both groups had received. I calculated the average for each of the groups. Firstly, I compared the average marks for the control group in the **first essay task** against the average marks of the **second essay task**. Then I did the same for the experimental group.

For the questionnaire I first recorded the numbers of "yes" and "no" responses to the closed questions (questions 1 and 3). I read through the responses to question 2 and determined what the most frequent comments were. Then I counted the number of learners who had responded in the different categories. I repeated this process for question 4.

I then transcribed the focus group recording, read through it and selected data that would support my next intervention.

e. Reflecting as part of the research process

The second cycle of action research was originally planned as the end of the study. However, because of the following comment in the focus group, I had to consider extending the study.

"...I think that the mark stresses you out, but you also need some indication whether you're in the eighties or in the forties. So a good way of doing this is by using colours. So if they had to give us pink as saying, you're doing well, and if you improve in this you'll do very, very

well. And then if they had to use, let's say orange, that means that you need to start improving on this but you are still okay. And if they use red, then you'll have to improve on the comments that they've given you so you can pass."

After consultation and advice, I decided that I did need another cycle of action research to investigate the effectiveness of this intervention. I also felt that more data would be useful in coming to more informed conclusions.

3.5.4.1.iii. The third cycle of action research

a. Planning

Before the **third formal essay** on the 11th of August, I set a practise essay task. The practise essay (Appendix P) was a selection of topics that had been used the previous year for grade 9 and it was given as a task under test conditions, which meant a time limit of 50 minutes and no collaboration or planning.

b. Action and observation

The control group's practise essays were marked according to the GDE rubric (Appendix A) and marks and comments were given and recorded. The practise essays with marks and comments were returned to the learners on 10 August. The formal essay, with different topics, was written on 11 August (Appendix Q). This essay was marked using the GDE rubric and returned to the learners with marks and comments.

For the experimental group the essays were marked according to the GDE rubric and comments were given and recorded. The marks for the practise essay were recorded but not shared with the learners. The practise essays with comments and a colour indicating their urgency were returned to the learners on 10 August. The colour codes were given as follows: for marks from 30-59% orange, 60-69% yellow and 70%+ green. The formal essay, with different topics, was written on 11 August. This essay was marked using the GDE rubric and returned to the learners with comments.

c. Organising the data for analysis

I recorded the marks for the control group's practise and formal essays on the Excel spreadsheet and worked out the class average for each task. I repeated this process for the experimental group. I then recorded the comments that the control group received for the

practise essay and formal essays and determined which of them had been successfully addressed. I repeated this process for the experimental group.

As I wanted to observe the effect of the intervention on the weakest learners, I identified the learners in both groups who had achieved less than 60% for the third formal essay. I compared the marks that they received for the practise essay to the marks that they had received for the third formal essay and transferred this information onto a bar graph.

Finally, I isolated the marks that the learners from the control group had achieved for each of the formal essay tasks only. I worked out the class averages for each task and recorded this information on a table. I transferred the marks for each learner for each of the three tasks onto a bar graph so that their performance on an individual level could be seen. I repeated this process for the experimental group.

d. Reflecting as part of the research process

With all of the data recorded I could reflect on the entire research process. I now had the benefit of being able to review the effects of the intervention over a long time period and start formulating claims.

3.6. Conclusion

Taking the hypothesis of an experimental study and applying it to a real life setting, almost 30 years later, in a different educational context requires careful planning. The choice of action research as a methodology was the best fit as it provided the immediacy of observation and reflection that an experimental study does not. Furthermore, as the both the teacher and the researcher, I got to engage first hand with the research tools and participants, which would not necessarily happen in an experimental setting. The planning for the three cycles of action research required much record keeping and reflection, therefore it was a process that really encouraged a mindful approach to assessment.

CHAPTER 4: DATA DESCRIPTION

This chapter depicts and organises the data that was collected in the pilot study and each of the action research cycles. The control group and experimental group are depicted separately in each section. Furthermore the quantitative and qualitative data has also been separated to make analysis easier.

4.1.THE PILOT STUDY

The learners from the control group and the experimental group were given a short comprehension task (Appendix I) to complete.

4.1.1. Control group

The completed comprehensions from the control group were marked according to a memo (Appendix I). Ticks were used to indicate correct responses and marks were allocated in the margin next to each response. The marks were added up and a total was indicated at the end of the exercise. For incorrect responses learners received a comment that could help them to correct their answer. The learners were given their books back and told to re-do only the questions that they had not been awarded the full possible mark for and to then hand their books in again.

In the control group, 9A, 2 out of 23 learners submitted the corrections on the comprehension. The corrections were marked using the memo and the marks from the corrected responses were added to the total that they received for the first attempt.

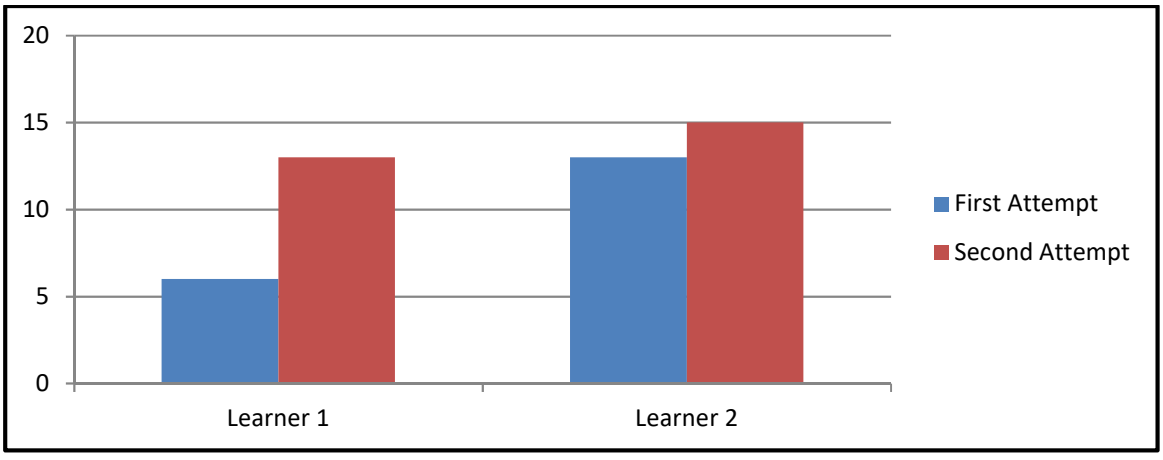


Figure 4.1.a Graph demonstrating learners' marks in the pilot study from the control group.

The marks of both of the learners who re-submitted their work improved.

4.1.2. Experimental Group

The completed comprehensions from the experimental group were marked according to a memo. No mark of any kind was used on the learners' work to indicate whether an answer was correct or incorrect. The total number of marks was tallied and recorded in a mark book but no total was given to the learners. Learners received comments that could assist them in correcting incorrect responses.

Learners were given their books back and told to re-do the entire comprehension and hand it in again. 18 of the 23 learners in the experimental group, 9D, handed their books in.

The second attempt at the comprehension was re-marked using the memo. This time ticks were used to indicate correct answers, the mark for each answer was indicated in the margin, these marks were tallied and a total given at the end of the exercise. These marks were then recorded next to the original mark.

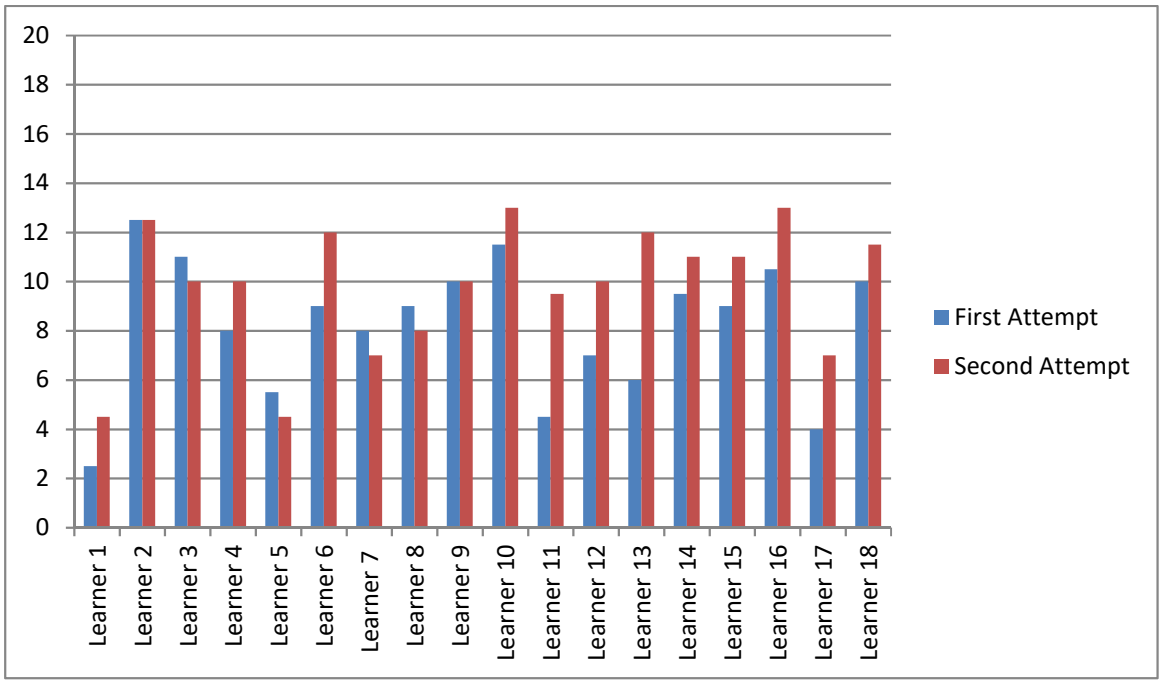


Figure 4.1.b Graph demonstrating learners' marks in the pilot study from the experimental group.

Twelve of these learners improved their performance, two remained the same and four of the learners got lower marks on the second attempt.

4.2. FIRST CYCLE OF ACTION RESEARCH

4.2.1. The Quantitative Data

a. The Control Group

The first essay task (Appendix N) was written on 4 February 2016. The control group, 9A, was awarded a mark using a rubric (Appendix A) and given comments describing how their performance could be improved. This group was then given the opportunity to re-work the same essay, using the comments to assist them, and then re-submit the task. The re-worked essay was then given a mark using the same rubric as before.

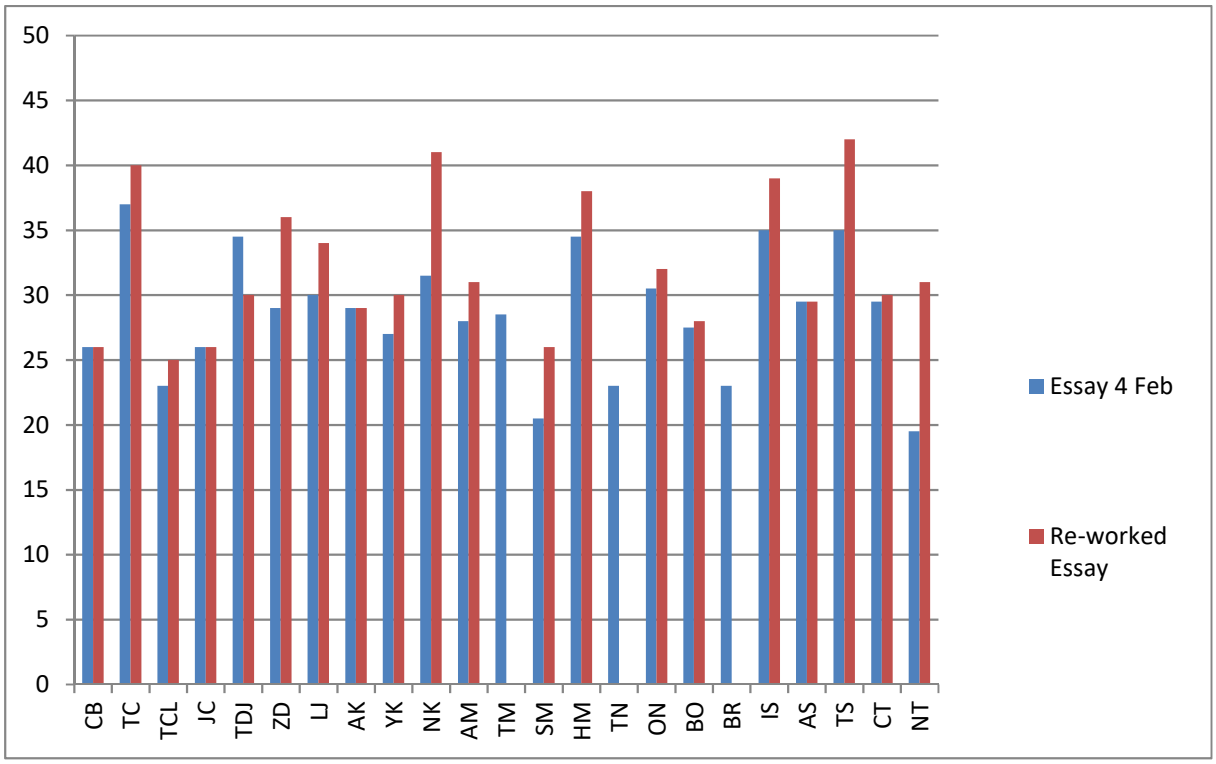


Figure 4.2.1.a. Bar Graph comparing the marks of the first essay task with the marks of the re-worked essay. Control Group.

Explanation of results for the control group

Three of the learners from the control group did not submit the re-worked essay and therefore no mark could be awarded. Fifteen of the learners improved their mark in the second attempt, four learners' marks stayed the same and one learner's mark decreased. The class average for the first essay was 28.56 out of a possible 50 marks and the average for the remarked essay was 32.17 showing an average increase of 3.61 marks or 7.2%.

b. The Experimental Group

The first essay task (Appendix N) was written on 4 February 2016. The experimental group, 9D, was awarded a mark using a rubric (Appendix A) but this mark was recorded in the teacher's mark book only and not revealed to the learners. The learners were only given comments describing how their performance could be improved when their essays were returned to them. This group was then given the opportunity to re-work the same essay, using the comments to assist them, and then re-submit the task. The re-worked essay was then given a mark using the same rubric as before.

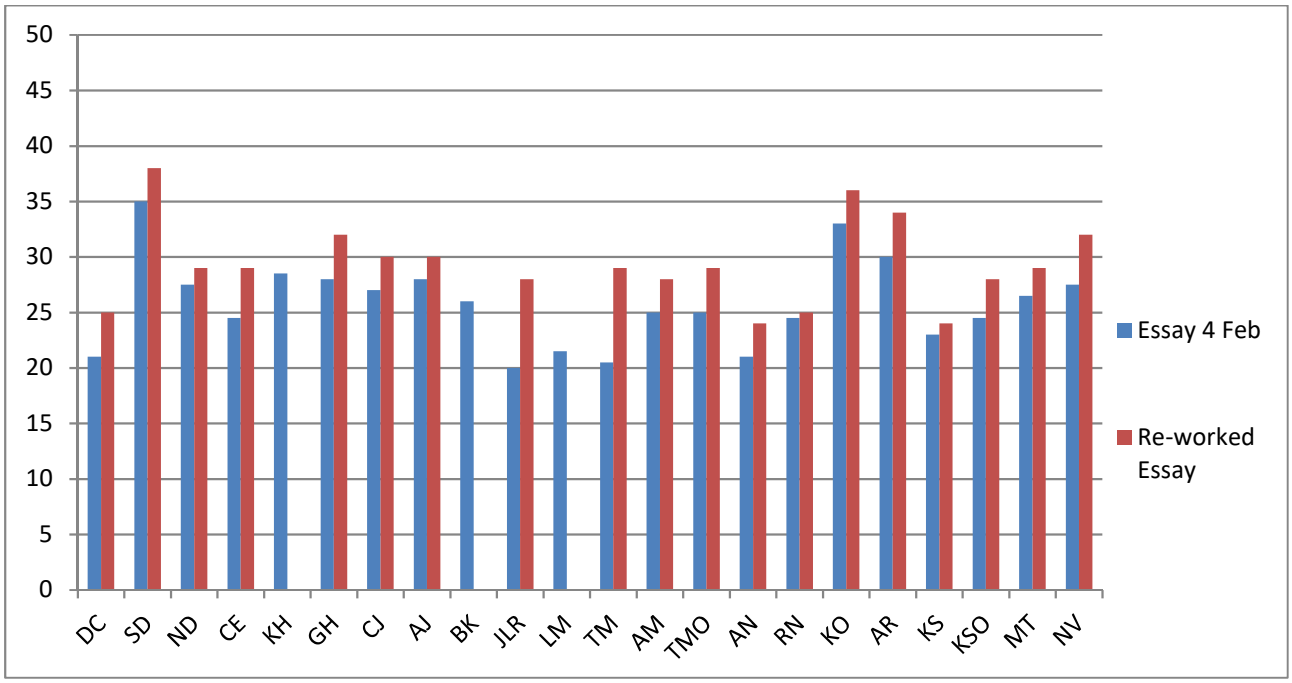


Figure 4.2.1.b. Bar Graph comparing the marks of the first essay task with the marks of the re-worked essay. Experimental Group.

Explanation of the results for the experimental group

Three learners from this group did not submit a re-worked essay and could not be awarded a mark. Nineteen of the learners improved their mark in the second attempt. The class average for the first essay was 25.7 out of a possible 50 marks and 29.4 for the second. This showed an average increase of 3.7 marks or 7.4%.

4.2.2. The qualitative data

a. The control group's comments

When the learners wrote the reworked essay they had the comments from the previous essay with them. Therefore this was not a test of how well the comments were remembered, rather it gives an indication of how well the learners could apply the comments that they were given.

Whilst marking the re-worked essay, the researcher took note of how many of the original comments were addressed. The researcher indicated that a comment had been addressed when it was clear that the learner had tried to apply the feedback to their work, regardless of whether it was partially or fully successful.

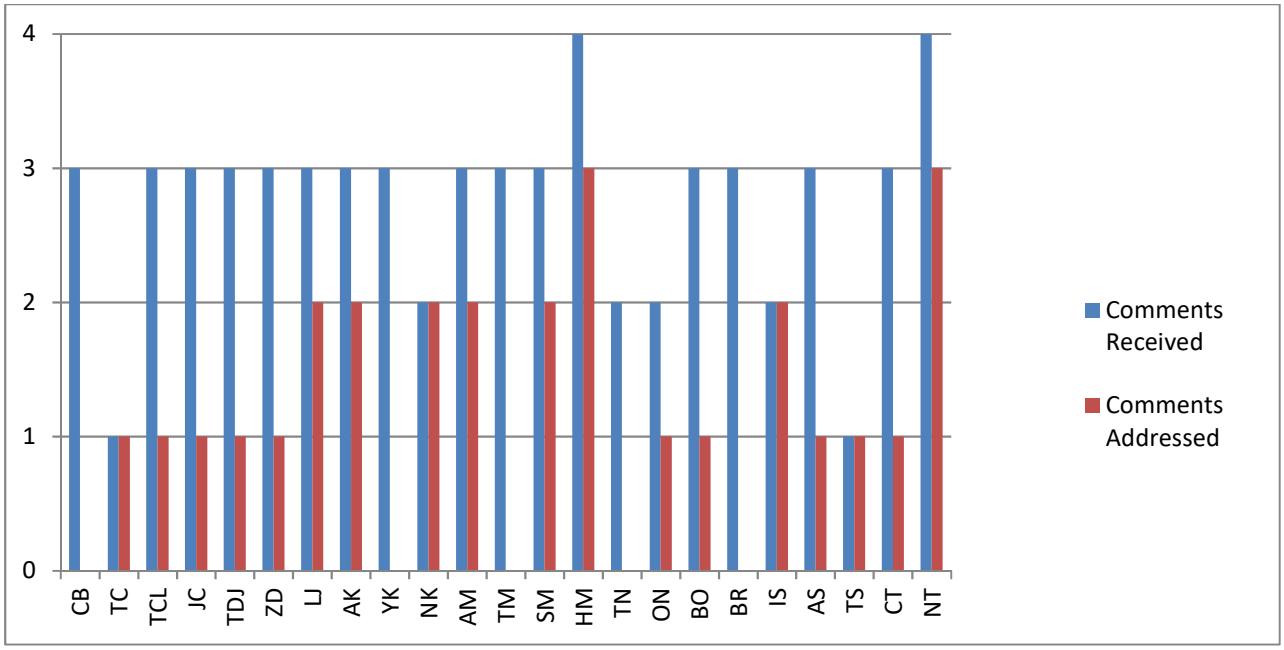


Figure 4.2.2.a. Bar graph depicting the number of comments given to the number of comments addressed. Control Group.

Explanation of bar graph

Each learner was given comments on the work that they had presented. The maximum number of comments given was four and the highest number of comments that were addressed was three. The three learners who did not address any comments did not submit the re-worked essay. Therefore each learner who submitted was able to address at least one comment in the re-worked essay.

Detailed depiction of comments for the control group

63 Comments were given to the control group in total. As three learners did not submit the re-worked essay, and it could not be determined if they had addressed the comments given to them or not, the researcher has used the number of comments given to the learners who did re-work their essays. 55 comments were given; 17 regarding accuracy, 19 for structure and 20 relating to vocabulary. 26 of these comments were successfully addressed; 12 relating to accuracy, 10 for structure and 6 for vocabulary.

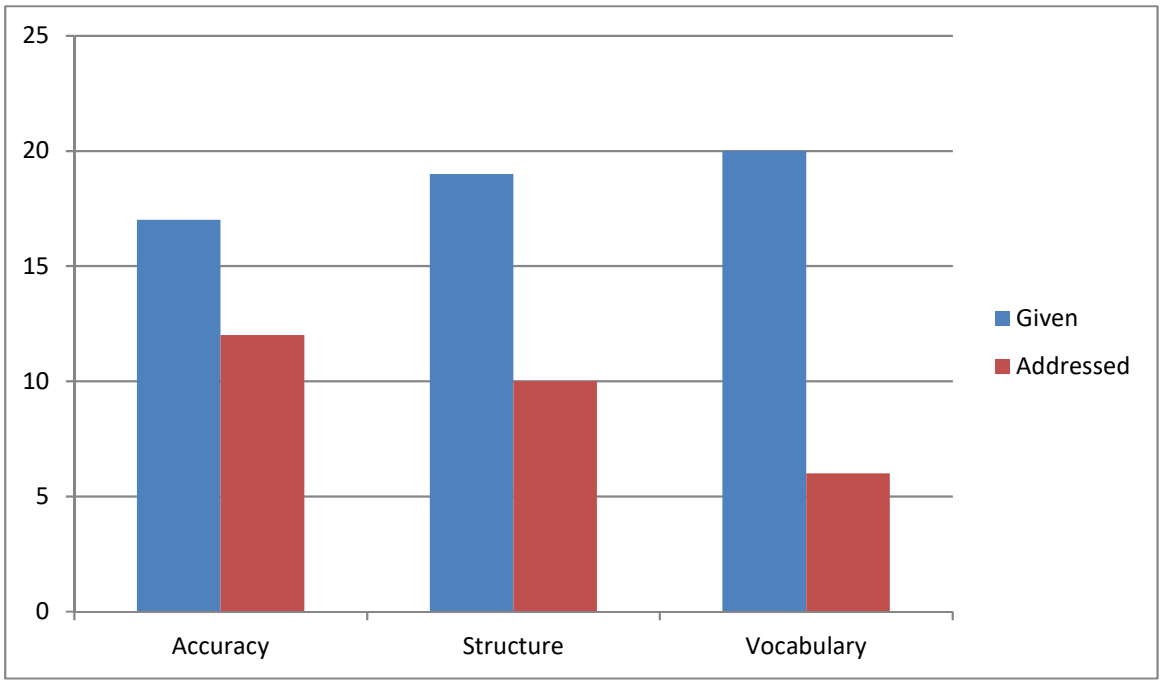


Figure 4.2.2.a.(i) Bar graph comparing the number of comments given in each category to the number of comments addressed in each category. Control Group.

The control group addressed 26 out of the 55 comments they received, 47.2%. 70.5% of the comments regarding accuracy were successfully addressed, 52.6% for structure and 30% for vocabulary.

b. The experimental group's comments

When the learners wrote the reworked essay they had the comments from the previous essay with them. Therefore this was not a test of how well the comments were remembered, rather it gives an indication of how well the learners could apply the comments that they were given.

Whilst marking the re-worked essay, the researcher took note of how many of the original comments were addressed. As with the control group, the researcher indicated that a comment had been addressed when it was clear that the learner had tried to apply the feedback to their work, regardless of whether it was partially or fully successful.

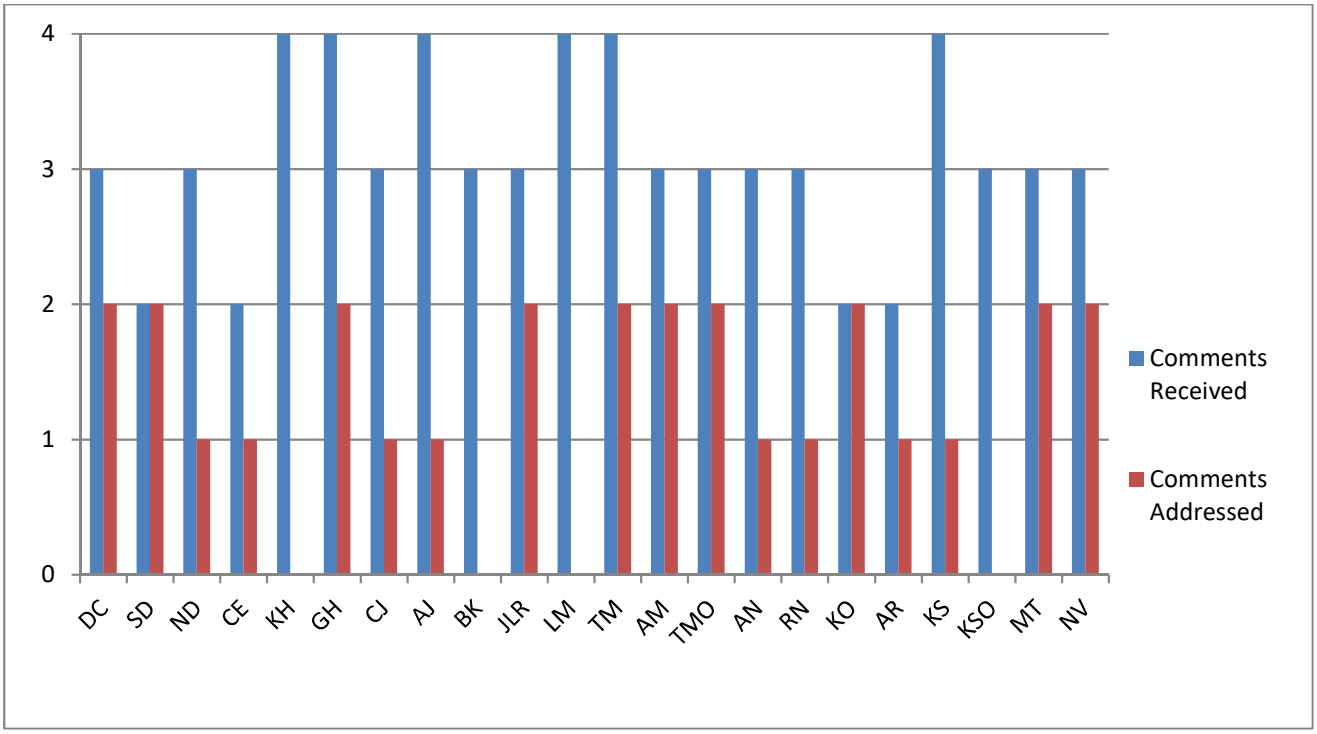


Figure 4.2.2.b. Bar graph comparing the number of comments received by each learner to the number of comments addressed. Experimental Group.

Explanation of bar graph

Each learner was given comments on the work that they had presented. The maximum number of comments given was four and the highest number of comments that were addressed was two. Four learners did not address any comments that they were given. Three of these learners did not submit the re-worked essay and the one did not address the comments in the re-worked essay.

Detailed depiction of comments for the experimental group

68 comments were given to the experimental group in total. From the learners who submitted the reworked essay, 57 comments were given and 28 of these were addressed. 18 were given for accuracy and 13 were addressed 18 for structure and 11 addressed and 20 were given for vocabulary and 5 only addressed.

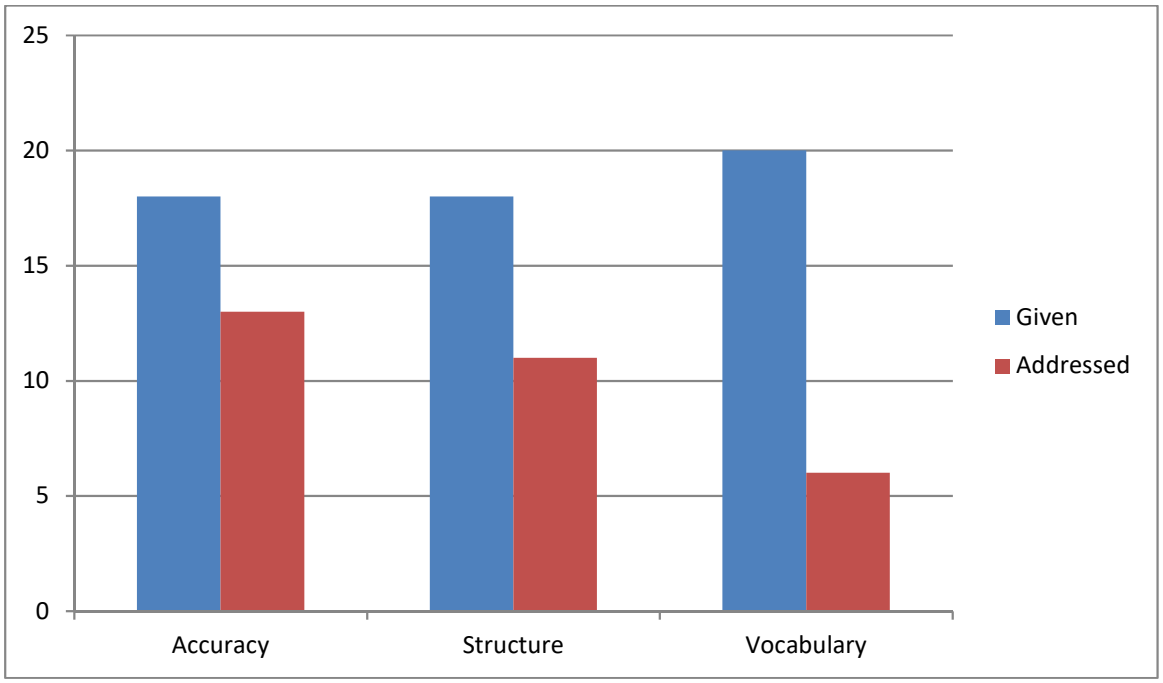


Figure 4.2.2.b.(i). Bar graph comparing the number of comments given in each category to the number of comments addressed in each category. Experimental Group.

The experimental group addressed 28, 49%, of the 57 comments given to the group as a whole in their re-worked essay. 72.2% of the comments for accuracy were addressed, 61% of the comments for structure were addressed and 25% of the comments for vocabulary were addressed.

4.3. SECOND CYCLE OF ACTION RESEARCH

4.3.1. Quantitative data

a. The control group

In the second cycle of action research, the learners wrote another formal essay on 4 May 2016 (Appendix O). This essay was marked using the GDE rubric (Appendix A). The mark that each learner received for the essay on the 4th of February was then compared to the mark that each learner received for the essay on the 4th of May.

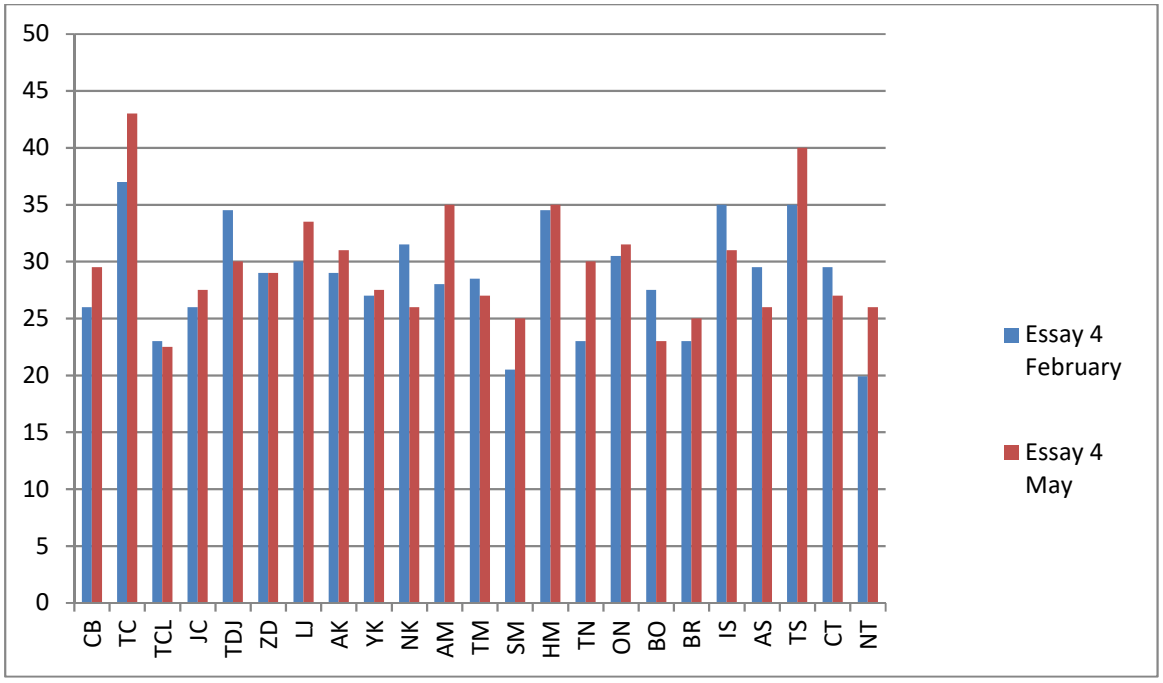


Figure 4.3.1.a. Bar graph comparing the marks for the first formal essay with the marks for the second formal essay. Control group.

14 of the learners improved their performance from the first essay to the second. One learner received the same mark and 8 learners received lower marks for their second essay. The class average for the essay on the 4th of February was 28.5 out of 50, 57%, and the average for the second essay was 29.6 out of 50, 59.2%. Thus the control group's marks improved by 2.2% in the second essay.

3.1.b. The Experimental Group

In the second cycle of action research, the learners wrote another formal essay on 4 May 2016 (Appendix O). This essay was marked using the GDE rubric (Appendix A). The mark that each learner received for the essay on the 4th of February was then compared to the mark that each learner received for the essay on the 4th of May.

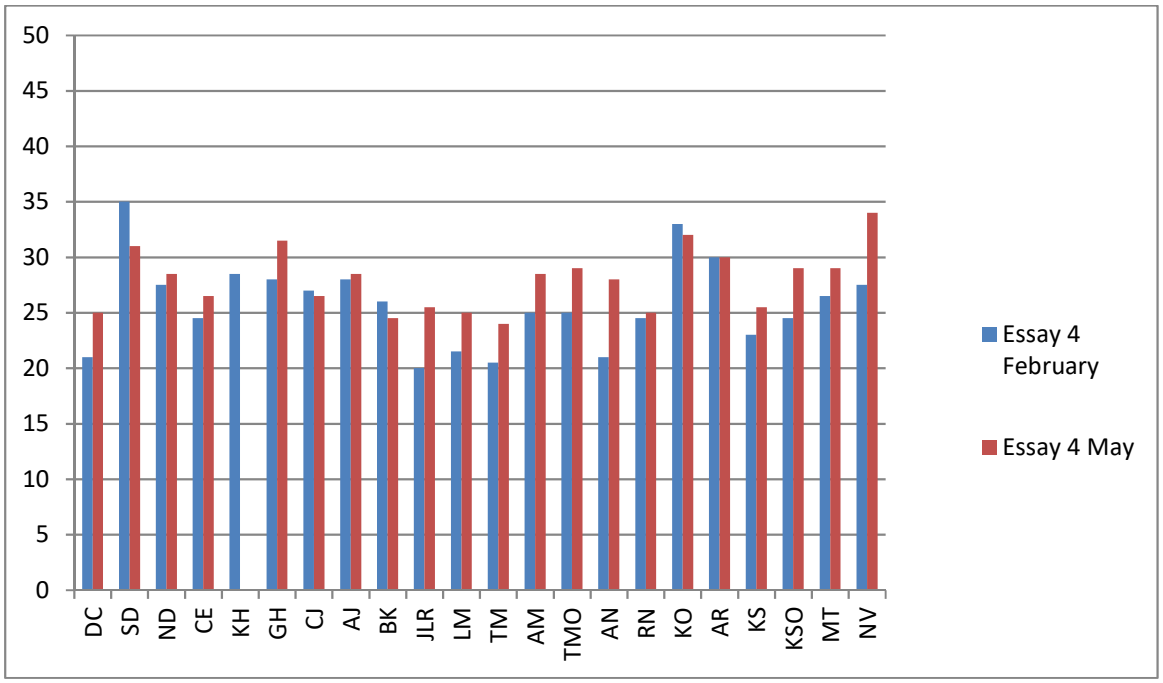


Figure 4.3.1.b. Bar graph comparing the marks for the first formal essay with the marks for the second formal essay. Experimental group.

21 of the 22 learners in the experimental group wrote the essay on the 4th of May. 16 of the learners improved their mark, one received the same mark and 5 received lower marks for the second essay. The class average for the first essay was 25.7 out of 50, 51.8% and the average for the second essay was 27.9 out of 50, 55.8%. The experimental group's marks improved by 4.4% from the first essay to the second essay 2.2% more than the improvement shown by the control group.

4.3.2. Qualitative Data

a. The control group's responses to the questionnaire

i. Responses to Question 1

Do you find it easy to remember the comments that you have received on your work so far?

I gave the control group the same questionnaire as the experimental group (Appendix D) on 6 May after the second formal essay was written. 20 learners answered the questionnaire (3 learners were absent from school that day).

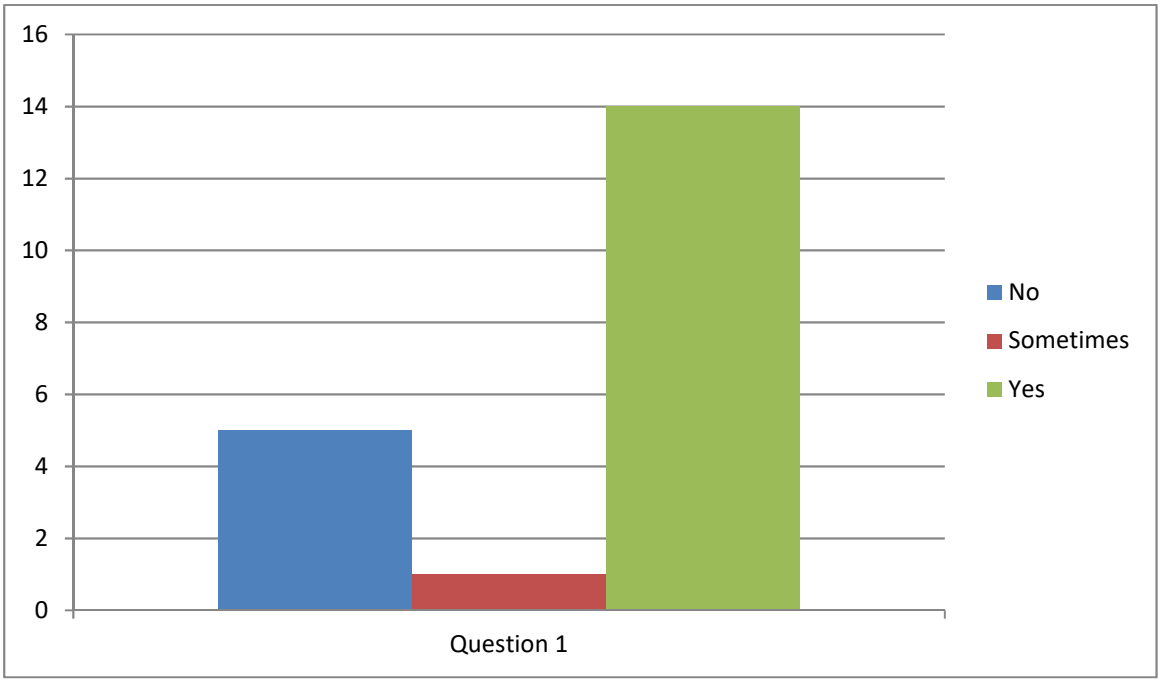


Figure 4.3.2.a (i). Control group's responses to question 1.

In the control group, 5 learners responded that they did not find it easy to remember the comments, one responded, "sometimes" and 14 responded, "yes".

ii. Responses to Question 2

Write down as many of these comments as you can remember in the space below.

I checked the comments that the learners had written against the record of the comments given for the first essay and addressed in the re-worked essay.

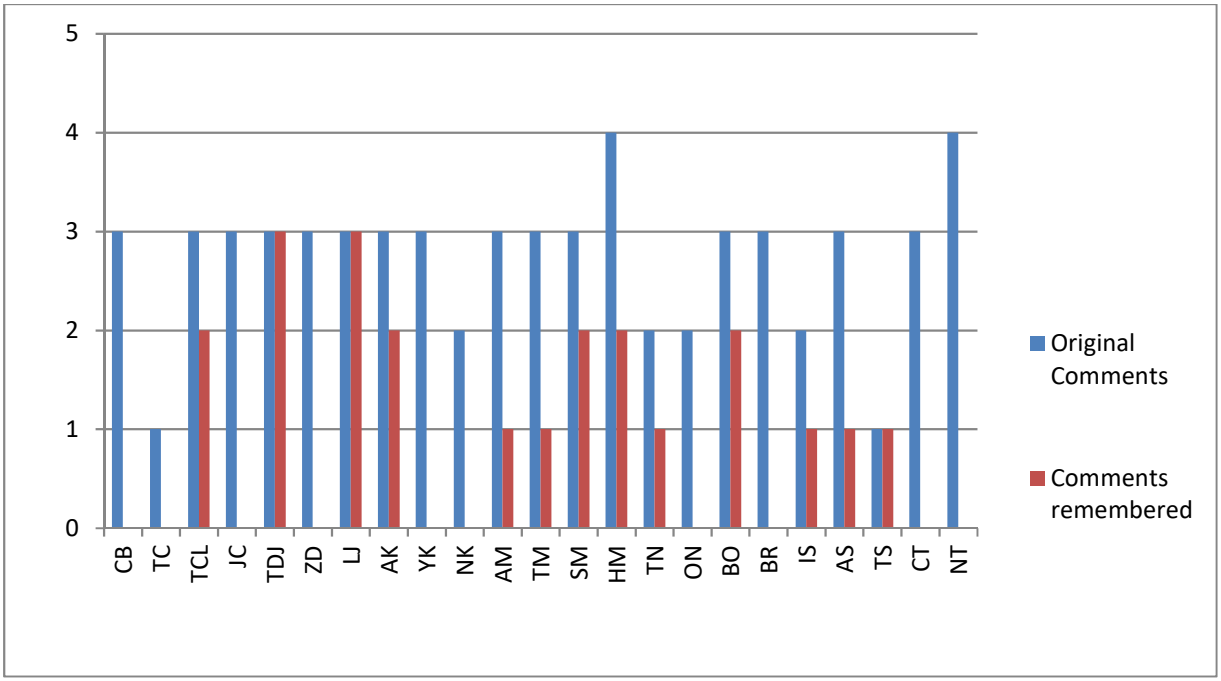


Figure 4.3.2.a.(ii). Bar graph comparing the number of comments remembered to the number of comments originally given. Control group.

20 of the 23 learners in the control group answered the questionnaire. 7 learners did not remember any of the comments that they had been given, 6 learners remembered one comment correctly, 5 remembered 2 comments and 2 remembered 3 comments.

iii. Responses to Question 3

Do you feel that any of these comments have helped you to improve your work?

Of the 20 learners who answered the questionnaire (Appendix D) 18 replied "yes" to the question and two replied "a little" to this question.

iv. Responses to Question 4

If so, how?

I divided the responses to this question into four categories; "Question not understood", "Comment didn't help", "Comment did help (non-specific)" and "Comment did help (specific)". The researcher differentiated between the responses where learners simply stated that the comments helped them improve or correct mistakes and those where learners gave specific information on which comments helped them and how.

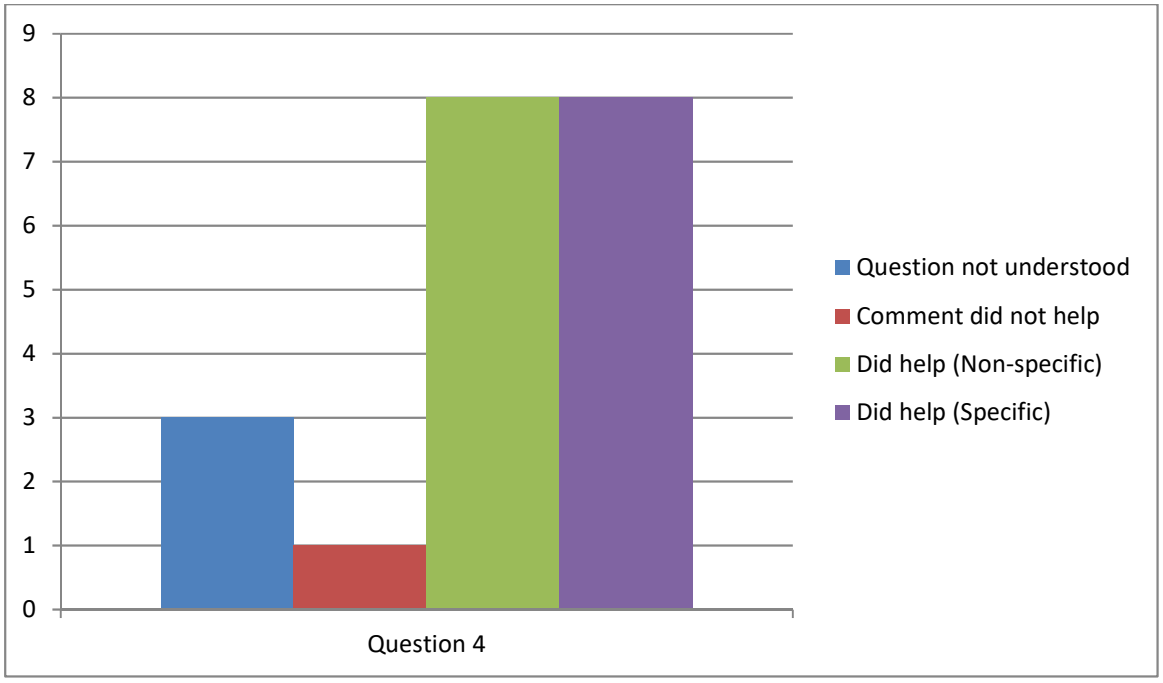


Figure 4.3.2.a.(iv). Bar graph depicting responses to question 4 of the questionnaire. Control group.

Three learners did not appear to understand the question as their responses did not even mention the comments themselves. One learner stated reasons why the comments did not help, 8 learners gave non-specific explanations of how the comments helped and eight learners gave specific explanations.

4.3.2.1. a. Comments applied compared to comments remembered in the second formal essay.

During the process of marking the second formal essay I made a note of which of the original comments had been addressed and compared them to the number of comments originally given and then to the number of comments explicitly remembered by each learner.

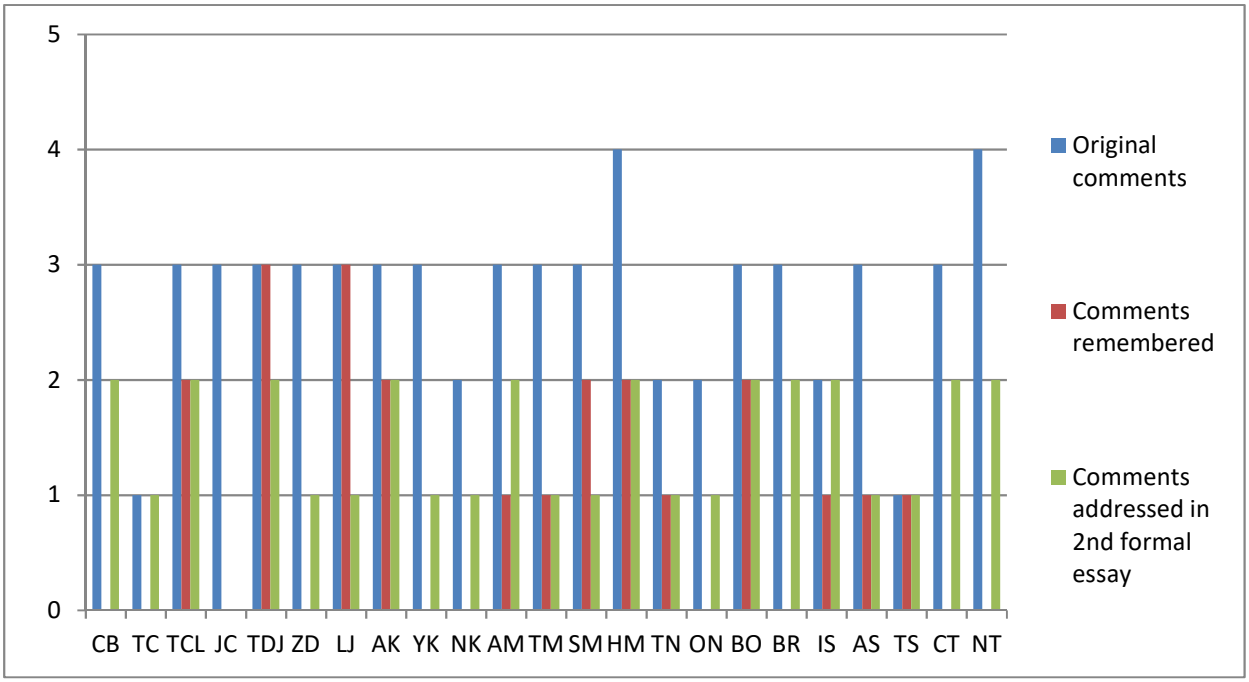


Figure 4.3.2.1.a. Graph depicting the original number of comments compared to comments remembered and comments applied by each learner. Control group.

In the second formal essay an interesting phenomenon emerged in the control group; learners seemed to apply information from the comments to their essays but did not necessarily remember being given that information in the form of the comments that they had originally received. One learner neither applied nor remembered any comments, 13 learners addressed and remembered comments from the original essay and 9 learners addressed comments to their second essay but did not explicitly remember having been given them in the original essay.

4.3.2.2. a. Longitudinal depiction of comments received, remembered and applied for the control group.

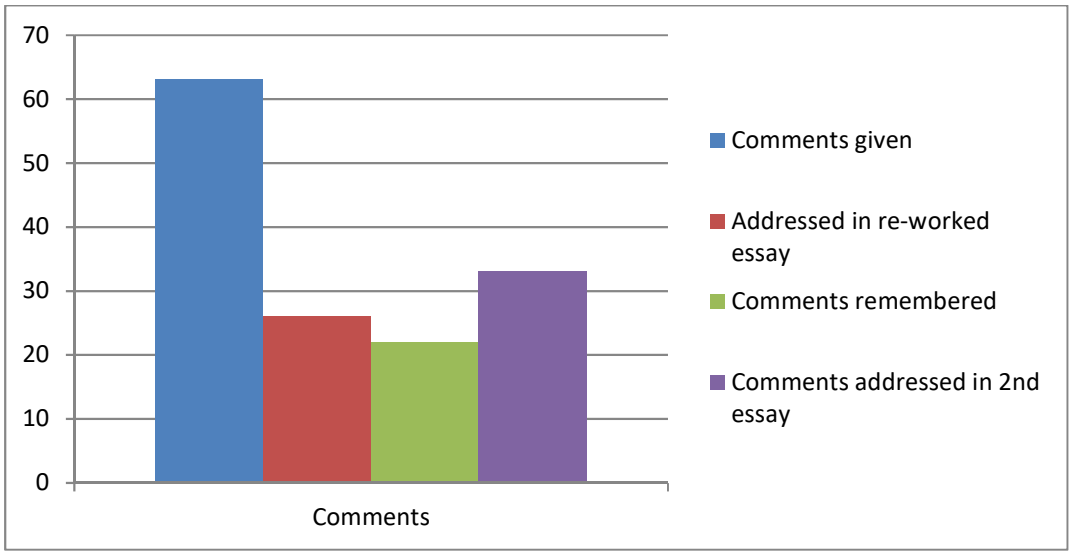


Figure 4.3.2.2.a. Graph depicting how comments were remembered and applied over time. Control Group.

The control group were originally given 63 comments. In the re-worked essay, as a group, they were able to address 26 out of 55 (47%) of these comments. When they wrote the second formal essay, three months later, they were able to address 33 (52%) of the comments but only explicitly remembered 22 (35%).

4.3.2.b. The experimental group's responses to the questionnaire

I gave the experimental group the same questionnaire as the control group (Appendix D) on 6 May after the second formal essay was written. 21 learners answered the questionnaire (2 learners were absent from school that day).

i. Responses to Question 1

Do you find it easy to remember the comments that you have received on your work so far?

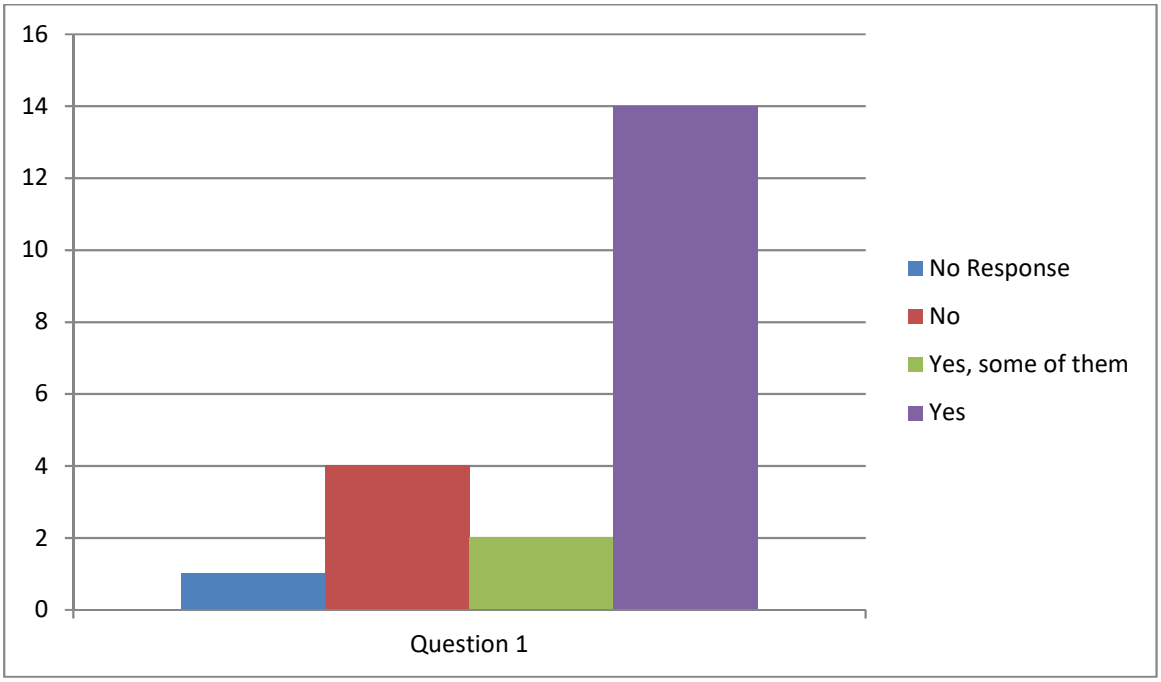


Figure 4.3.2.b. (i). Bar graph depicting the number of learners who say that they do not remember comments compared to those say that they do.

One learner did not respond to the question, four learners said that they did not remember any of the comments that they had received, two learners stated that they remembered some comments and 14 simply said, "yes".

ii. Responses to Question 2

Write down as many of these comments as you can remember in the space below.

I checked the comments that the learners had written against the record of the comments given for the first essay and corrected in the re-worked essay.

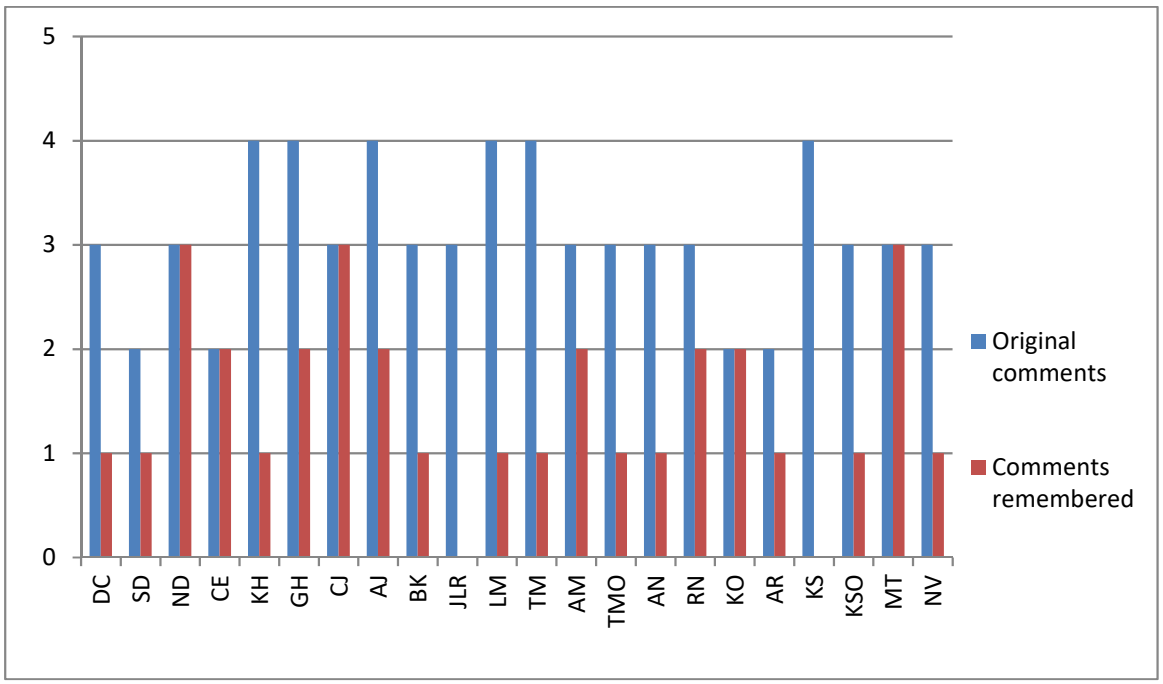


Figure 4.3.2.b(ii). Bar graph comparing the number of comments remembered to the number of comments originally given. Experimental group.

One of the learners did not remember any of the comments, 11 remembered one comment, 6 remembered 2 comments and 3 learners remembered 3 comments.

iii. Responses to Question 3

Do you feel that any of these comments have helped you to improve your work?

Of the 21 learners who answered the questionnaire (Appendix D) 19 replied "yes" to this question and two replied "no" to this question.

iv. Responses to Question 4

If so, how?

I divided the responses to this question into four categories; "Question not understood", "Comment didn't help", "Comment did help (non-specific)" and "Comment did help (specific)". The researcher differentiated between the responses where learners simply stated that the comments helped them improve or correct mistakes and those where learners gave specific information on which comments helped them and how.

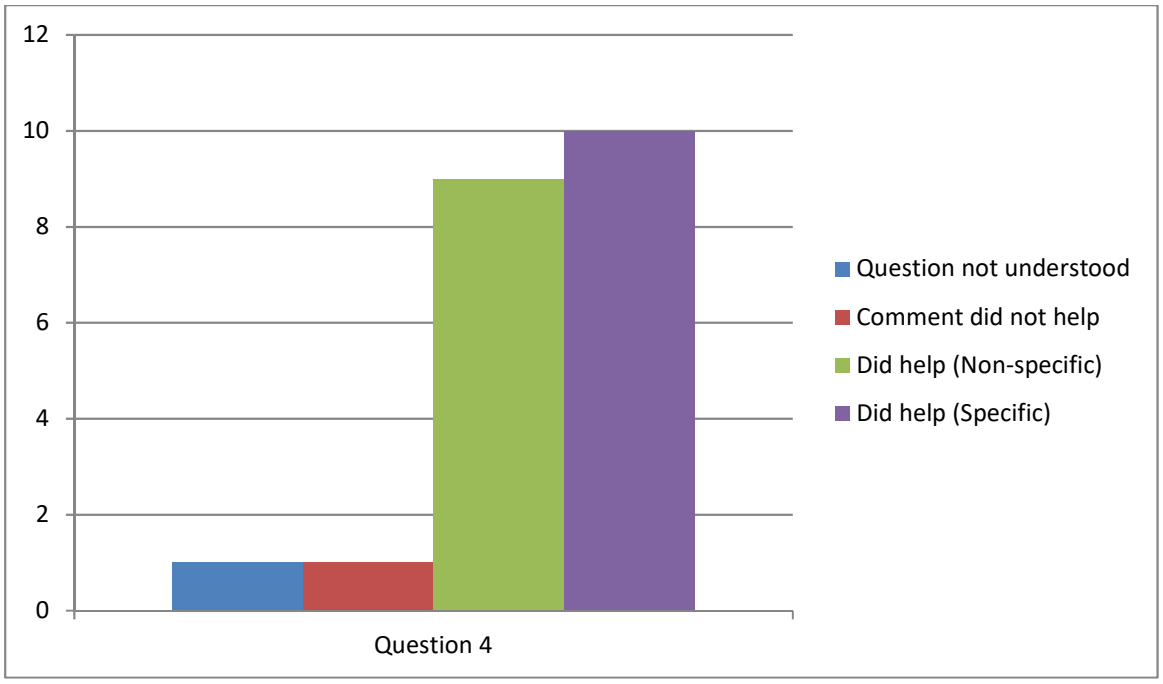


Figure 4.3.2.b(iv). Bar graph depicting responses to question 4 of the questionnaire. Experimental group.

One learner did not understand the question and one learner responded that the comments had not helped. Of the learners who stated that the comments did help, 9 gave non-specific answers that simply stated that the comment had helped them improve or correct mistakes and 10 gave specific answers on which aspects of their work they had started focusing on in order to improve.

4.3.2.1.b. Comments applied compared to comments remembered in the second formal essay.

During the process of marking the second formal essay I made a note of which of the original comments had been addressed and compared them to the number of comments originally given and then to the number of comments explicitly remembered by each learner.

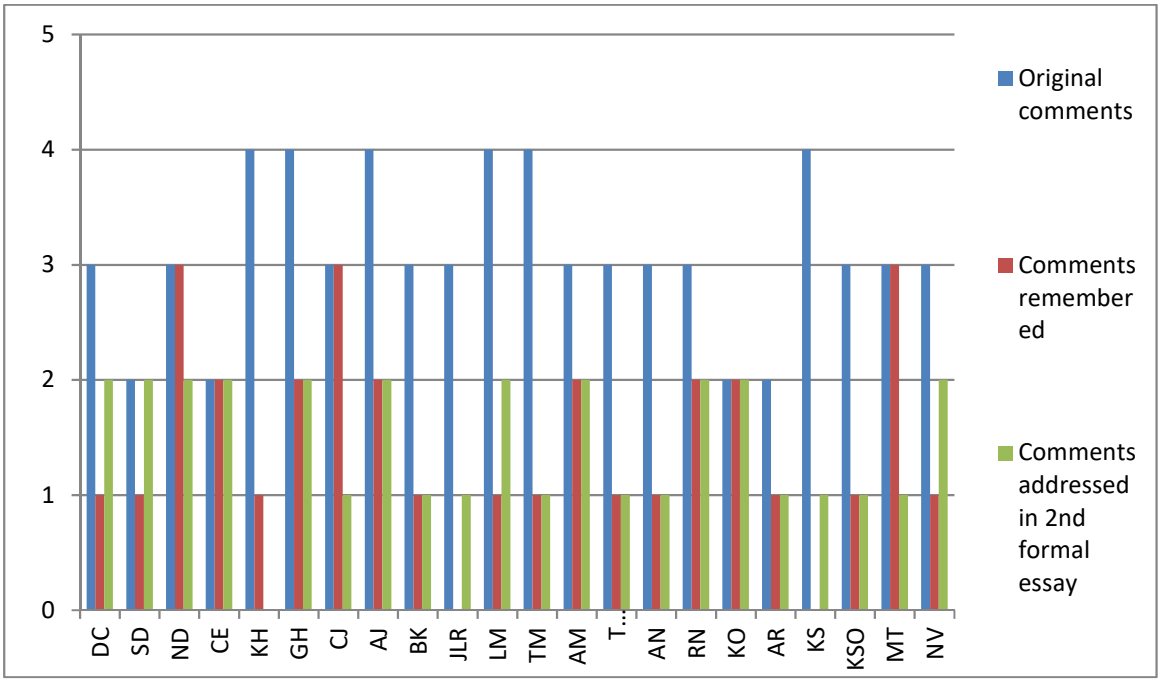


Figure 4.3.2.1.b. Graph depicting the original number of comments compared to comments remembered and comments applied by each learner. Experimental group.

As with the control group, the experimental group seemed to be able to address some of the original comments in their work without explicitly remembering them. Six learners seemed to address comments in the second essay without explicitly remembering them and 19 learners both remembered and addressed at least one comment. KH did not write the second essay and so had no opportunity to address comments.

4.3.2.2.b. Longitudinal depiction of comments received, remembered and applied for the experimental group.

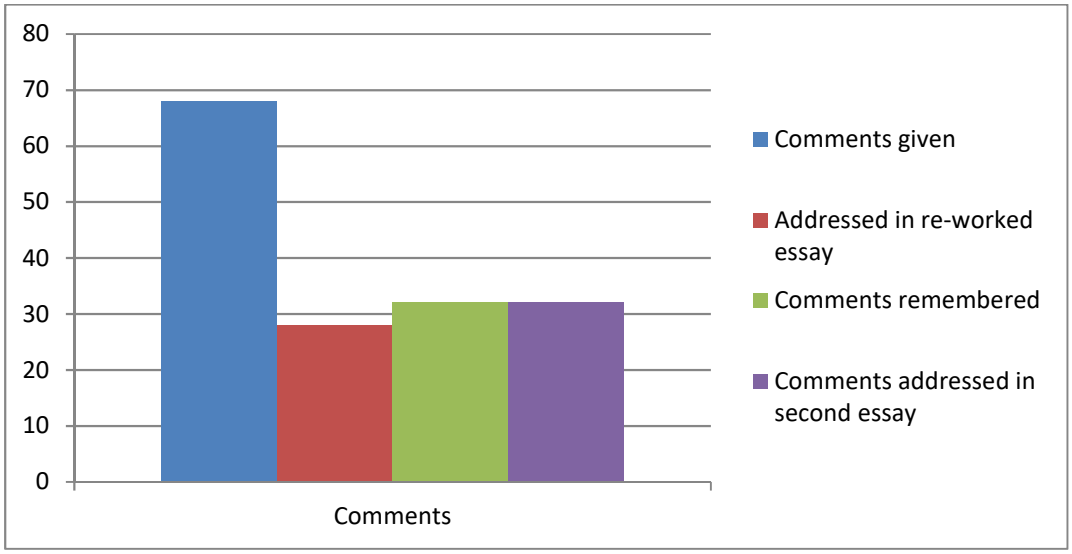


Figure 4.3.2.2.b. Graph depicting how comments were remembered and applied over time.

The experimental group were originally given 68 comments. In the re-worked essay as a group they were able to address 28 out of 56 (50%) of these comments (as 3 learners did not submit the reworked essay). When they wrote the second formal essay, three months later, they were able to address 32 out of 64 (KH did not write) (50%) of the comments and explicitly remembered 32 (50%).

4.3.3. FOCUS GROUP WITH EXPERIMENTAL GROUP

Duration: about 6 mins Transcribed by Cané Lake

Colour Coding:

Yellow= motivation for keeping marks secret

Green= suggestions for future tasks

Gwen	What is it like to get an essay back where there are only comments and no marks? What do you think and feel?
Learner	My name is Mastermind. It feels weird, because like you don't know if you did good or if you failed, and you just think the teacher was just being sweet and she didn't want to write the mark and she just gave you comments to improve.
Learner	My name is Casey Harper, and I believe that if there's just comments, you can see more of what you could have done, rather when you see your mark you just think, well I failed, or like I excelled and I don't need to improve, or now you know what you can improve on.
Gwen	When you sit down for the next task and you still don't know your mark from the previous task, but you have some comments, what are you…
Learner	Okay, so when I go for my next task, I think about the comments and not about the mark. Like I'm not stressing because I failed the last one. I'm concentrating on the comments so I can improve.
Learner	My name is Tulip, and I think it's okay. I prefer to get the marks as well because then I know how much I need to improve by, so with the comments I know with what I need to improve by.
Gwen	Having been through this process of getting comments but no marks,

	would you recommend that other people go through this process as a learning process, and also tell me why…?
Learner	I think it's a good idea, because normally when you get the test back you always look at the mark and you forget to…like you read the comment but it doesn't really like go through your mind. Like what goes through your mind is the mark, so it's better to get a comment rather than a mark, so you know how to improve in the future.
Learner	Okay, so I think it depends on the person. If you think you did well, then you would like your mark back. And if you think that you did badly the comment will help.
Learner	Okay, so maybe like if you get the comment you might be thinking that you will be…so when you get a comment you're driven by the motivation and if you get a low mark and the comment maybe you'll be thinking of improvement for your essay. But if you get a high mark maybe you will think, no, it's cool, I can ace this.
Gwen	What's coming up next are general comments and questions that were outside the focus…
Learner	I think us as students will feel more comfortable as teachers not marking to a memo, more reading it, so they'll understand, they'll actually read our essays without using a memo and marking it to the memo.
Learner	Okay, so I think that maybe when you just get a mark and the teacher didn't mark according to the rubric, it lets you feel confused because you're not sure of where you went wrong.
Learner	I feel like comments are important, but marks are just as important for us and our parents, because our parents care more about marks than a comment. Comments may be helping us but our parents will be more concerned that there's no mark.
Gwen	So that makes me want to ask, what is really important to you?

Learner	What's really important to me are comments mostly, because I can see what I've done wrong.
Learner	I feel like it would be better if we had practice tests where we get comments and we get to see where we're going wrong and where we should add in, and then we get one final one where we get a mark.
Learner	I think it's worthwhile because we get to know where we've improved and some of us don't know just by the mark where exactly we need to improve. It also helps with like…because sometimes we just get corrections and sometimes it's not specific. So like we could have like a line going through symbolising new paragraph but sometimes we don't know. So comments also help.
Learner	I feel like if you failed your first test, and they just gave you the comments but the comment doesn't reflect that you failed, so you don't try as hard for the next test, because you thought you did okay. But giving you the mark would help you and would make you want to do better for the next…
Learner	I feel we should get the marks and the comments because some people are motivated by the mark and not the comments…well, both of them, but…ja…
Learner	So my idea was that instead of giving us marks out of something, it would be easier if you put like little smiley faces, so a happy face for, okay, you're doing well, a straight face meaning you need to work harder, and a really sad face meaning you've gone completely wrong, which means like you need to improve.
Learner	So I think that the mark stresses you out, but you also need some indication whether you're in the eighties or in the forties. So a good way of doing this is by using colours. So if they had to give us pink as saying, you're doing well, and if you improve in this you'll do very, very well. And then if they had to use, let's say orange, that means that you need to start improving on this but you are still okay.

	And if they use red, then you'll have to improve on the comments that they've given you so you can pass.
Learner	Okay, so I think that teachers need to think more about the comments that they are writing on our essays, because sometimes we don't know where we went wrong and they're kind of vague to us…so…

4.4. THIRD CYCLE OF ACTION RESEARCH

I decided to conduct a third cycle of action research to test the hypothesis offered by GH in the experimental group; that if learners had an idea of what the level of their writing was, they would be able to apply the comments better.

The third cycle of action research involved a practise essay for both groups in which the control group would receive marks and comments and the experimental group would receive a colour code and a comment.

4.4.1. Quantitative Data

a. Control group

I gave the learners a set of practise essay topics (Appendix P) to choose from. The practise essays were written during class time on 2 August and handed in. The control group's essays were marked according to the GDE rubric (Appendix A) and comments were given and recorded. The practise essays with marks and comments were returned to the learners on 10 August. The formal essay (Appendix Q), with different topics, was written on 11 August. This essay was marked using the GDE rubric and returned to the learners with marks and comments.

The marks given for the practise essay were compared to those given for the formal essay.

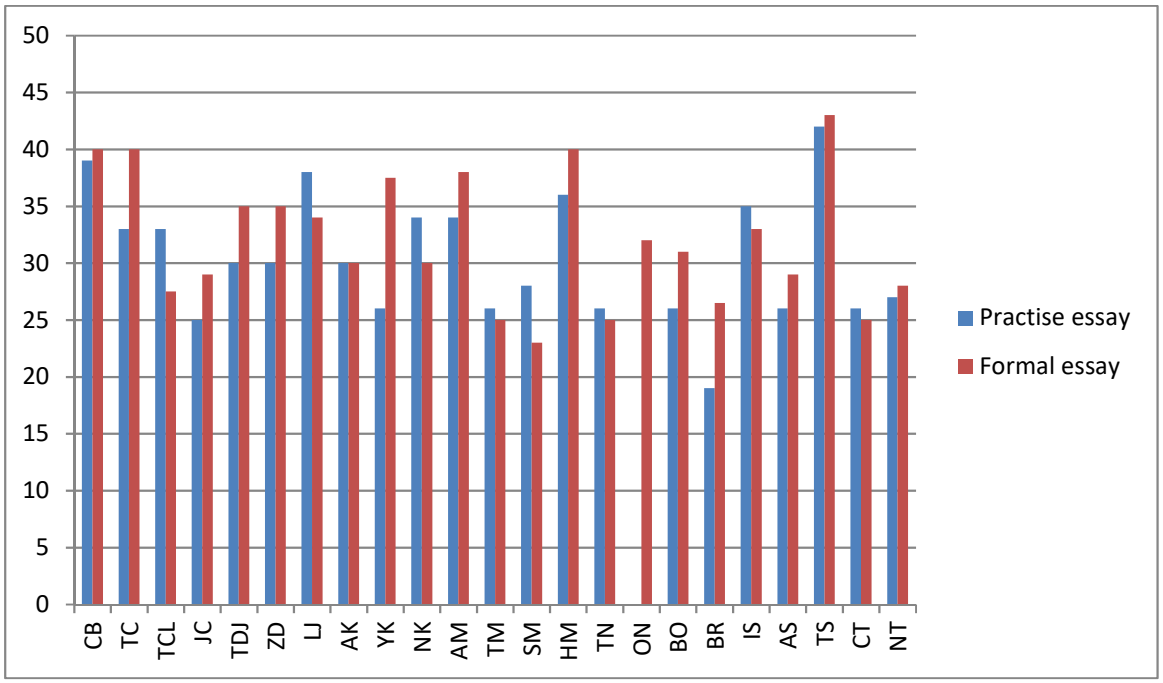


Figure 4.4.1.a. Graph depicting the results for the practise essay compared to the third formal essay. Control Group.

ON did not write the practise essay. 13 learners improved their marks from the practise essay to the formal essay. One learner achieved the same mark for both essays and 8 learners' marks dropped from the practise essay to the formal essay. The class average for the practise essay was 30.4 out of 50, 60.8% and this increased by 3.8% to 32.3 out of 50, 64.6% for the formal essay.

4.4.1.b. Experimental Group.

I gave the learners from the experimental group the same set of practise essay topics to choose from that the control group were given (Appendix P). The practise essays were written during class time on 2 August and handed in. The experimental group's essays were marked according to the GDE rubric and comments were given and recorded. The marks for the practise essay were recorded but not shared with the learners. The practise essays with comments and a colour indicating their urgency were returned to the learners on 10 August. The colour codes were given as follows: for marks from 30-59% orange, 60-69% yellow and 70%+ green.

The formal essay (Appendix Q), with different topics, was written on 11 August. This essay was marked using the GDE rubric and returned to the learners with marks and comments.

The marks given for the practise essay were compared to those given for the formal essay.

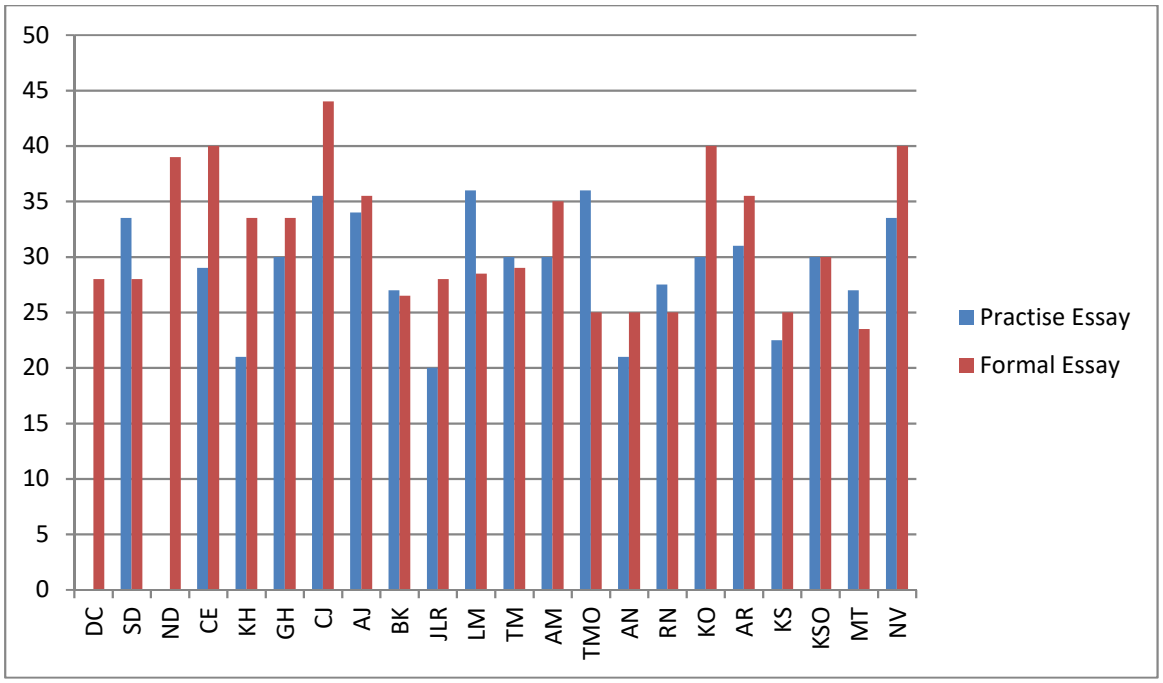


Figure 4.4.1.b. Graph depicting the results for the practise essay compared to the third formal essay. Experimental Group.

DC and ND did not write the practise essay. 12 learners improved their marks from the practise essay to the formal essay, one kept the same mark and 7 received a lower mark. The class average for the practise essay was 29.2 out of 50, 58.4% this increased by 5% to 31.7 out of 50, 63.4% a 1.2% higher increase than that shown by the control group.

4.4.1.1. The effect of marks vs. the effect of colour coding on the weakest learners

I wanted to investigate whether the colour coding method was more effective than awarding marks, especially for the weakest learners in both classes.

a. Control Group

Figure 4.4.1.1.a. Graph depicting the performance of the weakest learners from the control group.

Of the 10 learners who received less than 60% for the practise essay (which is equivalent to the "orange" code given to the experimental group) 6 improved and 4 received lower marks. The average mark for this group in the practise essay was 51% and the average mark for the formal essay was 55.8% showing an increase of 4.8%.

b. Experimental Group

Figure 4.4.1.1.b. Graph depicting the performance of the weakest learners from the experimental group.

Of the 7 learners who received "orange" comments 4 improved and 3 received lower marks. The average mark in this group for the practise essay was 23.7 out of 50, 47.4% which increased by 5.9% to 26.6 out of 50 53.3% in the formal essay. Therefore the experimental group increase by 1.1% more than the control group.

4.4.1.2. An overview of the quantitative data

In order to assess whether any significant gains had been made by either group of the course of the whole study, I have depicted the marks for the formal essay tasks for each learner in the control group and the experimental group.

a. The Control Group

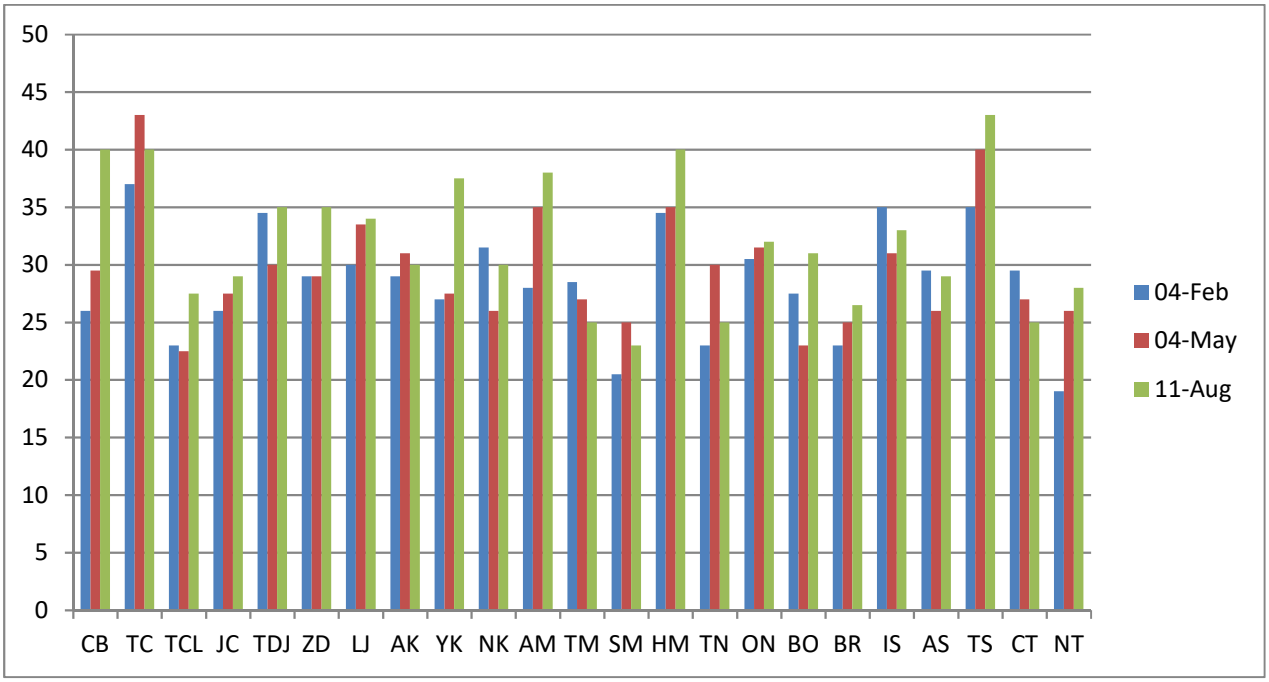


Figure 4.4.1.2.a. Graph depicting the control group's performance in all of the formal essay tasks.

The control group started with an average of 28.6 out of 50 (57.2%) for the first formal essay. This increased by 2% to 29.6 out of 50 (59.2%) in the second formal essay and by 4.8% to 32.0 (64%) in the third formal essay. The control group's average increased by 6.8% over the course of the study.

On an individual level 18 learners improved between the first essay task and the third essay task, however, only 10 (43% of the class) of those learners improved by the average 6.8% (or more) improvement of the group. The marks of 5 of the learners decreased from the first essay task to the third essay task.

b. The Experimental Group

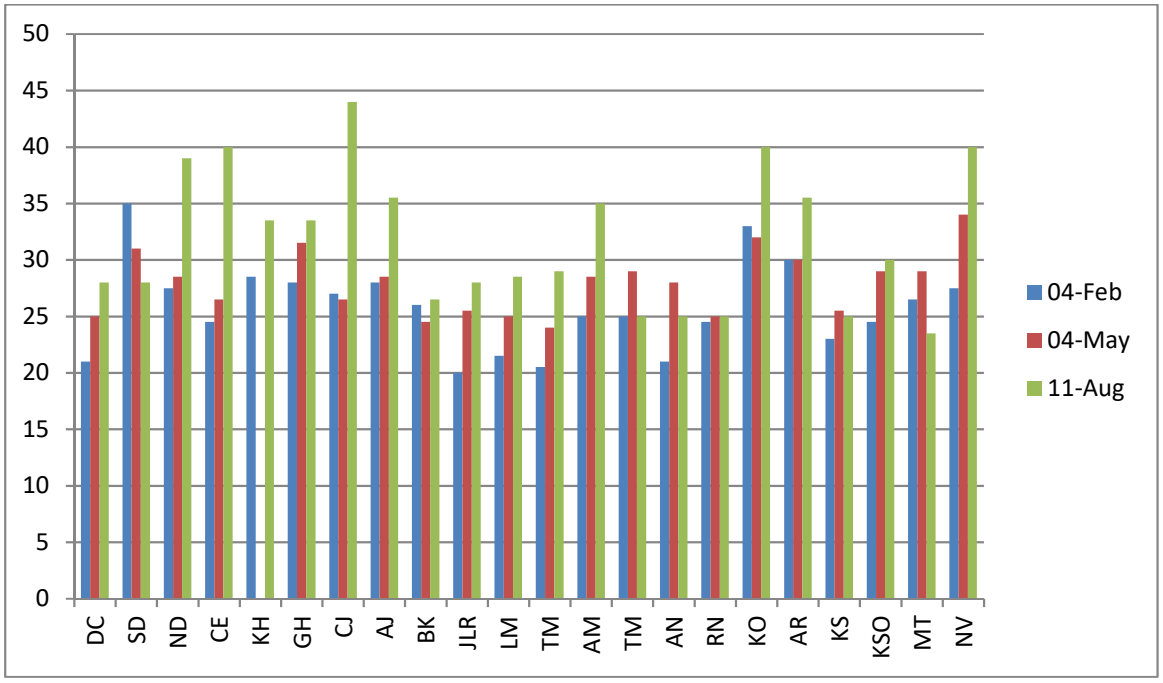


Figure 4.4.1.2.b. Graph depicting the experimental group's performance in all of the formal essay tasks.

The experimental group started with an average of 25.8 out of 50 (51.6%) for the first formal essay task. This increased by 4.2% to 27.9 out of 50 (55.8%) for the second formal essay task and by 7.6% to 31.7 out of 50 (63.4%) in the third formal essay task. The experimental group showed an 11.8% increase in their average from the first formal essay task to the third formal essay task. The control group's average increased by 6.8% from the first formal essay to the third formal essay. Overall, the experimental group increased their average mark on their essay by 5% more than the control group.

On an individual level 19 learners improved overall, with 13 (59% of the class) learners improving by the average improvement of the group (11.8%) or more. One learner's mark stayed the same and two learners' marks decreased overall.

Overall the control group's average improved by 6.8% from the first formal essay task to the third formal essay task. The experimental group's average improved by 11.8% from the first formal essay to the third formal essay, 5 % bigger improvement than the control group.

c. The overall performance of the control group compared to that of the experimental group.

	Control Group	Experimental Group	Difference
First Formal Essay	28.6 (57.2%)	25.8 (51.6%)	2.8 (5.6%)
Second Formal Essay	29.6 (59.2%)	27.9 (55.8%)	1.7 (3.4%)
Third Formal Essay	32.0 (64.0%)	31.7 (63.4%)	0.3 (0.6%)

Table 4.1. Representation of the overall performance of the control group compared to that of the experimental group.

At the beginning of the study it was clear that the control group was the group with the better ability- their average being 5.6% higher than that of the experimental group. At the end of the study it is clear that both groups improved, however, the experimental group improved at a faster rate than the control group did, bringing their averages to within 0.6% of each other.

4.4.2. The Qualitative Data

a. The control group

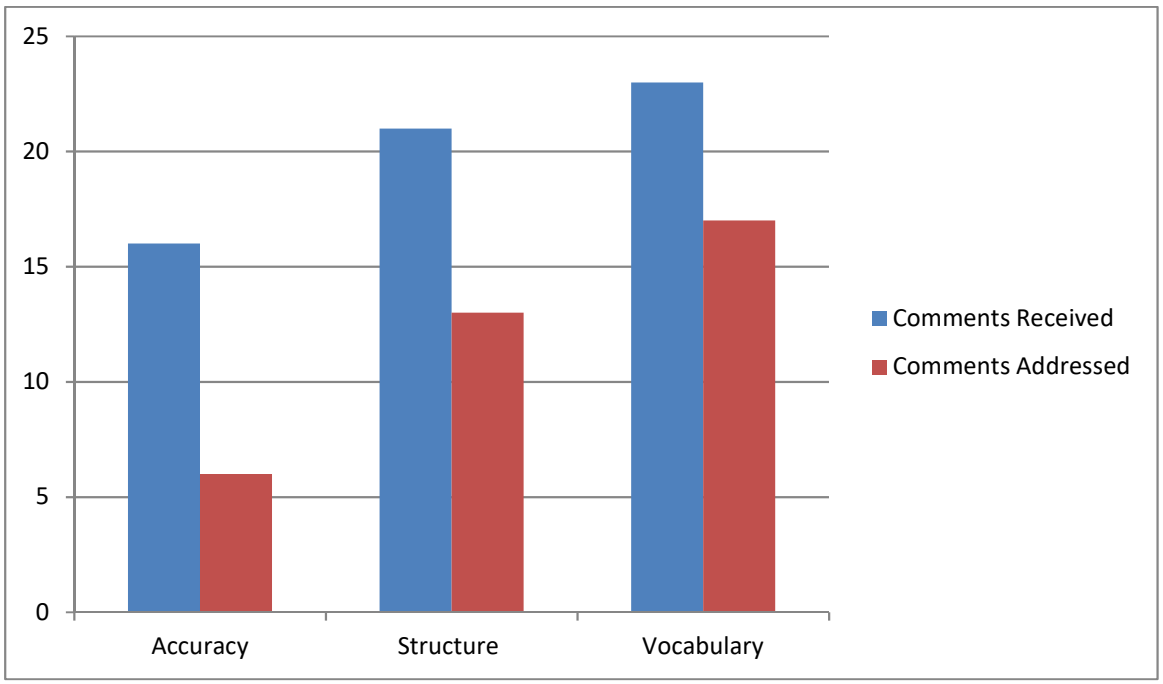


Figure 4.4.2.a. Graph depicting the control group's comments.

The control group received 60 comments in total and addressed 36 (60%) successfully. 16 comments were given for accuracy, of which 6 (37.5%) were addressed. 21 comments were given for structure of which 13 (61.9%) were addressed and 23 comments were given for vocabulary of which 17 (73.9%) were addressed successfully.

b.The experimental group

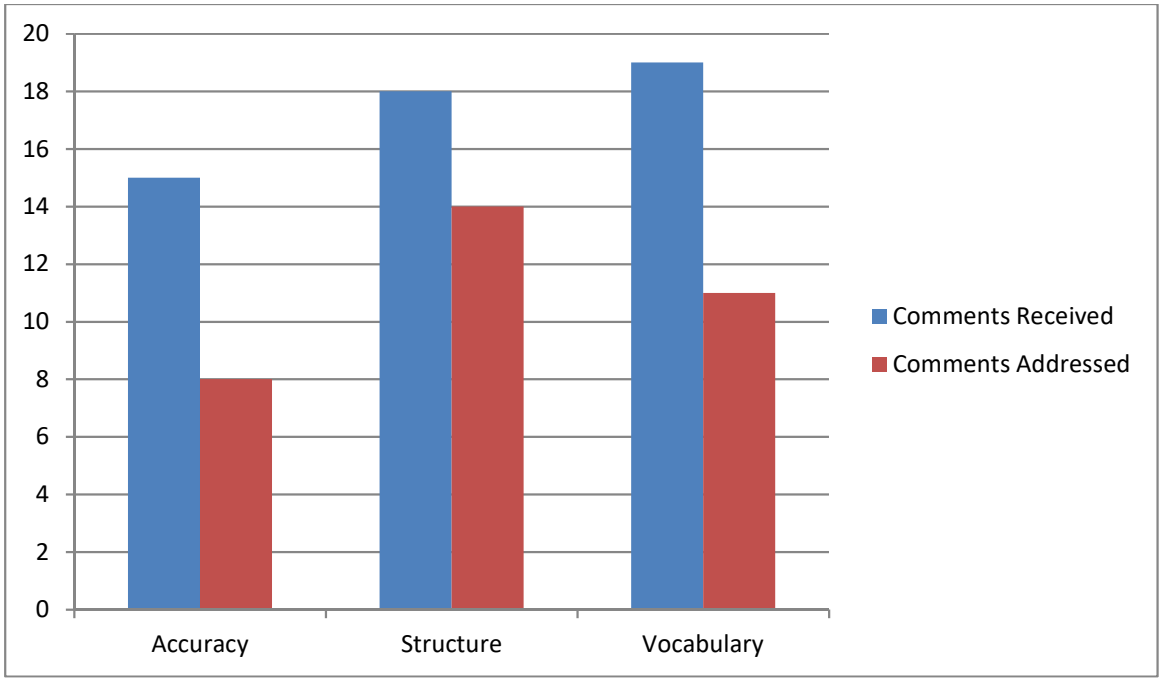


Figure 4.4.2.b. Graph depicting the experimental group's comments.

The experimental group received 52 comments in total and successfully addressed 33 (63%). Of the 15 comments given for accuracy, 8 (53%) were addressed. For structure 14 of the 18 comments (77%) of the comments were addressed and for vocabulary 11(57%) of the 19 comments were addressed.

4.5. CONCLUSION

The data collected during the pilot study and the three cycles of action research was depicted separately for the control group and experimental group so that comparisons may be easily made. The quantitative and qualitative data has also been depicted separately as different claims will arise from each.

5. CHAPTER 5: DATA ANALYSIS

This chapter used the data collected from the pilot study and each of the action research cycles to make direct comparisons between the performance of the control group and experimental group. As a result of the comparisons, certain claims are formed that can be supported by the data that has been collected.

5.1. THE ANALYSIS PROCESS

Cohen et.al (201, p. 554) warn that a researcher must be particularly cautious when analysing qualitative data for two reasons. A researcher approaches data with certain pre-conceived ideas, which are difficult to overcome. In addition to the possibility of focusing on preconceived ideas, there are new insights and reflections that occur during the research that were not necessarily present in the proposal.

a. Quantitative data

The quantitative data from all three cycles of action research is represented using bar graphs. As the study progressed different colours were used to represent the different marks that each learner received for each assessment on the graphs. The class average for each task was calculated. The averages were used to determine how much each group had improved from task to task and to measure the improvement of the control group against that of the experimental group.

b. Qualitative data

The qualitative data from the questionnaires was represented systematically using bar graphs to demonstrate the frequency of certain responses. Individual responses were then looked at more closely in the data analysis. When deciding on the categories to use for the graphs I did this in response to the kinds of answers that were given. Great care had to be taken to be responsive to what the learners said on the questionnaire, rather than what I had hoped to find.

Focus Group responses were colour coded according to two categories: motivation for withholding marks and suggestions for future tasks or interventions. This was a very lose form of Content Analysis in which a researcher tries to identify patterns of trend

in the data that support both existent and emerging theories (Cohen et.al 2011, p. 564-5)

The systematic and verifiable nature of content analysis makes it attractive for this study, as the process is transparent and therefore easy to apply in this context or reapply in another context (p. 563) where the aim is to test the same hypothesis as Butler's study proposes.

5.2. THE PILOT STUDY

The interesting thing about the results of the **pilot study** is that they actually indicated a situation that was very different from what I had been experiencing in the classroom with these two groups. It has been my observation that the learners in the so-called weak classes work out very early in the year that they have been streamed and that they are in "weak" classes.

9A, the control group, has actually got a much better overall work ethic than 9D. They settle down to work much faster, they bring their books, they are generally good with doing their homework and complete tasks in class in a much more focused manner. This class appears to be more focused and more anxious to improve and yet only two learners resubmitted work in order to get better marks.

9D, the experimental group, is an example of how learners can become disruptive in an attempt to deflect attention from school work and academic performance (Appendix S). There are only a few learners who always have their books and have always completed their homework. 9D has somewhat of a bad reputation in the school as they tend to steal and hide board markers and projector remotes and many of the learners have their own individual problems that add to a difficult group dynamic. Yet, surprisingly, fifteen of them resubmitted work for further feedback, knowing that they would not receive marks.

The results of the **pilot study** lead me to the following conclusions as a starting point for the actual study:

1. Only two learners from 9A, the control group, submitted their comprehension answers again, whereas fifteen learners from 9D, the experimental group, resubmitted. **I**

concluded from this that the absence of marks seemed to encourage learners to seek more feedback opportunities.

2. I was surprised that two of the learners in 9D got lower marks in their second attempt with, what I felt was, the benefit of the comments. **The absence of marks seemed to make learners unable to distinguish between the answers that were correct and those that were incorrect** and these two learners changed answers that were originally correct. I concluded from this that it would be too difficult to use comments without marks on tests that were marked according to a memo, like comprehension and literature tests, because it seemed to create confusion for the learners. I didn't really understand why but I did not want to take the risk of this happening in the actual study.

3. **I decided not to use comprehension or literature tests in the actual study.** Firstly, because I did not want to involve other teachers due to the difference in the consistency of the marking and the extra demands that it would make on their time to mark the learners in my classes differently. The other reason for excluding tests where marks are awarded for each answer was because they posed a structural problem in the study; it would be impossible to indicate how accurate the learner's answer is without awarding marks or giving information that would enable them to work out their marks for themselves.

CLAIM 1: When learners receive qualitative, task-oriented feedback on their work without quantitative feedback in the form of marks, they seek more opportunities for further feedback.

5.3.THE FIRST CYCLE OF ACTION RESEARCH

In the first cycle of action research, the learners were given a formal essay to write under test conditions. The control group received marks and comments for the essay and were told to re-work it and re-submit it. The experimental group's marks were kept secret and they only received comments for the formal essay. They were also told to re-work and re-submit the essay.

5.3.1. DISCUSSION OF THE QUANTITATVE DATA

a. The control group began with an average of 28.5 out of 50, or 57%. After the opportunity to re-work the essay the class average rose to 32.1 out of 50, 64%. Showing an increase of 7%. (Chapter 4. p. 49)

b. The experimental group began with an average of 25.7 out of 50, 51.4%. After the opportunity to re-work the essay the class average rose to 29.4 out of 50, 58.8%. Showing an increase of 7.4%. (Chapter 4, p. 50)

Two things are immediately clear from these results. Firstly, the control group is the better performing group with the higher average in both tasks. Secondly, the intervention appeared to have only had a slightly more positive effect on the experimental group.

However, a closer look at the effect of the intervention on each individual learner shows that the intervention was more effective on a personal level for the experimental group than for the control group. In the control group 20 of the 23 learners re-worked their essays. 15 of these learners improved their marks. Out of those 15 learners, the 3 of the 4 who improved the most (by more than 10%) were high achieving learners and one was the lowest achieving learner. This fits perfectly with Butler's (1988) statement that, "Individualised information which relates specifically to aspects of the task which have or have not been mastered should maintain initial task involvement in both more and less successful students (p. 2)". Butler goes on to say that grades maintain, "…interest and performance in high achievers who expect these to be self-enhancing, these should be undermined in low achievers (p.2)". All of the four learners whose marks stayed the same in the re-worked essay achieved an initial mark that was lower than 60%. The class average of the control group improved because of a few high achievers who used the comments to further improve their marks in expectation of further ego-enhancing feedback.

In the experimental group, who only received comments, 20 of the 22 learners re-worked their essays and all but one improved their performance even if the improvement was only slight. This result is a confirmation of Butler's hypothesis that low achievers will perform better after comments only (1988, p 3). Therefore the average improvement shown by the experimental group was because the entire group improved as individuals.

As a researcher and a teacher, I had quite a naïve expectation that the intervention of withholding marks was going to produce significant results from the beginning. What in fact happened was that the experimental group only showed a 0.4% higher improvement than the control group.

The following advice from McIntosh (2010) was very helpful in showing the way forward, "When values define the ends of a practice, such ends should not be viewed as concrete objectives or target which can be perfectly raised at some future point in time… Ends are defined in practice and not in advance of it (p. 33-4)". In order to keep the study ethical and valid I had to really control my disappointment and let the results speak for themselves.

It was always my intention to conduct a second cycle of action research and I decided to go into this cycle with the aim of investigating whether the results of the first cycle could be improved upon. As McIntosh (2010) cautions, data can provide a strong basis for taking action but the understanding of the true situation is always somewhat incomplete (p. 35). I went into the planning of the second cycle of action research hoping that a clearer picture would emerge.

CLAIM 2: Although both groups showed a similar average improvement, the intervention was more beneficial on an individual level for the learners who received comments but no marks.

5.3.2. DISCUSSION OF THE QUALITATIVE DATA

a. The control group addressed 26 of the 55 (47.2%) comments that the learners who re-submitted the essay were given. Each learner was able to address at least one comment in the re-worked essay. (Chapter 4, p. 51-52)

b. The experimental group addressed 49% of the comments that they were given as a group (28 out of 57). In this group all of the 22 learners were able to address at least one comment in the re-worked essay. (Chapter 4, p. 53-54)

The experimental group was only able to address 1.8% more of the comments that were given, which is not a significant number in itself. Furthermore, all of the learners were able to address at least one comment.

The detailed analysis of the comments given to both groups across the three achievement criterion indicate an interesting pattern: in both groups, each learner received a comment on vocabulary. Each group as a whole received 20 vocabulary comments and learners in both groups appeared to struggle to address these comments. Only 30% of the comments were addressed by the control group and 25% by the experimental group. This is in contrast to comments on accuracy and structure which both groups were able to address with more than 50% success. I don't understand this – are you saying that the control group did better than the experimental group at responding to comments? Doesn't that contradict your finding that the experimental group benefitted slightly more?

This may be attributed to what Stevenson and Palmer describe as the difference between implicit and explicit learning (1994, p.3). Explicit learning is conscious and easier to recognise and apply than implicit learning. The comments about accuracy and structure may have been easier to remember and apply because they relate more to the conscious activities of schooling. Learners are taught facts and are taught how to construct pieces of writing. Whereas vocabulary is largely an inherent and implicit product of learning, it takes time for new and unfamiliar words to be applied naturally.

I couldn't make any particular claim about the significance of this finding but I did find that it brought me closer to answering a question that had always puzzled me as a teacher: Why is it so difficult to get learners to improve their vocabulary? It is an intensive process that a

teacher needs to be mindful of at all times. One needs to model the use of good vocabulary, pick up words that learners are unfamiliar with and explain them and then use the words often with explicit reminders. In addition to all of this effort one needs to be prepared for the possibility that learners may only use these words many years later, when they feel comfortable, or not at all.

With the pressures of assessment I can often feel like I don't have time to labour over words that I feel the learners should be using. As Black and William (2006) mention, external pressure to perform causes teachers to emphasise grades and real learning gains are often ignored. Learners also become reluctant to take risks in trying out new and unfamiliar words because they want to get good marks. The irony of all of this is that the real benefits of education, like having an extensive vocabulary, tend to be ignored (p. 18).

5.4.THE SECOND CYCLE OF ACTION RESEARCH

The second cycle of action research compared the performance of the two groups on their first formal essay to their performance on a second formal essay.

5.4.1. DISCUSSION OF THE QUANTITATIVE DATA

a. The control group had an average of 28.5 out of 50, 57% for the first formal essay. Their average for the second formal essay increased to 29.9 out of 50, 59.2% showing an overall increase of 2.2%. (Chapter 4, p. 55)

b. The experimental group had an average of 25.7 out of 50, 51.4% for the first formal essay. Their average for the second formal essay increased to 27.9 out of 50, 55.8% showing an overall increase of 4.4%. (Chapter 4, p. 56)

Here the gains made by the experimental group are very apparent. As a group their improvement was double that of the control group. By the time the second cycle of action research happened, the learners in the experimental group had experienced what it was like to not to receive marks and only to receive task-oriented feedback on their work in the form of comments. They also knew that for the second formal essay they would only receive

comments again. It can be hypothesised that the learners in the experimental group were more focused on the demands of the task and their ability to improve the way in which they addressed the demands of the criteria of the task than trying to mask low ability (Butler 1988, p. 2).

A look at the individual results of the second cycle of action research appears to support the hypothesis that the absence of marks as feedback does indeed improve task-involvement. In the control group 14 learners improved their marks and 16 learners improved in the experimental group.

What is interesting, however, is the learners whose marks dropped in both groups; in the control group 8 learners achieved lower marks in the second formal essay. In the experimental group 4 learners received lower marks in the second formal essay. Double the number of learners received lower marks form the second formal essay in the control group than in the experimental group. The average by which the marks decreased in the control group was 3.2% which is 1.4% more than the 1.8% by which the learners in the experimental group dropped.

According to Nicol and Macfarlane-Dick (2006) learners use external feedback to support their own internal feedback processes (p. 202). If the external feedback confirms the poor view that a learner already has about their performance, the result often is what Brookhart (2008) calls "academic fatalism" (p. 21) where learners perceive poor performance as a result of low intelligence and do not believe that any strategies can be implemented to change this.

5.4.2. DISCUSSION OF THE QUALITATIVE DATA

5.4.2.1.The data from the questionnaire

Qualitative data was collected by means of a questionnaire and an observation of how the comments were addressed in the second formal essay.

i. In response to Question 1, "**Do you find it easy to remember the comments that you have received on your work so far?",** 14 learners from each group responded that they did.

ii. The response to Question 2, "**Write down as many of these comments as you can remember in the space below**, was able to expand on the data collected in Question 1.

13 learners in the control group were able to remember their comments, but only 2 could remember all three of the comments they were given. In the experimental group 20 learners were able to remember at least one comment and 3 remembered all of the comments they were given. Therefore, again, on an individual level, comments seem to be remembered better when marks are not present.

CLAIM 3: Learners explicitly remember comments better when marks are not present.

iii. Question 3 asked, **"Do you feel that any of these comments have helped you to improve your work?"**
There was a similar response from the control group and the experimental group to this question. 20 in the control group responded "yes" as did 19 in the experimental group.

iv. As with Question 1 and 2, the responses to Question 4 gave a clearer indication as to the real effect of the comments on each of the two groups. The question asked the learners to explain how the comments had helped them.

In the control group, of the 20 learners that replied that the comment had helped them, only 8 (35%) were able to identify exactly what they had done to change their work as a result of the comment. In the experimental group 10 learners (45%) could state specifically what they had done to change the quality of their work as a result of receiving the comments.

As the difference between the groups is essentially only that of two learners, I don't feel that it is enough to make the claim that comments without marks help learners to consciously apply specific changes to their work any better than comments with marks do.

5.4.2.2.Longitudinal depiction of comments received, remembered and applied.

a. The control group addressed 47% of the comments in the re-worked essay, 52% in the second formal essay and explicitly remembered only 35% of the comments. (Chapter 4, p. 61)
b. The experimental group addressed 49% of the comments in the re-worked essay, 50% in the second formal essay and explicitly remembered 50%. (Chapter 4, p. 66)

As stated earlier, the differences between the ability of the control group and the experimental group to address the comments are not substantial.

What is interesting to note is that the learners in the control group addressed 17% of their comments without explicitly remembering them. A possible explanation for this could come from Stevenson and Palmer's description of implicit learning. They state that explicit learning becomes implicit learning when it is practised enough (1994, p.3).

This appears confusing when we do not consider the qualitative data together with the quantitative data. It appears that the control group have practised addressing the comments more, in previous assessments, than the experimental group has. The control group also started off as the slightly stronger group with the better work ethic. They would have been together as a group from grade 8 and perhaps been more conscious of the requirements of essay writing and this is what has made the explicit learning from the comments become implicit, unconscious learning.

The experimental group has also been together as a class since grade 8 and as a group they could have been experiencing the "academic fatalism" that Brookhart (2008, p. 21) describes. The absence of marks in this study could be the reason that they are focusing on, remembering and practising the requirements of essay writing.

The fact is that the comments all highlight various aspects of essay writing that all of the learners have been familiar with since grade 8. The learners in both the control group and the experimental group would recognise the content of the comments that they had been receiving. What has happened with the experimental group could be explained by Stevenson and Palmer's explanation of learning as understanding:

"Learning through understanding involves deliberate attempts to make sense of new material by using prior knowledge, and deliberate attempts to rethink one's ideas in light of the new material."

"Explicit understanding, then, goes beyond the simple addition of new material to long-term memory… It involves the deliberate use of pre-existing knowledge to interpret new material to modify and update pre-existing beliefs and ideas (1994, p. 8-9)."

When the experimental group were forced to notice the comments and work with them as the only form of feedback on their performance, they had the opportunity to deliberately modify their approach to the task of essay writing. The comments were the only way in which they could evaluate their work and because this feedback was task-oriented, they modified their approach to the task and were not distracted by possible negative emotions about their performance.

When the control group sees the comments that they receive with marks, it could be that they recognise those comments as reminders of material that they have seen before. They recognise the information but because the mark, as ego-involving information, is more salient, their conscious attention is diverted away from the comment. This is possibly explained by what Brookhart terms "academic fatalism" (2008, p. 21). Where learners believe that the mark that they achieve is the result of their fixed intelligence, they do not perceive feedback as a helpful tool but recognise it as evidence of their inability to perform well.

Learners in the control group are possibly following the process of transforming explicit information into implicit information through the sheer number of times that they are seeing information relating to essay writing in the form of revision lessons and comments. However, because the presence of marks prevents their interpretation of the comments as helpful rather than judgemental, they do not engage with them on an explicit and conscious level. This can been seen both by the fact that, as a group and as individuals, they do not show the same overall improvement in marks as the experimental group and that they do not explicitly remember or focus on the comments when they approach a new assessment.

The learners in the experimental group had the advantage of not being distracted by the ego-involving feedback of marks. The only feedback that they received was task-oriented in the form of comments. In this case the feedback made the learners aware of what they needed to

do, as a conscious effort, in order to achieve the goals of the assessment (Brookhart 2008, p. 21). Therefore, the learners in the experimental group were aware that the comments needed to be remembered explicitly and they put real conscious attention into addressing them as much as possible.

There is further evidence of this claim from the focus group held with the learners from the experimental group.

"I believe that if there's just comments, you can see more of what you could have done"

"Okay, so when I go for my next task, I think about the comments and not about the mark. Like I'm not stressing because I failed the last one. I'm concentrating on the comments so I can improve."

"I think it's a good idea, because normally when you get the test back you always look at the mark and you forget to…like you read the comment but it doesn't really like go through your mind. Like what goes through your mind is the mark, so it's better to get a comment rather than a mark, so you know how to improve in the future."

In a focus group interview with the learners of the experimental group many of them could verbalise the internal process that they employed when they set out to write the next essay.

5.5. THE THIRD CYCLE OF ACTION RESEARCH

In the third cycle of action research, the control group and the experimental group were both given practise essay topics. The control group received marks and comments for these topics whilst the experimental group only received comments and a colour code that represented the urgency with which the comments needed to be applied and I recorded their marks in secret. Both groups then wrote the third formal essay and the marks for the practise essay and formal essay were compared. The two groups were also compared in terms of the comments that they were able to address.

5.5.1. DISCUSSION OF THE QUANTITATIVE DATA

a. The control group's average for the practise essay was 30.4 out of 50, (60.8%). For the formal essay, this increased by 3.8% to 32.3 out of 50, (64.6%). (Chapter 4, p. 71-72)

b. The experimental group's average for the practise essay was 29.2 out of 50, (58.4%). For the formal essay, this increased by 5% to 31.7 out of 50, (63.4%) - a 1.2% higher increase than that shown by the control group. (Chapter 4, p. 72-73)

The experimental group's average did increase by 1.2% more than that of the control group. However, this in itself is not a very significant margin. When exploring the individual results of both groups there seems to be very little difference there too, 13 out of 23 learners improved in the control group and 12 out of 19 in the experimental group. If this is looked at as a percentage 56% of the control group improved as opposed to 63% of the experimental group, however the actual number of learner who improved is too small to be significant.

5.5.1.1. The effect of marks vs. the effect of colour coding on the weakest learners

I wondered whether the introduction of the colour coding for the experimental group could account for the similar increase in averages between the two groups. I compared the learners in both the control group and the experimental group who received less than 60% or an "orange" code for their practise essay tasks.

a. In the control group, 10 learners received less than 60% for the practise essay. The average mark for the practise essay in the group was 51%. In the formal essay, 6 learners improved and 4 got lower marks, however, as a group, these learners improved by 4.8% to 55.8%. (Chapter 4, p. 73)

b. In the experimental group, 8 learners received the "orange" code representing a mark below 60%. The average mark in the practise essay was 47% which increased by 5.9% to 52.9% in the formal essay, a 1.1% bigger increase than that of the control group. (Chapter 4, p. 74)

This increase is very similar to that of the entire group so it was not the weakest learners who were particularly affected by the colour coding as such. What is important to remember is that now the learners from the experimental group were receiving a form of ego related feedback. Therefore, in this instance they could have been experiencing the same feelings of academic fatalism, as the weak learners from the control group seem to be experiencing.

Looking at Figure 4.1.b. it is interesting to note that, of the 7 learners whose marks decreased between the practise essay and the formal essay, the two learners who showed the most significant decrease were LM and TMO who both received a "green" code initially. The possibility here is that the inverse of what Stough and Emmer (1998) describe when they say that learners can have such a negative emotional reaction to a low mark that they ignore further feedback (p. 350). It appears here that LM and TMO had such a positive reaction to the ego-involving feedback that they ignored the task-involving feedback for the formal essay. This confirms Brookhart's (2008) claim that even praise can be harmful in certain instances because it is an external regulator of motivation and interest as opposed to the learner being interested in the task itself (p. 50).

CLAIM 5: Any form of ego-involving feedback can detract from task-involving feedback.

5.5.1.2. OVERVIEW OF THE QUANTITATIVE DATA

The third cycle of action research did not seem to produce significant gains in itself but when the two groups are compared longitudinally, the differences between them are clearer.

a. The control group started with an average of 28.6 out of 50 (57.2%) for the first formal essay. This increased by 2% to 29.6 out of 50 (59.2%) in the second formal essay and by 4.8% to 32.0 (64%) in the third formal essay. The control group's average increased by 6.8% over the course of the study. (Chapter 4, p. 75)

b. The experimental group started with an average of 25.8 out of 50 (51.6%) for the first formal essay task. This increased by 4.2% to 27.9 out of 50 (55.8%) for the second formal essay task and by 7.6% to 31.7 out of 50 (63.4%) in the third formal essay task. The experimental group showed an 11.8% increase in their average from the first formal essay task to the third formal essay task. (Chapter 4, p. 76-77)

Both the control group and the experimental group showed an improvement over the course of the study. What is noteworthy is that the experimental group improved far more with each task than the control group did and far more overall. Another significant thing is that in the beginning of the study the control group was clearly the stronger group with an average mark that was 5.6% higher than that of the experimental group. By the end of the study, the achievement gap between the two groups had closed to 0.6% making the experimental group almost on par with the control group.

On an individual level, the gains that have been made by the experimental group are also significant. 13 out of 22 learners (59% of the class) improved by 11.8% or more reflecting that the majority of the group was responsible for the overall average improvement. In the control group only 10 out of 23 (43% of the class) learners improved by the overall average, which indicates that a smaller percentage of learners were responsible for the increased average of the class as a whole.

CLAIM 6: Marks and comments do help some weak learners to improve, but comments without marks help more weak learners to improve at a faster rate.

5.4.2. DISCUSSION OF THE QUALITATIVE DATA

a. The control group received a total of 60 comments for the practise essay and addressed 36 (60%) successfully in the formal essay. 16 comments were given for accuracy, of which 6 (37.5%) were addressed. 21 comments were given for structure of which 13 (61.9%) were addressed and 23 comments were given for vocabulary of which 17 (73.9%) were addressed successfully. (Chapter 4, p. 78)

b. The experimental group received 52 comments in total for the practise essay and successfully addressed 33 (63%) in the formal essay. Of the 15 comments given for accuracy, 8 (53%) were addressed. For structure 14 of the 18 comments (77%) of the comments were addressed and for vocabulary 11(57%) of the 19 comments were addressed. (Chapter 4, p. 79)

Overall, the experimental group only addressed 3% more of the comments than the control group did. Yet what is interesting to note, is that the experimental group addressed more of the comments for accuracy and structure than the control group, while the control group addressed more vocabulary comments than the experimental group did.

This further strengthens my claim that the experimental group put more conscious effort into addressing the comments because improving the accuracy and structure of one's work are tasks that involve the application of explicit learning. This requires the learner to explicitly remember and deliberately apply the information in the comments (Stevenson and Palmer 1994, p. 8).

Vocabulary and spelling are implicitly learned constructs that can be the result of explicit learning. The ability of the control group to address vocabulary comments better than the experimental group could be because, as the stronger group, they are simply being reminded of information that they already know.

5.4. CONCLUSION

In essence, my action research study has confirmed Butler's hypothesis in a real educational setting. The problem is that the scope of the study and its results are not statistically significant. The study would have to be conducted again on a much larger scale to provide significant quantitative data.

The importance of the study as it stands is lies in its implications for my personal educational practise and perhaps for educational practise as a whole. Both the experimental group and control group showed an improvement due to receiving comments, however, the experimental group improved at a faster rate because their marks were absent. The experimental group also remembered and addressed their comments more effectively than the control group did. Finally, it was clear that the learners in the experimental group were uneasy about not receiving marks or any form of evaluative feedback, yet when the colour coding was introduced; it had the same effect as any evaluative and ego-involving feedback. To me this indicates that there is no substitute or alternative to providing marks, it would appear that learners are better off without them even if this does make them uncomfortable.

6. CHAPTER 6: CONCLUSION

The aim of this study was to test Butler's 1989 hypothesis that intrinsic motivation will be differentially affected by task-involving and ego-involving evaluation (p.1) in a real classroom setting with high-stakes assessments. The practices of Butler's original study, in which she had a control group to whom she gave marks and comments and an experimental group to whom she gave comments only, were replicated in an ecologically valid setting. As the researcher, what I wanted to determine was whether weak learners who received comments only would increase their engagement in the demands of the task and thereby improve their marks.

I believe that the use of action research as a paradigm for performing this study was one of its most valuable aspects. This does, however, have some limitations as the study was performed on too small a scale and the data that was captured did not reflect gains substantial enough to be statistically significant. However, the nature of action research is such that it is an intimate paradigm designed for the classroom, the teacher and the learners. It serves primarily to inform mindful practice and promote behaviours that enhance the learning process. It is not always possible to derive meaningful statistics from such a situation but it is possible to change the lives of individual learners, which is in essence what teaching is about.

This study in its present form will not be the basis of a reform in the way that all teachers practise assessment; some may try it to enhance their own practice but the chances of withholding marks from learners becoming part of a high-stakes system are very slim. The true significance of the study is for my own understanding and practice as a teacher and for the learners that I have taught this year and will teach in the years to come. One day perhaps it will be adopted by the English department at the school and by other subjects that it can be adapted to.

6.1. An evaluation of the claims made as a result of the study

I made six claims based on the findings during the study:

1. When learners received qualitative, task-oriented feedback on their work without quantitative feedback in the form of marks, they sought more opportunities for further feedback.
2. Although both groups showed a similar average improvement, the intervention was more beneficial on an individual level for the learners who received comments but no marks.
3. Learners explicitly remembered comments better when marks were not present.
4. Whilst learners who received comments with marks could recognise and apply the information that they received in comments, they did not apply it with the same conscious and explicit effort that learners who do not received marks could.
5. Any form ego-involving feedback detracted from task-involving feedback.
6. Marks and comments did help some weak learners to improve, but comments without marks helped more weak learners to improve at a faster rate.

The claims are of a predominantly qualitative nature and are based on observation. Of course these observations can be depicted on a graph but, as said before, the study did not yield data that is of statistical significance. The significance of the claims in the study is that they are significant on an individual level for each of the learners in the experimental group.

Marks are essentially an evaluation of work that has been presented. To evaluate is to form an idea of a numerical representation of that work and it contains an element of finality as well as judgement. This is an interesting aspect of assessment and can be seen very clearly in the third cycle of action research where the experimental group was given a colour code with their comment.

The learners who were most negatively affected by the colour coding were the top and bottom achievers for that task and my suspicion is that their reaction was a response to their perception of the colour code as a final evaluation. The orange code confirmed suspicions of poor performance and the green code confirmed good performance. This appeared to produce the effect of considering the next essay task to be completely separate from the practice essay. Some of the lowest achievers may have felt that the effort of addressing comments and improving their engagement with the task was futile because their performance was already

poor in their minds. Two of the top achieving learners were so satisfied with their performance that they did not see the need to continue their engagement with the task as they had already achieved well. Thus, any form of ego-involving feedback detracts from task-involving feedback.

Closely linked to how ego-involving feedback detracted from task-involving feedback is the observation that the absence of evaluative feedback promoted interest in task-involving feedback. Learners actively sought more feedback opportunities and consciously remembered the feedback that they received. Therefore, I count claims 5, 1 and 3 as the most significant and educationally valuable conclusions.

What surprised me most about this study was that during the course of the data collection process it was easy to see that both groups had improved. At first I perceived this as an indication that the data would show the opposite of what Butler's study had. I wondered if perhaps my comments were not well formulated or if perhaps the experimental group was in fact on a very much lower functioning level than the control group. It was only when the study was complete and the information was clear on graphs that the significance of the improvement for the individual learners could be seen as mentioned in claim 2.

I found that during the study the experimental group was very interested in their progress and mine! They were fascinated by the thought of a teacher trying to change the age-old assumption about marks, and once they had overcome their initial concerns about being judged by outsiders, they were keen to participate and advise in the process.

What was particularly striking was the enthusiasm they showed in the third cycle of action research when a suggestion from one of their classmates (about the colour coding) was implemented. They loved being consulted and included in the assessment process and it felt like they had a sense of ownership when it came to their performance. The pinnacle of this process came when I put up the graph that tracked their performance in each of the formal assessments. When the learners could see their own progress and account for it, it was a very exciting moment for me and for them – it was about a teacher and her learners collaborating to do assessment well.

6.2. How has the study answered the original research question and empirical sub questions?

The original research question was:

Will the practice of withholding the marks of at risk Grade 9 learners and providing detailed comments encourage them to make an effort to improve the quality of their work, apply the guidelines suggested by the teacher, and thus improve their marks?

With regards to this question my research found that comments, whether accompanied by marks or not, do encourage learners to apply the guidelines suggested by the teacher and thus improve their marks. However, the absence of marks helps more of the weaker learners to apply comments as individuals (claim 2), they remember comments more explicitly when marks are not present (claim 3) and apply them with more conscious effort (claim 4) and thus improve at a much faster rate (claim 6).

Empirical sub-questions:

1. Does withholding marks induce learners to take more interest in and remember a remedial comment?

2. Will learners be encouraged to apply the guidelines of that comment to future assessment tasks and thereby improve their marks?

The **pilot study (Chapter 3 Section 1)** demonstrated very clearly that when marks are absent, learners do seek further opportunities for feedback on the task (claim 1). The absence of marks appeared to stimulate their curiosity about their performance on the task, whereas learners who received marks for the comprehension task were largely uninterested in further feedback of any kind.

The results of the questionnaire in the **second cycle of action research (Chapter 3 Section 3.2)** demonstrated that the learners from the experimental group, who had received comments only, explicitly remembered those comments three months later far better than the learners in the control group did (claim 3).

The learners in both the control group and the experimental group applied the guidelines in the comments and thus improved their marks. However, when marks were absent, the

learners of the experimental group addressed the comments with more conscious effort (claim 4) and improved at a faster rate than the control group did (claim 6).

6.3.The implications of the study for future practice

What also stood out for me as a teacher was how difficult it was for learners to address comments about vocabulary. Ask any learner and they will tell you that they need to improve their vocabulary. Ask any teacher how to improve vocabulary and they will tell you to read. This is not enough support for very weak learners. They simply cannot overcome the deficiencies of poor primary school education or lack of resources on their own, while teachers do not have the time to help each of them individually.

I was shocked to discover that many of the learners that I teach really struggle to read on a level appropriate for grade 9. They mis-read words, make up their own words to replace unfamiliar ones and often just skip entire chunks of text. No person can improve their vocabulary from reading in this way. The reading needs to be tackled first.

This insight led me to initiate an extra-mural project with the students. The project was inspired by my own experiences as a learner. While I was at school, I enjoyed reading and perceived myself as a good reader. I offered to read aloud to my grandfather, who was blind, and to my horror, I discovered that I really struggled. It was extremely difficult to read aloud for sustained periods of time, recognise words immediately and keep pronouncing them correctly. This was for a first language English speaker! However, at the end of reading *King Solomon's Mines* and *Alan Quatermain* aloud, I had become an excellent reader and this confidence filtered into my comprehension and writing skills.

I wondered whether a similar project might benefit some of my weakest learners. The school works closely with a retirement village for community service projects and I enquired whether some of the residents would be interested in volunteering to be read to. The learners were each randomly assigned a resident and then each pair chose a book from a selection that I had provided. The learners and residents read together for an hour a week and already after only five weeks there is a marked improvement in the performance of these weaker learners in English assessments.

The project has taken on a life of its own with little guidance from me: Learners record new words in notebooks, the residents draw pictures to illustrate concepts and many of them bring their own dictionaries. One learner even remarked to me that she had never realised that

reading was “like having TV in your head”. I am very excited about this project because even after only five weeks these weaker learners are performing better in tests and written assignments. One learner in my class is passing English for the first time this year.

Another significant realisation that I had because of this study was about the importance of writing as a discipline in itself and how important it is for learners to master this discipline. It was my observation that learners, especially the presumably academically weak ones who participated in my study, really struggle to translate their experiences into words on paper.

Through the school’s community service project at a local library, I have made contact with a local author who has actually written a book about a young girl who grows up in the same street where the school is located. He has agreed to come to the school and give workshops on how to translate one’s experiences into words and other tips for creative and transactional writing.

6.4. Conclusion

This study has been a very interesting and enriching experience that has led to a number of realisations and changes in the way that I do things. I am excited about the spin-off projects that it has inspired too. At the beginning of the study, I was somewhat sceptical of action research as a paradigm and I wondered if it could yield workable results. Now I see that it is the personal significance of such a study and the way that it inspires change in one’s teaching practice that is its greatest significance.

7. LIST OF REFERENCES

Black, P. (2003). The nature and value of formative assessment for learning. *Improving Schools*. 6.3. 7-22.

Black, P. (2015). Formative assessment – an optimistic but incomplete vision. *Assessment in Education: Principles, Policy and Practice*. 22.1. 161-177.

Black, P. and William, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice*. 5.1.7-74

Black, P. & Wiliam, D. (2006) Developing a theory of formative assessment. In J. Gardner (Ed.), Assessment and learning (pp. 81-100). London: Sage.

Bloch, G. (2009). *The Toxic Mix. What's wrong with South Africa's schools and how to fix it.* Tafelberg. Cape Town.

Brookhart, S. (2004). Classroom Assessment: Tensions and Intersections in Theory and Practice. *Teachers College Record*. 106.3. 429-458.

Brookhart, S. (2008). *How to Give Effective Feedback to Your Students.* Association for Supervision and Curriculum Development Alexandria, Virginia USA

Burke Johnson, R. and Onwuegbuzie, A. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*. 33.7. 14-26.

Butler, R. (1988). Enhancing and Undermining Intrinsic Motivation: The Effect of Task-Involving and Ego-Involving Evaluation on Interest and Performance. *Br. J. educ. Psych*. 58. 1-14.

Butler, D and Winne, P. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*. 65.3. 245-281.

Cheng, L. and Sun, Y. (2014). Teachers' grading practices: meaning and values assigned. *Assessment in Education: Principles, Policy and Practice*. 21:3. 326-343.

Cohen, L. Manion, L. and Morrison, K. (2011) *Research Methods in Education. Seventh Edition*. Routledge. London and New York.

Conley, L. de Beer, J. Dunbar-Krige, h. du Plessis, E. Gravett, S. Lomofsky, L. Merkel, V. November, I. Osman, R. Petersen, N. Robinson, D. and van der Merwe, M. (2010). *Becoming a Teacher*. Pearson. Cape Town.

Department of Basic Education. (2013). *National Policy Pertaining to the Program and Promotion Requirements of the National Curriculum Statement Grades R-12*.

Department of Basic Education (2011). *Curriculum and Assessment Policy Statement English Home Language Grades 7-9.*

Elwood, J. & Murphy, P. (2015). Assessment systems as cultural scripts: a sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy and Practice.* 22.2. 182-192.

Emmer, E.T. and Stough, L.M. (1998). Teachers' emotions and test feedback. *Qualitative Studies in Education*. 11.2. 341-361.

Gipps, C. (1999). Socio-Cultural Aspects of Assessment. *Review of Research in Education*. 24. 355-392.

Govender, B and Wood, L. (2013). 'You learn from going through the process': The perceptions of South African school leaders about action research." *Action Research*. 11.2. 176-193.

Graue, E and Johnson, E. (2011) Reclaiming Assessment Through Accountability That Is "Just Right". *Teachers College Record*. 113.8. 1827-1862.

Gumbiner, J. (2003). *Adolescent Assessment*. Wiley and Sons Inc. New Jersey.

Hargreaves, E. (2013). Enquiring into children's experiences of teacher feedback: reconceptualising Assessment for Learning. Oxford Review of Education. 39.2.229-246.

Hugo, W. (2013) *Cracking the Code to Educational Analysis*. Pearson

Jiang, H. (2015). Dual Process of Assessment: The field experience of pre-service teachers in a rural school in China. In Garbett, D and Ovens A. (Eds). *Teaching for Tomorrow Today*. 176-184. ISATT, Edify Ltd, Auckland , New Zealand.

Koshy, V. (2008). *Action Research for Improving Practice. A Practical Guide*. Sage Publications. London.

Leedy, P. and Ormond, J. (2013). Practical Research. Planning and Design. Pearson. New Jersey.

Madaus, G. and O'Dwyer, L. (1999). A Short History Of Performance Assessment Lessons Learned. *Phi Delta Kappan*. 688-695.

Macfarlane-Dick, D. and Nicol, D. (2006). Formative Assessment and self-regulated learning: a model and seven principles of good feedback practice. Studies in Higher Education. 31.2. 199-218.

McIntosh, P. (2010). *Action Research and Reflective Practice*. Creative and visual methods to facilitate reflection and learning. Routledge. New York and London.

Palmer, J. and Stevenson, R. (1994). *Learning: Principles, Processes and Practices*. Biddles Ltd. Guildford

Rubie-Davies, C. Stephens, J. and Watson, P. (Eds) (2015). *The Routledge International Handbook of Social Psychology of the Classroom.* Routledge. London

Scott, D. (2008). *Critical Essays on Major Curriculum Theorists*. Routledge. London and New York.

Shepard, L. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*. 29.7. 4-14.

Steinberg, C. (2008). Assessment as an "emotional practice". *English Teaching: Practice and Critique*.7.3. 42-64.

Steinberg, C. (2013). PhD dissertation: Teacher Emotions towards Assessment: What can be learnt from taking the emotions seriously? *Chapter 5: Teachers' Emotions towards their Assessment Practise.*

Swafield, S. (2011). Getting to the heart of authentic Assessment for Learning. Assessment in Education: Principles, Policy and Practice. 22.1. 161-177.

Weaver, M. (2006). Do students value feedback? Student's perceptions of tutors' written responses. Assessment and Evaluation in Higher Education. 31.3. 379-394.

GAUTENG PROVINCE
Department: Education
REPUBLIC OF SOUTH AFRICA

APPENDIX A

DEPARTMENT OF EDUCATION RUBRIC FOR ESSAYS

Essay Marking Rubric (50) Suggested

	Code 7: 80-100%	**Code 6: 70-79%**	**Code 5: 60-69%**	**Code 4: 50-59%**	**Code 3: 40-49%**	**Code 2: 30-39%**	**Code 1: 0-29%**
CONTENT and PLANNING 30 MARKS	**24-30** -Vivid, exceptional detail. -Content highly original, sincere. -Ideas thought-provoking, mature. -Planning and/or drafting has produced a presentable essay. -Coherent development of topic.	**21-23½** -Content original, sincere. -Ideas imaginative, interesting. - Planning and/or drafting has produced a well-crafted and presentable essay. Logical development of details. Coherent.	**18-20½** -Content sound, reasonably coherent. -Ideas interesting, convincing. - Planning and/or drafting has produced a presentable and good essay. -Several relevant details developed.	**15-17½** -Content appropriate, coherent. -Ideas interesting, adequately original. - Planning and/or drafting has produced a satisfactory, presentable essay. -Some points, necessary details developed.	**12-14½** -Content mediocre, ordinary. Gaps in coherence. -Ideas mostly relevant. Limited originality - Planning and/or drafting has produced a moderately presentable and coherent essay. - Most necessary points evident.	**9-11½** -Content not always clear, lacks coherence. -Few ideas, often repetitive. - Content in poor taste. -Inadequate for home language level despite planning/drafting. Essay not well presented. - Sometimes off topic but general line of thought can be followed.	**0-8½** -Content largely irrelevant. No coherence. - Content filled with vulgarity and/or profanity. -Ideas tedious, repetitive. -Inadequate planning/drafting. Poorly presented essay. - Plagiarism -Off topic.
LANGUAGE and STYLE 15 MARKS	**12-15** -Critical awareness of impact of language. -Language, punctuation effectively used. -Uses highly appropriate language. -Choice of words exceptional, mature. -Style, tone, register highly suited to topic. -Virtually error-free.	**10½-11½** -Critical awareness of impact of language. -Language, punctuation correct; able to use language effectively. -Choice of words varied and creative. -Style, tone, register appropriately suited to topic. -Largely error-free.	**9-10** -Critical awareness of language evident. -Language and punctuation mostly correct. -Choice of words suited to text. -Style, tone, register suited to topic. -Mostly error-free.	**7½-8½** -Some awareness of impact of language. -Language simplistic, punctuation adequate. -Choice of words adequate. -Style, tone, register generally consistent with topic requirements. -Still contains a few errors.	**6-7** -Limited critical language awareness. -Language mediocre, punctuation often inaccurately used. - Flippant tone. -Choice of words basic. -Style, register lacking in coherence. -Contains several errors.	**4½-5½** -Language and punctuation flawed. -Choice of words limited. -Style, tone, register inappropriate. -Error-ridden.	**0-4** -Language and punctuation seriously flawed. -Choice of words inappropriate. -Style, tone, register flawed in all aspects. -Error-ridden and confused.

English Home Language CAPS

CONTENT and PLANNING 30 MARKS		**Total:**
LANGUAGE and STYLE 15 MARKS		
STRUCTURE and EDITING		

APPENDIX B

INFORMATION SHEET LEARNERS- CONTROL GROUP

DATE: 11 January 2016

Dear Learner

In addition to being your English teacher I am also a Masters student in the School of Education at the University of the Witwatersrand.

I am doing research on **The Effects of Withholding Marks and Providing Written Feedback for Grade 9 in English Home Language**.

My investigation involves developing comments that can be given to learners that will help them to understand how to improve their understanding of the work for future assessments. I will be investigating the way in which comments can be structured as well as the conditions under which they can be received. I am hoping that you will learn more about how to improve your work from the comments and be able to apply these skills for the rest of your schooling.

I am asking you to join my research study for term 1 in 2016. The aim of the study is to determine the best ways in which comments on assessments can be structured and administered in order to have the maximum effect on learning at school.

If you agree to participate in this research study, there will be no additional work required from you nor any disadvantage as the teaching and assessment program will be the same as normal. The benefit for you in participating is that for the duration of the study you will be receiving very carefully constructed feedback that should help you to improve your skills in English.

This study is not a test of your ability, nor will you gain any marks for participating. It is voluntary, which means that you don't have to do it. Also, if you decide halfway through that you prefer to stop, this is completely your choice and will not affect you negatively in any way.

I will not be using your own name but I will make one up so no one can identify you. All information about you will be kept confidential in all my writing about the study. Also, all collected information will be stored safely and destroyed between 3-5 years after I have completed my project.

Your parents have also been given an information sheet and consent form, but at the end of the day it is your decision to join me in the study.

Please feel free to speak to me if you have any questions.

Thank you

MS. G Barry
Tel: 011 593 5900
Email: gclbarry@gmail.com

APPENDIX C

INFORMATION SHEET LEARNERS- EXPERIMENTAL GROUP

DATE: 11 January 2016

Dear Learner

In addition to being your English teacher I am also a Masters student in the School of Education at the University of the Witwatersrand.

I am doing research on **The Effects of Withholding Marks and Providing Written Feedback for Grade 9 in English Home Language**.

My investigation involves developing comments that can be given to learners that will help them to understand how to improve their understanding of the work for future assessments. I will be investigating the way in which comments can be structured as well as the conditions under which they can be received. I am hoping that you will learn more about how to improve your work from the comments and be able to apply these skills for the rest of your schooling.

As a part of this study your marks will be kept secret from you until the end of term1. You will get your assessment back with very detailed comments which I will help you to work on. Your marks will still be carefully recorded and calculated as usual for your report and I will give you a record of all the individual marks at the end of the term.

I am asking you to join my research study for term 1 in 2016. The aim of the study is to determine the best ways in which comments on assessments can be structured and administered in order to have the maximum effect on learning at school.

If you agree to participate in this research study, there will be no additional work required from you nor any disadvantage as the teaching and assessment program will be the same as normal. The benefit for you in participating is that for the duration of the study you will be receiving very carefully constructed feedback that should help you to improve your skills in English.

This study is not a test of your ability, nor will you gain any marks for participating. It is voluntary, which means that you don't have to do it. Also, if you decide halfway through that you prefer to stop, this is completely your choice and will not affect you negatively in any way.

I will not be using your own name but I will make one up so no one can identify you. All information about you will be kept confidential in all my writing about the study. Also, all collected information will be stored safely and destroyed between 3-5 years after I have completed my project.

Your parents have also been given an information sheet and consent form, but at the end of the day it is your decision to join us in the study.

I look forward to working with you!

Please feel free to speak to me if you have any questions.

Thank you

Ms G. Barry
Tel: 011 593 5900
Email: gclbarry@gmail.com

APPENDIX D

QUESTIONNAIRE

Please read the questions carefully and fill in your responses below:

1. Do you find it easy to remember the comments that you have received on your work so far?

_____

2. Write down as many of these comments as you can remember in the space below.

_____
_____
_____
_____
_____

3. Do you feel that any of these comments have helped you to improve your work?

_____

4. If so, how?

_____
_____
_____
_____

APPENDIX E

INFORMATION SHEET PARENTS- CONTROL GROUP

DATE: 11 January 2016

Dear Parent

My name is Gwen Barry and I am a Masters student in the School of Education at the University of the Witwatersrand as well as your daughter/ward's English teacher.

I am doing research on **The Effects of Withholding Marks and Providing Written Feedback for Grade 9 Learners in English Home Language**

My research involves developing specialized and accurate comments on learner assessment responses written with a view to helping learners identify the areas in which they can perform better and the strategies that they can use in future assessments. Each learner will receive detailed written comments as part of normal classroom feedback on assessments. I will keep a record of these comments to determine their effectiveness as the term progresses.

The reason why I have chosen your child's class is because Grade 9 is a year in which the skills that have been learnt in Grade 8 are built upon and refined in preparation for the senior grades. It is my belief that comments on assessments play a particularly important role in helping learners gain the metacognitive skills that will help them to improve their won understanding of the content and engage more fully in the learning process so that they may achieve to their full potential. I am interested in finding ways to make feedback as relevant and helpful as possible.

Your child will not be advantaged or disadvantaged in any way. She will be reassured that she can withdraw her permission at any time during this project without any penalty. There are no foreseeable risks in participating and your child will not be paid for this study.

Your child's name and identity will be kept confidential at all times and in all academic writing about the study. Her individual privacy will be maintained in all published and written data resulting from the study.

All research data will be destroyed between 3-5 years after completion of the project.

Please let me know if you require any further information.

Thank you very much for your help.

Yours sincerely,

Ms G. Barry
Tel: 011 593 5900
Email: gclbarry@gmail.com

APPENDIX F

INFORMATION SHEET PARENTS- EXPERIMENTAL GROUP

DATE: 11 January 2016

Dear Parent

My name is Gwen Barry and I am a Masters student in the School of Education at the University of the Witwatersrand as well as your daughter/ward's English teacher.

I am doing research on **The Effects of Withholding Marks and Providing Written Feedback for Grade 9 Learners in English Home Language**

My research involves developing specialized and accurate comments on learner assessment responses written with a view to helping learners identify the areas in which they can perform better and the strategies that they can use in future assessments. Each learner will receive detailed written comments as part of normal classroom feedback on assessments. I will keep a record of these comments to determine their effectiveness as the term progresses.

The literature suggests that providing comments on assessments without providing the mark is particularly effective, therefore I would like to keep the marks for assessments secret for the duration of the first term. This will not affect the calculating of the report mark in any way and learners and parents will have full access to each individual mark at the end of term.

The reason why I have chosen your child's class is because Grade 9 is a year in which the skills that have been learnt in Grade 8 are built upon and refined in preparation for the senior grades. It is my belief that comments on assessments play a particularly important role in helping learners gain the metacognitive skills that will help them to improve their own understanding of the content and engage more fully in the learning process so that they may achieve to their full potential. I am interested in finding ways to make feedback as relevant and helpful as possible.

Your child will not be advantaged or disadvantaged in any way. She will be reassured that she can withdraw her permission at any time during this project without any penalty. There are no foreseeable risks in participating and your child will not be paid for this study.

Your child's name and identity will be kept confidential at all times and in all academic writing about the study. Her individual privacy will be maintained in all published and written data resulting from the study.

All research data will be destroyed between 3-5 years after completion of the project.

Please let me know if you require any further information.

Thank you very much for your help.

Yours sincerely,

Ms G. Barry
Tel: 011 593 5900

Email: gclbarry@gmail.com

APPENDIX G

Graph comparing grade 9 English term 1 marks to promotion marks

In the first term of Grade 9 there is a natural drop in marks as learners adjust to the demands of the new grade. This improves over the course of the year as indicated in Figure 3. This graph represents the number of learners who scored in each percentage distribution in the first term as opposed to later in the year.

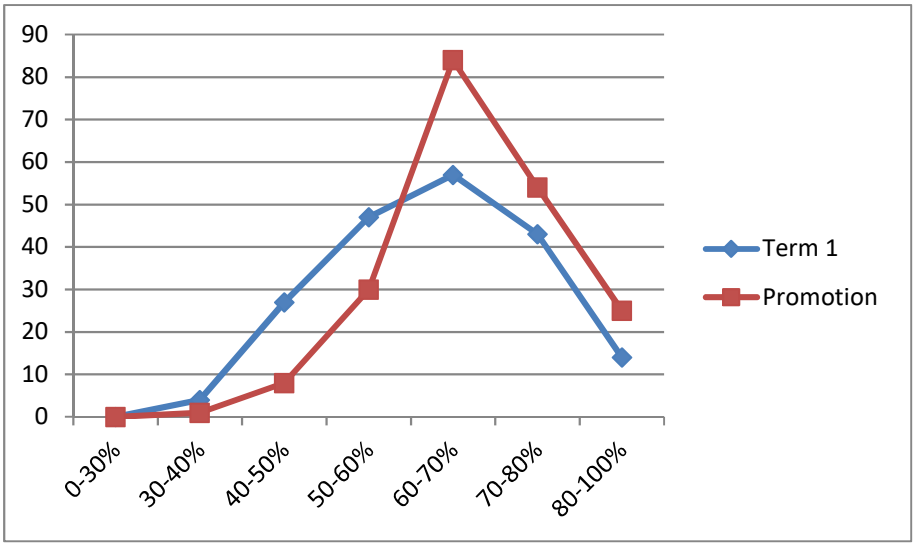


Figure 3: 2014. Comparison of Term 1 marks to Promotion marks.

Grade 9 English Home Language 2014.

APPENDIX H

LESSON PLAN FOR COMPREHENSION SKILLS

ENGLISH HOME LANGUAGE
GRADE 9 GET
LESSON PLAN: COMPREHENSION SKILLS

LEARNING TARGET:

- Revise the three main types of comprehension questions and how to recognise them.
- Revise the correct way of answering comprehension questions.
- How to de-code unfamiliar words

SUCCESS CRITERIA:

- Learners will be able to construct a well thought-out answer to each kind of comprehension question using a full sentence.

RESOURCES:

- White board and markers
- Practise comprehension exercise

METHODS:

- Question and answer
- Problem solving

STUDY METHODS TAUGHT:

- Metacognition
- Problem Solving
- Accessing and using prior knowledge

ACTIVITIES:

- Learners write down and discuss their own comprehension strategy and point out how it suits their personal working style.
-

ASSESSMENTS:
INFORMAL:

- Practise comprehension

FORMAL:

- Language and comprehension term test

Comprehension

- understand
- relate → recognise → classify
  - knowledge related to it
  - relate it to what you know
  - classify
  - you can explain

40% - basic recall
- answer found in text.

40% - explain
- go inbetween reason

20% - applying to outside world

My Comprehension Strategy

1. Read Passage (I want to get an idea of the story)
2. read questions (to know what's important in passage)
3. read Passage (to highlight important words)
4. Answer questions (I still have an idea of where)?

How to decode unfamiliar terms:

1. Read the text and come up with synonym
2. break the word up into prefix - root - suffix
   - apply similar
   - use meaning

① Answer in full sentence
- Subject + verb + object
- Every answer needs to make a statement about the Passage
- Always name the Subject (DO Not Start with a Pronoun)

APPENDIX I

PRACTISE COMPREHENSION AND MEMO

Read the extracts below and answer the questions that follow:

Anne Frank was a Jewish girl who lived in Amsterdam. During World War II, she and her family hid from the Nazis in an attic above an office. Anne kept a diary that she addressed as Kitty, during this time.

Extract 1.

Tuesday 20 October 1942

Dearest Kitty

My hands are still shaking, though it's been two hours since we had the scare. I should (2.4.1.) explain that there are five extinguishers in the building. The office staff forgot to warn us that the carpenter, or whatever he is called, was coming to fill the (2.4.2.) extinguishers. As a result we didn't bother to be quiet until I heard the sound of hammering on the landing (across from the bookcase). I immediately assumed it was the carpenter and went to warn Bep, who was eating lunch, that she couldn't go back downstairs. Father and I stationed ourselves at the door so we could hear when the man left. After working for about fifteen minutes, he laid his hammer and some other tools on our bookcase (or so we thought!) and banged on our door. We turned white with fear. Had he heard something after all and did he now want to check out this (2.4.3.) mysterious looking bookcase? It seemed so, since he kept knocking, pulling, pushing and jerking on it. I was so scared I nearly fainted at the thought of a total stranger managing to discover our wonderful hiding place…

Extract 2.

Monday 26 July 1943

Dearest Kitty

Yesterday was a very tumultuous day, and we're still wound up. Actually, you may wonder if there's ever a day that passes without some kind of excitement.

The first warning siren went off in the morning while we were at breakfast, but we paid no attention, because it only meant that the planes were crossing the coast. I had a terrible headache so I lay down for an hour after breakfast, and then went to the office at about two. At two-thirty Margot had finished (2.4.4.) her office work and was just gathering her things together when the sirens began wailing again. So she

and I trooped back upstairs. None too soon, it seems, for less than five minutes later the guns were booming so loudly that we went and stood in the passage. The house shook and the bombs kept falling. I was clutching my 'escape bag', more because I wanted to have something to hold on to than because I wanted to run away. I know we can't leave here, but if we had to, being seen on the streets would be just as dangerous as being caught in an air raid. After half an hour the drone of the engines faded and the house began to hum with activity again. Before long (2.4.5.) the smell of fire was everywhere, and outside it looked as if the city was enveloped in a thick fog.

A big fire like that is not a pleasant sight but fortunately for us it was all over and we went back to our various jobs. Just as we were starting dinner: another air-raid alarm. The food was good, but I lost my appetite the moment I heard the siren. Nothing happened, however, and fourty-five minutes later the all-clear was sounded.

QUESTIONS

1.1. What period of time has elapsed between the two diary entries? (1)

Refer to extract 1.

1.2.1. How do we know that Anne had got a fright about something? (1)

1.2.2. How do we know that the fright was quite serious? (1)

1.3. What alerted Anne to the fact that there was someone in the building? (1)

1.4.1. Why did Anne and her father stay next to the door? (1)

1.4.2. Explain why Anne perceived the presence of the carpenter as a threat. (2)

1.5. What does the carpenter do that causes Anne to nearly faint with fear? (1)

Refer to extract 2

1.6.1. What does the word "tumultuous" mean? (1)

1.6.2. What does is this word indicate about the kind of life that the people in the attic . live? (2)

1.7. Why didn't the family pay attention to the first warning siren? (1)

1.8. Why would it have been just as dangerous for the family to be out on the street being caught in an air raid? (2)

1.9. Why did Anne lose her appetite during dinner? (1)

TOTAL QUESTION 1: 15

MEMO

1.1. Nine months and six days have elapsed between the two diary entries. (1)

1.2.1. We know that Anne had gotten a fright because her hands were shaking. (1)

1.2.2. Anne's hands were still shaking for two hours after the event, which means that the fright was serious. (1)

1.3. The sound of hammering on the landing alerted Anne to that fact that someone was in the building. (1)

1.4.1. Anne and her father stayed next to the door so that they could hear when the carpenter left. (1)

1.4.2. Anne perceived the carpenter as a threat because he could possibly discover their hiding place (1) and report them to the Nazis which would result in them being arrested (1). (2)

1.5. The carpenter seems to be investigating the bookcase by pulling on it and this makes Anne nearly faint with fear. (1)

1.6.1. Tumultuous means loud, confused and disorderly (any one) (1)

1.6.2. This indicates that the people in the attic live a life that can be easily disrupted or disorganized (1) and that there is often tension and confusion (1) (2)

1.7. The family doesn't pay attention to the first siren because they believe that it only indicates that planes are crossing the coast and there is no immediate danger. (1)

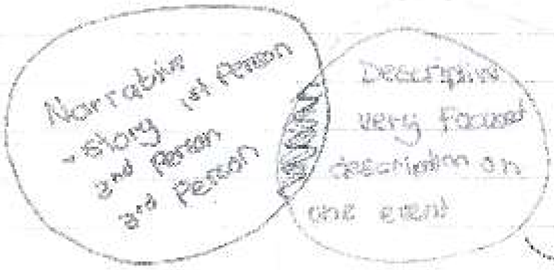
1.8. If the family were out on the street they could be recognised as Jews and arrested (1) however if they remain in the attic no one would know to look for them if the building was bombed because it is a secret hiding place (1) (2)

1.9. Anne lost her appetite because there was another air-raid siren and she became afraid again. (1)

APPENDIX J

LESSON ON ESSAY BASICS

ENGLISH HOME LANGUAGE GRADE 9 GET LESSON PLAN: REVSION OF ESSAY BASICS
LEARNING TARGET: • Revising the structure and elements of a creative essay • Learners will know how to go about constructing a narrative or descriptive essay SUCCESS CRITERIA: • Producing a coherent and interesting text
RESOURCES: • Discussion • White board and markers
METHODS: • Question and answer • Visual representation STUDY METHODS TAUGHT: • Metacognition • Planning
ACTIVITIES: • Writing of examples
ASSESSMENTS: INFORMAL: Exercise on strong verbs FORMAL: Formal essay task



Will have happend in the Past. Use PAST Tense

Reflection = Memory

What it means now.

Begining · introduction

Middle · body

End · conclusion

Set the scene, introduce event

Main event · explaining background

reflection, Second event.

Vocabulary

*Meat + Potatoes

complex *simple

Accuracy

are you giving the Full Picture

Said -

Walked -

Looked -

APPENDIX K

REVISION OF ESSAY BASICS

ENGLISH HOME LANGUAGE GRADE 9 GET LESSON PLAN: REVISION OF ESSAY BASICS AND ALTERNATIVE VOCABULARY
LEARNING TARGET: • Revising the structure and elements of a creative essay • Leaners will be able to use some new vocabulary to express emotions SUCCESS CRITERIA: • Producing a coherent and interesting text
RESOURCES: • Discussion • White board and markers • Vocabulary wheel
METHODS: • Question and answer • Visual representation STUDY METHODS TAUGHT: • Metacognition • Planning
ACTIVITIES: • Writing of examples
ASSESSMENTS: INFORMAL: Colour coding vocabulary wheel and chosing words to add to personal vocabulary FORMAL: Formal essay task

VOCABULARY WHEEL

APPENDIX L

MARKS SPREADSHEET

CONTROL GROUP

	Essay 4 Feb	Re-worked Essay	Essay 4 May	Practise essay	Essay 11 August
CB	26	26	29.5	39	40
TC	37	40	43	33	40
TCL	23	25	22.5	33	27.5
JC	26	26	27.5	25	29
TDJ	34.5	30	30	30	35
ZD	29	36	29	30	35
LJ	30	34	33.5	38	34
AK	29	29	31	30	30
YK	27	30	27.5	26	37.5
NK	31.5	41	26	34	30
AM	28	31	35	34	38
TM	28.5		27	26	25
SM	20.5	26	25	28	23
HM	34.5	38	35	36	40
TN	23		30	26	25
ON	30.5	32	31.5		32
BO	27.5	28	23	26	31
BR	23		25	19	26.5
IS	35	39	31	35	33
AS	29.5	29.5	26	26	29
TS	35	42	40	42	43
CT	29.5	30	27	26	25
NT	19.5	31	26	27	28
Average					

MARKS SPREADSHEET

EXPERIMENTAL GROUP

	Essay 4 Feb	Re-worked Essay	Essay 4 May	Practise Essay	Essay 11 August
DC	21	25	25		28
SD	35	38	31	33.5	28
ND	27.5	29	28.5		39
CE	24.5	29	26.5	29	40
KH	28.5			21	33.5
GH	28	32	31.5	30	33.5
CJ	27	30	26.5	35.5	44
AJ	28	30	28.5	34	35.5
BK	26		24.5	27	26.5
JLR	20	28	25.5	20	28
LM	21.5		25	36	28.5
TM	20.5	29	24	30	29
AM	25	28	28.5	30	35
TMO	25	29	29	36	25
AN	21	24	28	21	25
RN	24.5	25	25	27.5	25
KO	33	36	32	30	40
AR	30	34	30	31	35.5
KS	23	24	25.5	22.5	25
KSO	24.5	28	29	30	30
MT	26.5	29	29	27	23.5
NV	27.5	32	34	33.5	40
Average					

APPENDIX M

COMMENTS SPREADSHEET

CONTROL GROUP

	Comments 4 February			Comments 4 May			Comments Practise Essay		
	A	V	S	A	V	S	A	V	S
CB	A2	V1	S4	A2	V1	S4		S5	V2
TC		V2			V2		A5	V2V3	
TCL	a2	v1	s3	**a2**	**v1**	s3		S3S4	V1
JC	A2	v1	s3	A2	v1	s3	A3	S1	V2
TDJ	a3	v1	s3	**a3**	**v1**	**s3**	A3		V1V4
ZD	A2	v1	s3	A2	v1	s3	A3		V3V4
LJ	a3	v1	s3	**a3**	**v1**	**s3**		S4	V3
AK	a3	v1	s5	a3	**v1**	**s5**	A3	S1	V1
YK	a3	v1	s6	a3	v1	s6	A3	S4	V4
NK		v3	s3		v3	s3	A3	S3	V1
AM	a3	v1	s1	a3	**v1**	s1	A3	S4	V1
TM	a3	v1	s3	a3	v1	**s3**	A3	S3	V2
SM	a3	v1	s4	**a3**	v1	**s4**	A3	S4	V1
HM	a5	v1	s5 s6	**a5**	v1	s5 **s6**			V2V3
TN	a3	v1		a3	**v1**		A3	S1	V1
ON		v1	s1		v1	s1			
BO	a2	v1	s3	a2	**v1**	**s3**	A3	S4	V1
BR	A2	v1	s3	A2	v1	s3	A5	S5	V1
IS		v3	s6		v3	S6	A3	S1	
AS	A2	V3	s1	A2	V3	**s1**	A3	S6	V1
TS		v1			**v1**			S5	V5
CT	a2	v1	s6	a2	v1	s6	A3	S1	V1
NT	a2	v1 v2	s3	a2	v1 v2	s3		S4	V2

Addressed
Remembered

COMMENTS SPREADSHEET

EXPERIMENTAL GROUP

	Comments 4 February			Comments 4 May			Comments Practise Essay		
	A	V	S	A	S	V	A	V	S
DC	a2	s3	v1	**a2**	s3	v1		S3	V1
SD	a3		v1	a3		v1			
ND	a3	s7	v1	**a3**	**s7**	**v1**	A3		V1
CE	a2		v1	**a2**		**v1**	A3	S3	V1
KH	a2	s3 s5	v1	**a2**	s3 s5	v1	A3		V1V3
GH	a3	s6	v1 v2	a3	s6	v1 v2	A3	S6	
CJ	a3	s5	v1	a3	s5	v1	A3	S6	V5
AJ	a2 a3		v1 v2	a2 **a3**		**v1** v2	A3	S4	V5
BK	a2	s3	v1	a2	s3	**v1**	A3	S7	V1
JLR	a4	v3	s5	a4	S5	V3		S4	V3
LM	a3	s3	v1 v2	a3	s3	**v1** v2	A3	S3	V1
TM	a2	s1 s3	v1	a2	s1 s3	**v1**		S3	V3
AM	a3	s3	v1	a3	s3	**v1**		S4	
TMO	a3	s4	v1	a3	s4	**v1**	A5	S1	V1
AN	a2	s3	v1	a2	s3	**v1**	A3	S5	V1
RN	a2	s3	v1	a2	**s3**	**v1**	A3	S1	V1
KO		s3	v1		**s3**	**v1**	A3	S1	V1
AR	a3		v1	**a3**		v1	A3	S6	V1
KS	a2	s1 s3	v1	a2	s1 s3	v1	A3	S4	V1
KSO	a2	s1	v1	**a2**	s1	v1		S1	V1
MT	a3	s3	v1	**a3**	**s3**	**v1**	A3	S3	V1
NV	a3	s1	v1	**a3**	s1	v1	A3	S3	V1

Addressed
Remembered

APPENDIX N

1ST FORMAL ESSAY TASK

Instructions

1. Choose ONE of the following topics and write a creative essay of between 250-300 words.
2. There is no need to write a rough draft but make sure you have adequate planning to accompany your essay.

Question 1.

"Don't you wish you could take a single childhood memory and blow it up into a bubble and live inside it forever?"

- **Sarah Addison Allen**

Write an essay of **250-300** words in which you describe your favourite childhood experience.

Question 2.

"The highlight of my childhood was making my brother laugh so hard that food came out of his nose."

-Garisson Kellior

Write an essay of **250-300** words in which you describe a humorous incident from your childhood.

Question 3.

"There is always one moment in childhood when the door opens and lets the future in"

-Deepak Chopra

In an essay of **250-300** words, describe a moment in your childhood when life as you knew it changed forever.

APPENDIX O

2ND FORMAL ESSAY TASK

Select one of the topics below and write a descriptive or narrative essay of 250-300 words.

Topic One:

I must have missed it when I blinked... (50 marks)

Topic Two:

(50 marks)

Topic Three:

"You must not fight too often with one enemy, or you will teach him all your art of war.' – **NAPOLEON BONAPARTE** **(**50 marks)

Topic Four:

(50 marks)

Topic Five:

We didn't realise we were making memories; we just knew we were having fun.

(50 marks)

Topic Six:

(50 marks)

APPENDIX P

PRACTISE ESSAY TASK

Write a creative essay of 250-300 words using ONE of the following topics for inspiration. Indicate which topic you have chosen ON THE MARKING GRID.

TOPIC 1.

TOPIC 2.

“It takes a great deal of bravery to stand up to our enemies, but just as much to stand up to our friends.”

-J. K. Rowling

TOPIC 3.

“Don't Make Assumptions. Find the courage to ask questions and to express what you really want. Communicate with others as clearly as you can to avoid misunderstandings, sadness and drama. With just this one agreement, you can completely transform your life.”

-Miguel Angel Ruiz

TOPIC 4.

TOPIC 5.

APPENDIX Q

3RD FORMAL ESSAY TASK

Choose **ONE** of the following topics and write a descriptive or narrative essay between **200-300** words.

Topic One:

Topic Two:

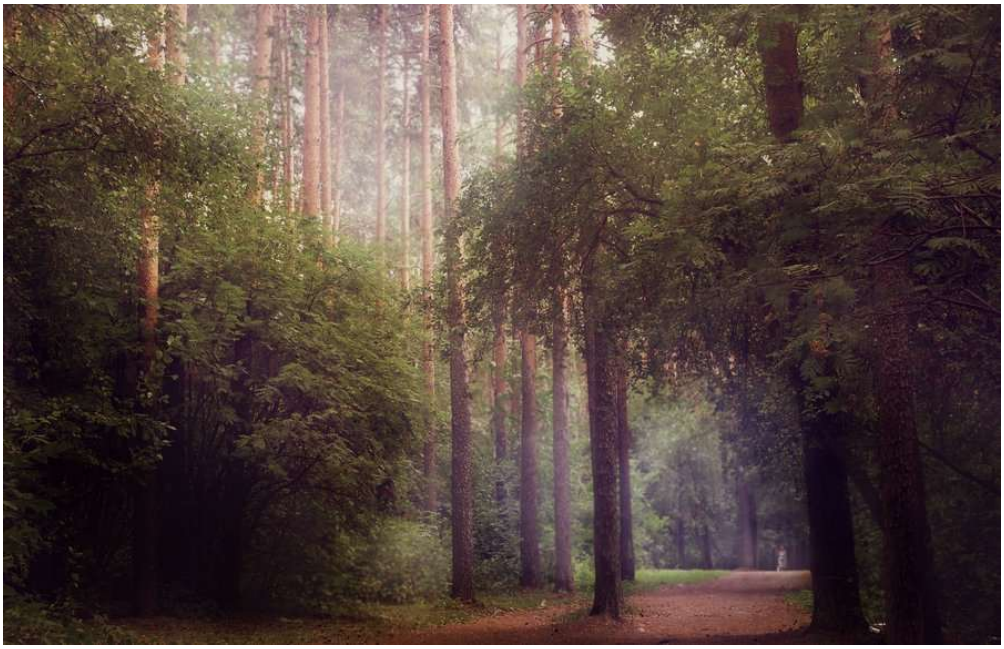
A villain is the hero of her own story.

Topic Three:

"A friend is someone who knows all about you and still loves you."
-Elbert Hubbard

Topic Four:

Topic Five:



Topic Six:

She unwrapped the gift and was shocked to see what was inside...

APPENDIX R

COMMENTS CODE

Comments Code

Structure

1. Start the sentence by naming the subject/ using a proper noun or noun.
2. Answer using a full sentence that also makes reference to the question.
3. Make sure you use a new paragraph for a new idea.
4. Sentences are too long.
5. Punctuate sentences.
6. Vary sentence length to enhance writing
7. Use paragraphs

Vocabulary

1. Work for a more varied and accurate vocabulary.
2. **Spelling**
3. Keep the tense constant
4. Use of slang
5. Misuse of words (common errors)

Accuracy

1. Provide the number of detailed reasons that match the mark allocation.
2. Read the question carefully (misinterpretation)
3. Answer in more depth and give detailed explanations.
4. Focus the essay more on the topic.
5. Writing vague and doesn't come to the point

APPENDIX S

GRADE HEAD BEHAVIOUR REPORTS

CONTROL GROUP

Grade 9 Feedback

9A

DATE	FEEDBACK	TEACHER
18/03	Class disruptive and weak. Hard to get learners to work	KH

GRADE HEAD BEHAVIOUR REPORTS

EXPERIMENTAL GROUP

Grade 9 Feedback

9D

DATE	FEEDBACK	TEACHER
03/02	Rude Weak Disruptive	ML,RS,GB
01/04	Class hid teacher's board markers and overhead remote	ML
25/04	Very weak Disruptive	VR, ML, RS
24/05	Learners are weak but are trying hard in certain classes (English, Biology and Maths)	Meeting of all teachers
03/06	Teasing of other learners	KH