

Johannes Sibeko

Nelson Mandela University

An evaluation of the assessment of creative writing essays in the Further Education and Training band

Abstract

The current process of assessing creative writing essays (using the correction code and the rubric to mark the pre-final and final drafts of essays in the Further Education and Training band) does not provide useful feedback that learners can use to improve their creative essay writing skills. This discursive paper highlights the basic flaws of the process. Amongst others, is the issue of the holistic nature of the rubric feedback presented to learners on their essays, the generic essence of the rubric in assessing different essay types as though they are similar and the divergent focus of the rubric feedback on

the macro-scale issues from the focus of the correction code on micro-scale essay features. The argument here is that the different foci of the two assessment tools leave learners with uncertainty on what they need to improve in their writing, thus impairing the whole process of assessment. This study recommends, inter alia, an inclusion of rubric feedback in pre-final drafts and the expansion of rubric feedback presented to learners.

Keywords: Assessment; Further Education and Training; Holistic rubrics; Analytic rubrics; Correction Codes; CAPS; Creative writing essays

Introduction

Desirably, feedback provided to learners should allow them to realise their own strengths and weaknesses, the result of which would be; understanding how to move on in their learning (Department of Basic Education, DBE, 2011a:11). Unfortunately, as will be argued in this paper, the process-approach to the assessment of creative writing essays does not yield this desired outcome. Following essay assessment, learners are left guessing what they could have improved.

Typically, the tools used for assessing essays in South African high schools are designed at the head office of the DBE and subsequently circulated by the district offices. Subject advisors are then taught to train teachers in the use of the tools (Sibeko, 2015:156). However, given that teachers have no contribution to the construction of rubrics and correction codes, it may follow that they hold somewhat different notions as to what each element of the rubric entails (Hattingh, 2009:150; McNamara, 1996:125). This is exemplified in Sibeko (2015:184), wherein teachers seem confused as to what the criterion of style entails. Another important finding in the same study is that the participating teachers are seemingly more content with the rubric than with the correction code (Ibid, 2015:169). This finding provides one of the bases for this paper. It follows that there is imbalance of interest if teachers are more content with one tool as opposed to the other. More so, if the two tools focus on different aspects. The generalising inference in this study is that the non-specific, i.e. the holistic feedback provided through the rubrics hinders transparent marking, thus allowing teachers not to motivate the marks that they assign to learners' essays.

It is not in the scope of this paper to investigate what goes on in the minds of the teachers when they assess essays. Insights on cognitive processes and the rating of essays are discussed elsewhere (c.f. Hout, 1990; Lumley, 2002; Cumming et al 2001; Cumming et al., 2002). Furthermore, this paper does not provide insight into personal factors that influence divergence in teachers' perceptions of essay qualities and in turn affect their assessments of essays (for that purpose, see Hamp-Lyons, 1991). This paper seeks to highlight the flaws in the assessment and feedback practices in creative writing essays in the Further Education and Training (FET) band, that is, grades 10 to 12 in South African schools. In this order, the principles of marking are discussed, followed by a consideration of the debate on feedback practices, an overview of correction codes and rubrics, and a comprehensive discussion of the current process of assessment in the FET band.

Principles of marking

Fry *et al.* (2009:144) lists six principles of marking, namely: (i) consistency, (ii) reliability, (iii) validity, (iv) levelness, (v) transparency, and (vi) inclusivity. Using an appropriate rubric can allow realisation of all six principles. One, a good rubric that is well understood by markers ensures uniformity in their marking by ensuring that specific essay aspects

are penalised or rewarded consistently and that the same aspects are assessed for throughout all the essays. Two, it should ensure that marks are reliable and not subjective by allowing different markers to grant similar grades to similar essays and ensuring that even the same teacher rates the essay similarly if blindly re-marking it at a later stage. This is achieved through consistency in marking for similar aspects as would be enforced by a clear rubric. Three, it should assess for valid aspects of the essay in that it covers only the aspects that are expected in the type of essay written. Currently, only one general essay rubric is used to assess different essay types. Although it is seemingly not desired to have a general essay, learners in the FET tend to write hybrid essay types. For instance, a learner can write a descriptive essay that has narrative essay characteristics. Thus, the general rubric works better than an essay type specific rubric. Four, it should be of the right level. For instance, it must not assess grade 10 essays for aspects that are only expected to be mastered in grade 12. Likewise, it should not assess for aspects that are expected to be mastered at lower grades as this would not result in learners' improvement to expected standards. However, the same aspects are assessed for throughout the FET band since one rubric is used from grades 10 to 12. It is therefore the teacher's responsibility to relax the assessment standard. Five, a good rubric ensures that learners know exactly why they are penalised or rewarded. This means that the rubric should have clearly explained criteria that learners can easily follow in a reflective process after receiving their marked essays. According to Whipp (2011:168) the nature and range of the knowledge necessary for successful completion of assessment tasks should be transparent before learners attempt the task. A clear analytic rubric can help attain this. Finally, a good rubric does not focus on one single skill. Instead, it ensures that learners' different strong points are rewarded, thus attaining inclusivity. This should not be confused with levelness in that although different abilities are assessed, the level should still be of the right standard. Accordingly, Brown (2004:83) states that "inclusivity involves deploying a variety of methods for assessment so that the same students are not always disadvantaged," or advantaged. A rubric that assesses multiple aspects in an essay such as those used in the assessment of creative writing essays in the FET band in South African schools stands a better chance of representing the different learners' skills than one that is restricted to limited essay aspects.

Assessment

According to Hounsell (1995:51), assessment is the act of "making informed and considered judgements about the quality of a student's performance on a given assignment." That is, teachers cannot haphazardly assign marks to learner essays, they need to make informed and considered judgements and be able to justify the grades they assign to each essay. In principle, Hounsell's definition restricts grading learner achievement to a specific task. According to Brown et al. (1997:08) when assessing, you take a learner's specific task, analyse it and decide on the marks they should get. At the end, the gathered evidence of learner performance is expected to inform pedagogical decisions (Chappuis et al., 2010:13). This means that assessing learners' essays should

allow us to make inferences on their improvement as writers. Such inferences should form basis for alterations to pedagogical approaches. In the end, learners ought to be informed of the quality of their performance, hence the provision of feedback.

According to Lockett and Sutherland (2000:98), assessment defines four important things for learners, (i) it tells learners what is important, (ii) it indicates what counts, (iii) it tells learners how they will spend their time and (iv) it reflects to them, what kind of learners they are. This would mean that what learners are assessed on is flagged as important, what they are not assessed on does not necessarily count, and that they will spend more time on what is important and counts than on what is not assessed. Furthermore, success or failure to achieve the expected outcomes reflects the types of learners they are. The correction code and the rubric are used for such a purpose in the teaching and assessment of creative writing essays in the FET band. The correction code indicates grammatical constructions that need improvement while the rubric provides overall feedback on their final performance. Unfortunately, using the correction code imposes negative marking. Inconveniently, the inclusion of positive codes cannot counter the negativity inherent in the code itself (Spencer, 2009:25; Sibeko, 2015:39).

Generally, Withers' (1987) description of the basic marking strategy in Australia in the 1980s, captures the current practice in South African high schools. According to Ibid (1987:7):

the basic marking strategy used by a teacher is to pick up the next piece of work in the pile, look carefully at it, and subject it to judgment. This judgment will be more or less agonised, according to how long the piece of work has been in the pile, and how big the pile is. The judgment will also be more or less subjective.

Unfortunately, this means that the workload of the teacher determines the effort and dedication with which each essay is assessed. Favourably, due to the use of the correction code as prescribed by the DBE, it can be guaranteed that the teacher will at least read the whole essay. Sadly, it cannot be guaranteed whether the teacher focuses on error identification or listens to what the learner is saying. Desirably, the marking teacher should read the essay and grade the first criteria (content and planning), re-read the essay and mark the second criteria (language, style and editing) and re-read the essay to grade the structure (c.f. DBE, 2018:3). Whether teachers assess essays in this fashion given their varied workloads has not been empirically proven.

Feedback debate

According to Louw (2009:86) feedback is a given in most South African teaching institutions. For instance, Spencer (1998:208) states that specific, encouraging and honest criticism feedback is expected by students. The problem, however, is

that although learners expect feedback, they do not know why they expect it (Louw, 2009:86), and they sometimes do not even know what to do with it (Louw, 2006:7). For instance, Hyland (2003:218) states that learners lack the ability to understand and use feedback such as language correction. Furthermore, Louw (2009:86) states that even though teachers provide feedback, they lack clarity on why they ought to provide it. Also, even though teachers provide feedback, they are not always able to explain learners' problems (Hyland, 2009:218).

Louw (2006:4-8) lists four problems with feedback. First, there is lack of consistency when there is no standard set. Second, feedback is labour intensive and time consuming. An interesting finding in Sibeko (2015:85) was that teachers reported marking an average of 1414 grade 10 Sesotho HL essays each year. Beyond doubt, the number of essays each teacher assesses coupled with the need to read each essay three times while grading (DBE, 2018:3) results in labour intensity. These issues of time consumption and labour-intensity limit any attempts at improving creative essay assessment and ensuring quality feedback. Third, students expect feedback but do not always know how to handle it. As such, Louw (2006:7) may be correct in stating that teachers' efforts in providing feedback are wasted when students do not know how to use it. Finally, students do not recognise recurring patterns of errors in their own writing. Even though learners receive feedback from the teacher and their peers they replicate them in their final drafts. This replication of errors indicates that providing feedback on drafts might not be as effective as envisaged. Even so, encountering the same error that was previously identified is not concrete evidence that there was no engagement with the feedback previously received. This notion of repeated errors can be explained using James' (2007:239) three reasons why learners write grammatically incorrect constructions, namely; errors, mistakes and occasional slips. First, a learner may not know how to spell some words [i.e. produce errors], misuse a certain punctuation sign [i.e. make a mistake] or sometimes get the spelling right while occasionally getting it incorrect [i.e. occasionally slip]. The correction code is used to point these three types of issues to the learner. However, if the use of the correction code by the teacher does not result in learners' improvement and correction of these issues, it means that the feedback is ineffective either due to disuse or lack of skills in interpreting it.

Correction codes

According to Spencer (2012:32), in the correction code strategy, students are given explanations of the codes used for marking and only the codes are used in the actual marking. Consider figure 1 following for an example use of correction codes in the marking of Sesotho HL essays in grade 10 essays discussed in detail in Sibeko (2015) and an illustration of marking in English FAL in figure 2:

feedback should be neutral and not red, they also report less than 1% difference in the harshness of marking between the red pen and the more neutral aqua pen.

A problem with basing grades on the errors learners produce is that the way in which teachers perceive errors is closely related to their experiences (Hyland and Anan, 2006:510). In other words, they mark and grade based on their “personal pedagogical systems – stores of beliefs, knowledge, theories, assumptions, and attitudes,” (Borg, 1998:9). As such, the effect an error has on the overall mark is subjective to an individual teacher’s perception of the impact of the specific error. Unfortunately, the issue of how teachers recognise, and judge errors has not been studied much (Hyland and Anan, 2006:510). The rubric does not help in justifying the effects of errors on the overall marks since it only evaluates the number of errors and not the types or effects thereof (c.f. table 2).

According to Spencer (2012:32) the correction code method can be used like road signs. The ability to incorporate the suggested changes would then be a clear indication of the extent to which the feedback is understood (Gilliland, 2015:295). There is however a concern of whether learners understand the codes (Spencer, 2012:32), although the codes are explained to them. Accordingly, participants in Sibeko (2015:154), report that Sesotho grade 10 learners are not always clear on what each code means, this is owed to the fact that learners neglect incorporating basic orthography-based changes

Holistic marking and rubrics

This paper contends that holistic marking allows teachers in the FET band to avoid making or accounting for grading judgments. It is however acknowledged that more standardised procedures used to ensure scoring validity and grades reliability have been tested in other investigations. For instance, the advantages of automated scoring (Whithaus, 2005; Ericsson and Hasswell, 2006) where student writing is judged by machines (Hamp-Lyons, 2007), an investigation of the theoretical best practice in using computer assisted language learning in the South African context (Spencer and Louw, 2008) and the benefits thereof in feedback (Louw, 2008). According to Weinberger et al. (2011:11) these standardised technology-based systems assist with countering the labour-intensity of essay marking. Sadly, as impressive and promising as these advancements may be, it is unfortunately still in the far future of most South African schools to take advantage of them. For instance, most township schools do not have computers for learners to type. In such schools, the teacher with the assistance of the rubric and correction code is the most valued participant in the assessment of the learners’ essays.

A rubric is a document that explains the expectations for an assessment task by listing all the criteria upon which it is or is to be judged and describes different levels of achievement such as poor and excellent (Andrade et al., 2008:03). In this way, the rubrics do not measure true scores (Thompson, 1995:198). Even so, rubrics can help to

legitimise grades (Kohn, 2006:12). This is achieved through ensuring that teachers focus on the assessment of similar features in performances to ensure an acceptable level of consistency for each teacher and agreement between different teachers (Hattingh (2009:136).

Weigle (2002:72) recommends choosing between holistic and (multiple trait) analytic rubrics in the assessment or evaluation of essays. The two types of rubrics are accepted as widely used and valued in the field of writing assessment (Hattingh, 2009:137; Shaw and Weir, 2007:149). The current rubric used for FET creative writing essays combines both analytic and holistic aspects. Unfortunately, even with such good intentions, the rubric is not optimal in the provision of best feedback.

The difference between the holistic and the analytic rubrics

The main difference between holistic and analytic rubrics is focus (Hatting, 2009:137). According to Carr (2000:228), only one construct is assessed in holistic scoring. As such, instead of focusing on the finer details, the holistic rubric looks at the whole category of marks. Conversely, the analytic rubric tries to show how the marks for the category were accumulated by showing how each sub-criterion thereof combines to make one whole thereby informing learners of their individual strengths and weaknesses (Weigle, 2007:203). Holistic rubrics are easier to use (*Ibid*) as there are fewer aspects to consider. However, they do not provide enough detail as analytic rubrics (*Ibid*) and are thus less desirable for learning (Bacha, 2001:380-1).

The holistic rubric allows the teacher to simply award a single score based on his/her overall impression and subjective interpretation of the text (Davies et al., 1999:75; Hattingh 2009:133). It is to this undesirable end that Davies (2004:262) refers to essays as a way to assess subjective skills. The rubric allows the teacher to be subjective in his marking and to mark the subjective skills of the learner regardless of the type of rubric used (*Ibid*). That is, both the analytic and the holistic rubrics fail to annihilate the subjectivity of essay assessment.

Although holistic rubrics are not preferred (Dipado et al., 2011:171), they are not devoid of benefits. For instance, they are fast and inexpensive (Shaw, 2002:11; Hattingh, 2009:133). These are welcome benefits in most schools with overcrowding issues and for teachers who teach a lot of language classes.

In contrast to the holistic rubric, although the analytic rubric does not totally obliterate subjectivity, it has the advantage of increased reliability of the final marks based on the fact that teachers assign multiple marks to judge a single performance (McNamara, 2000:44; Hattingh, 2009:136). The effect extends from intra-rater to inter-rater reliability. This abates the problem indicated by Yürekli and Üstünlüoğlu's (2007:56) concern that markers focus on different essay elements and thus respond to different facets of writing. This issue of focusing on different aspects is aggravated by the use of holistic rubrics as opposed to analytic rubrics.

Unfortunately, across different languages in South Africa, teachers have limited control on the issue of choosing rubrics. Although the rubrics are prescribed, teachers are not involved in their construction. This imposition is typical of holistic and sometimes analytic scales (Fulcher and Davidson, 2007:96; Hattingh, 2009:138). In fact, Hattingh (2009) indicates that the rubrics used in the FET band have not been validated. A case for the validation of the rubrics is made therein.

The types of essays

Comparatively, the *Setatementsa sa Leano la Kharikhulamu le Tekanyetso* (DBE, 2011c:47-9) and the Curriculum and Assessment Policy Statement (CAPS) (DBE, 2011a:45; DBE, 2011b:40) list five types of essays that learners in the FET phase should be taught for paper 3, namely, (i) narrative, (ii) argumentative, (iii) discursive, (iv) descriptive and (v) reflective. The problem is that although essays are differentiated in the CAPS documents, only one rubric is prescribed respectively at HL, FAL and Second Additional Language (SAL) levels, impeding the notion that there is dissimilarity between the essay types.

There are three possible rationalisations why only one rubric is used to assess different creative essay types. The first justification may be to prevent chances of confusing learners. The second contemplation is that teachers are not experts at identifying essay types from learners' writing and thus need a general rubric. The third possibility is that learners are expected to write somewhat hybrid essays. Each creative writing task presents learners with multiple essay topics and pictures to allow learners to choose any topic or come up with their own topics. The disadvantage to this practice is that multiple essay topics lower expectations of reliability as opposed to one compulsory topic (Brennan, 1996:8; Jonsson and Svingby, 2007:135). Consequently, the standard with which teachers assess the different essay topics is not consistent as different essay aspects are expected from each type of essay. Furthermore, topic-specific rubrics stand better chances of achieving the goal of dependable scores (DeRemer, 1998; Marzano, 2002). As such, introducing essay type specific rubrics would impact mark reliability by exposing the fact that learners are assessed on different tasks as though they are the same. Unfortunately, given time-constraints and labour-intensity of essay marking, adding the fourth reading of essays which would be aimed at identifying essay types might not be feasible. As a result, the general rubric suffices.

The micro and macro-scale issues

The correction code is used to provide feedback on both the drafts and final submissions marked by the teacher while the rubric is restricted to the final submission of the essay. However, the rubric and correction code focus on different essay aspects. Table 1 indicates the micro and the macro issues indicated by the correction code currently

used to provide feedback to learners. Each HL subject has its tailored correction code appropriate to the specific language. Table 2 identifies the micro and macro scale essay aspects for the rubric.

Table 1: Micro and Macro scale errors indicated by the correction code

Micro-scale errors		Macro-scale errors
Orthography	Split + combine words	Unclear constructions
	Orthography + spelling	Split paragraphs
Lexical	Word choice	Combine paragraphs
	Missing + Unnecessary items	
Punctuation	Punctuation + Capitalisation	

Table 2: Micro and Macro-scale issues indicated by the current rubric

	Micro-scale issues	Macro-scale issues
Content and planning		Content: Originality Appropriateness Coherence Ideas: Imaginative Development of details The visibility of planning and the obviousness of the outcome of planning
Language, style and editing	Use of punctuation marks Adequateness of word choice	Awareness of impact of language Standard of language used Appropriateness of style, tone + register Number of errors present post-editing
Structure	Adherence to length specifications	Construction of sentences + Paragraphs Suitability of the introduction and the conclusion

The classification of micro and macro elements in table 1 and table 2 explicates the fact that the correction code focuses on the identification of minor issues while the rubric largely focuses on feeding back on macro issues. Regrettably, this resultant inconsistency between the feedback for improvement (i.e. the correction code) and the feedback for judgment of the final product (i.e. the rubric feedback and correction code) means that one form of feedback cannot be used to give a clear picture of learner performance. In terms of grading, the focus of the correction code on minor issues might not be problematic as the rubric is seemingly used to grade independent of the correction code. However, in terms of essay improvement, this divergence of foci proves as problematic as learners are not guided on fixing the content of their essays and in expressing their creativity more explicitly.

The number of errors present in the essay after editing has been indicated as a macro-scale issue in table 2. This is because the learner's ability to correctly incorporate suggested changes into an essay is an indicator of the extent to which the learner is familiar with the correct form. A successful revision of the areas of concern indicated in the draft essay shows conversancy with the correct form. In contrast, an unsuccessful attempt to improve the draft indicates unacquaintance with the correct form. Unfortunately, most of the codes represent more than one error, resulting in a possible confusion of what the codes represent in the essay (Sibeko, 2017:82). That is, a learner might unwittingly misinterpret the correction code and effect wrong changes thereby introducing new errors.

Awarding marks through the current rubric

According to the CAPS (DBE, 2011c:45) teachers can choose not to provide feedback on assessments that do not form part of the Continuous Assessment (CASS) mark of the learner. Consequently, teacher feedback on writing may be limited to recorded assessment tasks. In cases where the teacher chooses not to mark CASS tasks, self- and peer generated feedback is still provided as mandated by the Subject Assessment Guidelines (DBE, 2008:1) that all tasks be assessed. This option to exclude teacher feedback means that every piece of feedback provided by the teacher should yield maximum effect.

Generally, the rubric used to grade creative essays in the FET band spreads marks across three main criteria. A total of 60% is reserved for content and planning, 30% for language, style and editing and 10% is earmarked for structure. The marks are divided into five levels of achievement for each of the three criteria in the HL rubrics across different languages. The rubric used for Home Language creative writing essays is presented in table 3 following:

Table 3: Rubric used for HL essay marking (DBE, 2018:7)

Criteria	Exceptional	Skilful	Moderate	Elementary	Inadequate
CONTENT AND PLANNING (Response and ideas) Organisation of ideas for planning; Awareness of purpose, audience and context 30 MARKS	28–30 - Outstanding/ Striking response beyond normal expectations - Intelligent, thought-provoking and mature ideas - Exceptionally well organised and coherent, including introduction, body and conclusion	22–24 - Very well-crafted response - Fully relevant and interesting ideas with evidence of maturity - Very well organised and coherent, including introduction, body and conclusion	16–18 - Satisfactory response - Ideas are reasonably coherent and convincing - Reasonably organised and coherent, including introduction, body and conclusion	10–12 - Inconsistently coherent response - Unclear ideas and unoriginal - Little evidence of organisation and coherence	4–6 - Totally irrelevant response - Confused and unfocused ideas - Vague and repetitive - Unorganised and incoherent
	25–27 - Excellent response but lacks the exceptionally striking qualities of the outstanding essay - Mature and intelligent ideas - Skilfully organised and coherent, including introduction, body and conclusion	19–21 - Well-crafted response - Relevant and interesting ideas - Well organised and coherent, including introduction, body and conclusion	13–15 - Satisfactory response but some lapses in clarity - Ideas are fairly coherent and convincing - Some degree of organisation and coherence, including introduction, body and conclusion	7–9 - Largely irrelevant response - Ideas tend to be disconnected and confusing - Hardly any evidence of organisation and coherence	0–3 - No attempt to respond to the topic - Completely irrelevant and inappropriate - Unfocused and muddled

LANGUAGE, STYLE AND EDITING Tone, register, style, vocabulary appropriate to purpose/effect and context; Word choice; Language use and conventions, punctuation, grammar, spelling 15 MARKS	Upper level				14-15	11-12	8-9	5-6	0-3
	<ul style="list-style-type: none"> - Tone, register, style and vocabulary highly appropriate to purpose, audience and context - Exceptionally impressive use of language - Compelling and rhetorically effective in tone - Virtually error-free in grammar and spelling - Very skilfully crafted 	<ul style="list-style-type: none"> - Tone, register, style and vocabulary appropriate to purpose, audience and context - Language is effective and a consistently appropriate tone is used - Largely error-free in grammar and spelling - Very well crafted 	<ul style="list-style-type: none"> - Tone, register, vocabulary appropriate to purpose, audience and context - Appropriate use of language to convey meaning - Tone is appropriate - Rhetorical devices used to enhance content 	<ul style="list-style-type: none"> - Tone, register, vocabulary less appropriate to purpose, audience and context - Very basic use of language - Tone and diction are inappropriate - Very limited vocabulary 	<ul style="list-style-type: none"> - Language incomprehensible - Tone, register, style and vocabulary not appropriate to purpose, audience and context - Vocabulary limitations so extreme as to make comprehension impossible 				
STRUCTURE Features of text; Paragraph development and sentence construction 5 MARKS	Lower level				13	10	7	4	0-1
	<ul style="list-style-type: none"> - Language excellent and rhetorically effective in tone - Virtually error-free in grammar and spelling - Skilfully crafted 	<ul style="list-style-type: none"> - Language engaging and generally effective - Appropriate and effective tone - Few errors in grammar and spelling - Well crafted 	<ul style="list-style-type: none"> - Adequate use of language with some inconsistencies - Tone generally appropriate and limited use of rhetorical devices 	<ul style="list-style-type: none"> - Inadequate use of language - Little or no variety in sentences - Exceptionally limited vocabulary 	<ul style="list-style-type: none"> - Necessary points lacking - Sentences and paragraphs faulty - Essay lacks sense 				
	5	4	3	2	0-1	<ul style="list-style-type: none"> - Excellent development of topic - Exceptional detail - Sentences, paragraphs exceptionally well-constructed 	<ul style="list-style-type: none"> - Logical development of details - Coherent - Sentences, paragraphs logical, varied 	<ul style="list-style-type: none"> - Relevant details developed - Sentences, paragraphs well-constructed - Essay still makes sense 	<ul style="list-style-type: none"> - Some valid points - Sentences and paragraphs faulty - Essay still makes some sense

A different rubric is used in the HL, FAL and SAL respectively. The rubrics are translated to different languages and are similar across all levels. For instance, all HL level rubrics are similar across different home languages in the FET band. One rubric is used from grade 10 to grade 12. For comparison, see the grade 10 paper 3 memorandum (DBE, 2017:8) and the grade 12 paper 3 memorandum (DBE, 2018:8).

As stated earlier, although the rubric used by teachers is analytic in theory, the feedback they provide to learners is holistic in practice. This is evident in the code used by teachers to provide rubric feedback. Table 4 presents the codes used in Sesotho, Zulu, Xitsonga and English home language level in the FET band as examples:

Table 4: The codes used to provide rubric feedback on learner essays

	Sesotho	Zulu HL	Xitsonga HL	English FAL
Content and planning	DM	L	VM	CP
Language style and editing	PST	Q	RXV	LSE
structure	Seb	ISAK	X	S

Although the codes vary, they refer to the same criteria since the rubrics are similar. The mark for each level and each criterion in the different languages is standard. Illustrations of rubric feedback presented to learners in Sesotho HL as discussed in Sibeko (2015) and English FAL are presented in Figure 3 following:

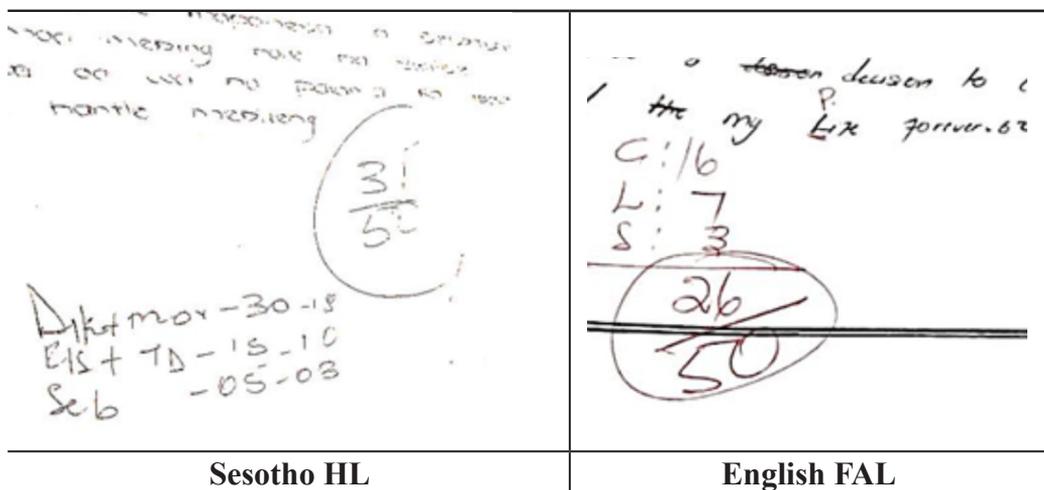


Figure 3: Examples of rubric feedback provided to learners

Final marks provided by use of the codes from rubrics in table 4 and illustrated in Figure 3 creates a problem for learners as they cannot be sure where their essays lacked and how they could have been improved. This is typical of holistic rubric feedback (Weigle, 2007:20). The problem with the overall score that does not indicate learners' weak and strong points in essay writing is that it neglects accounting for learners who perform at different levels for different rubric criteria (Weir, 2005:189; Hattingh, 2009:133-4). To illustrate, two essays may be graded at 50% in the category of content and planning. One essay might have presented outstanding details and lacked in the organisation of the content, and another might have been less relevant to the topic, but well organised. Unfortunately, the learners would be misled into thinking that they produced essays of similar quality because it would not be clear which components influenced the mark in each essay. In this way, learners cannot use the feedback to improve their subsequent creative writing essays.

Learners would benefit from more specific feedback than is provided through the use of the rubric feedback codes such as those in figure 3. Although rubric feedback is only provided on the final submission, if used correctly, learners can benefit through revising their internalised rules of language (Louw, 2006:63; Louw, 2009:89). To achieve this benefit, the rubric feedback would need to be clear enough for learners, which is not the case since the analytical rubric benefits teachers through guiding their marking while learners only receive holistic feedback. Furthermore, issues of cohesion and creativity remain unspecified when rubric feedback is summarised thereby causing the rubric feedback to be unusable beyond presenting marks. Moreover, this manner of awarding marks through summarised rubric-feedback based codes allows teachers to subjectively penalise and award learner performance. This is unavoidable since it is not clear what they award and penalise learners' essays for. This is cause for concern because in some cases, teachers do not regard the same issues as the most important ones when marking essays (Sibeko, 2015:79). There is an inherent need to standardise the marking of creative essays.

More teacher training, especially for teachers exempted from attending training sessions at marking centres, and different approaches to teaching essay marking need to be employed (Wolfe et al., 1998:485) for standardisation. According to Hamp-Lyons (2007:5) reliance cannot be placed on grades as indicators of standardisation as training sessions may not always provide clarity to teachers but result in confusion and frustration even if marks seem standardised which might be the case when teachers grant average marks for most essays. To this end, training teachers on marking creative writing essays needs to not only focus on achieving similar grades but also produce like-mindedness in understanding what is important in essays. Some attempts have been made to achieve like-mindedness in the interpretation of rubric criteria. For instance, Hattingh (2009) provides a guide for her proposed scoring rubric for grade 12 English FAL. Unfortunately, the guide is unusable in the FET band given that she suggests a scoring rubric different to the ones currently used in the FET band. Sibeko (2017) fashioned after Hattingh's (2009) guide, provides a rubric guide for Sesotho home language essays, thereby attempting to influence a standard understanding of

the rubric. However, there is no application of the rubric guide on actual learner essays to exemplify and validate the use thereof.

It may be deduced that teachers' contentment with the rubric over the correction code provides them with freedom to award marks without accounting for them by unintentionally opening a gap for teachers to nurse their biases. The lack of feedback on rubric items during the process of writing escalates the problem further as teachers are forced to focus on and provide feedback on minor matters and make their judgements on major issues at the end. This inconsistency of the type of feedback provided hampers the successful execution of assessment decisions and transparency on the part of teachers as there is no clear link between correction codes and rubric marks. Interestingly, Hattingh (2009:166) reports that markers of Grade 12 final examinations grant marks around 50 - 55% for English FAL creative essay writing, only 5% higher than Sibeko's (2015:103) finding that marks cluster around 60% for Sesotho HL grade 10 creative essays. This trend of teachers awarding average marks can be interpreted as their attempt to avoid making judgments and accounting for high achieving and low achieving essays. Unfortunately, even the process of moderation of essay marking where either a fellow or senior teacher re-marks already marked essays using a differently coloured pen to ensure marking credibility does not help (Partington, 1994:57). To this end, Partington suggests double blind-marking. However given the issues of labour intensity and time consumption, Partington's solution is impractical in the FET band in most South African high schools.

Summary

This paper set out to critically discuss the current practice followed in the assessment of creative essays with special focus on the use of the rubric and the correction code. First, the current analytic rubric used by teachers was argued to be inefficient as it is not specific to essay type. Second, the rubric feedback was argued as holistic in practice based on the actual feedback presented to learners (c.f. Table 4). Third, a mismatch between the formative feedback (using only the correction code) and summative feedback (using both the correction code and the rubric) presented on learners' creative essay tasks was highlighted. The correction code highlights grammar issues during essay writing. Considering Louw's (2008:83) contention that feedback is more than simply identifying errors which is the aim of the correction code, the lack of rubric feedback on drafts means that, in essence, no feedback is given during writing. The rubric looking into creativity, content and planning is only used during grading the final submission. A gap was thus highlighted between the different foci of the two tools used to feedback to learners. As such, the exclusion of the rubric in providing feedback during writing stages and the summarised rubric feedback are challenged as they do not help learners improve their essays (McKenna, 2007:25).

A few recommendations can be deduced from this brief discussion. First, since the rubric and the correction code clearly have divergent foci, they should not be separated

but be provided on all drafts submitted. This would ensure that macro issues that affect marks are highlighted for learners before they submit their final drafts and enable them to track their own performance. Second, more transparency in grading is needed. Learners need access to the analytic rubric feedback that their grades are based on. As such, rubric feedback more specific than the one indicated through codes such as those in table 4 might prove beneficial in helping learners improve their essay writing skills. Third, Soiferman (2017) argues a case for moving from mass production to improvement of creative writing. That is, the teacher should work with a learner until they get it right instead of assigning more creative writing tasks. This would entail that the concept of process writing in the FET band is revisited and reinterpreted.

Finally, future studies might benefit from investigating teacher training on creative writing essay marking. The most intensive trainings for teachers on marking creative essays are conducted at the grade 12 marking centres for paper 3. An investigation of how the training takes place, its benefits and outcomes and the longevity of the outcomes needs to be carried out. Furthermore, training for other teachers who do not attend to marking centres like grade 10 language teachers needs to be investigated and documented.

References

- Andrade, H.L., Du, Y. & Wang, X. 2008. Putting Rubrics to the Test: The Effect of a Model, Criteria Generation, and Rubric-Referenced Self-Assessment on Elementary School Students' Writing. *National Council on Measurement in Education. Educational Measurement: Issues and Practice*, 27(2):3–13.
- Bacha, N. 2001. Writing evaluation: what can analytic versus holistic essay scoring tell us? *System*, 29(2001):371–383.
- Borg, S. 1998. Teachers' pedagogical systems and grammar teaching: a qualitative study. *TESOL Quarterly*, 32(1):9–38.
- Brennan, R. 1996. Generalizability of performance assessments. In Phillips, G. Ed. 1996. *Technical issues in large-scale performance assessment*. Washington, DC: National Center for Education Statistics.
- Brown, G., Bull, J. & Pendlebury, M. 1997. *Assessing Student learning in higher education*. New York: Routledge.
- Brown, S. 2004. Assessment for Learning. *Learning and Teaching in Higher Education*, 1(1):81–89.

- Carr, N. 2000. A comparison of the effects of analytic and holistic composition in the context of composition tests. *Issues in Applied Linguistics*, 11(2):207–241.
- Chappuis, S., Commodore, C. & Stiggins, R. 2010. *Assessment balance and quality: an action guide for school leaders*. 3rd ed. USA: Pearson.
- Cumming, A., Kantor, R. & Powers, D.E. 2001. *Scoring TOEFL essays and TOEFL 2000 prototype writing tasks: An investigation into raters' decision making, and development of a preliminary analytic framework (TOEFL Monograph Series)*. Princeton, NJ: Educational Testing Service.
- Cumming, A., Kantor, R. & Powers, D.E. 2002. Decision-making while rating ESL/EFL writing tasks: A descriptive framework. *Modern Language Journal*, 86(1):67–96.
- Davies, A., Brown, A., Elder, C., Lumley, K. & McNamara, T. 1999. *Dictionary of language testing*. Studies in Language Testing Series, 7. Cambridge: Cambridge University Press/ UCLES.
- Davies, P. 2004. Don't write, just mark: the validity of assessing student ability via their computerized peer-marking of an essay rather than their creation of an essay. *ALT-J, Research in Learning Technology*, 12(3):261–217.
- DBE. See South Africa.
- Deremer, M.L. 1998. Writing assessment: Raters' elaboration of the rating task. *Assessing Writing*, 5(1):7–29.
- Dipardo, A., Storms, BA. & Selland, M. 2011. Seeing voices: Assessing writerly stance in the NWP Analytic Writing Continuum. *Assessing Writing*, 16(3):170–188.
- Dukes, R. L. & Albanesi, H. 2013. Seeing red: quality of an essay, colour of the grading pen, and student reactions to the grading process. *Social Science Journal*, 50(1):96–100.
- Fry, H., Ketteridge, S. & Marshall, S. 2009. *A Handbook for Teaching and Learning in Higher Education Enhancing Academic Practice*. 3rd ed. New York: Routledge.
- Fulcher, G. & Davidson, F. 2007. *Language testing and assessment: an advanced resource book*. Abingdon, Oxon: Routledge.
- Gilliland, B. 2015. High school teacher perspectives and practices: second language writing and language development. *Language & Education: An International Journal*, 29(4):287–301.

- Hamp-Lyons, L. 1991. The writer's knowledge and our knowledge of the writer. In: Hamp-Lyons, L. Ed. *Assessing second language writing in academic contexts*. Norwood, NJ: Ablex. pp. 51–56.
- Hamp-Lyons, L. 2007. Worrying about rating. Editorial. *Assessing Writing*, 1(12):1–9.
- Hattingh, K. 2009. *The Validation of a Rating Scale for the Assessment of Compositions in ESL*. Potchefstroom [Thesis - PhD Thesis]. North-West University.
- Hounsell, D. 1995. Marking and commenting on essays. *Tutoring and demonstrating: a handbook*. Available: <http://www.tla.ed.ac.uk/resources/tut-dem/Chap6.pdf> [Accessed: 27 May. 2020].
- Huot, B.A. 1990. Reliability, validity and holistic scoring: What we know and what we need to know. *College Composition and Communication*, 41(2):201–213.
- Hyland, F. 2003. Focusing on form: student engagement with teacher feedback. *System*, 31(2):217–230.
- Hyland, K. & Anan, E. 2006. Teachers' perceptions of error: The effects of first language and experience. *Science Direct System*, 34(4):509–519.
- James, E.A. 2007. Qualitative data collection. In James, E., Milenkiewicz, M.T. & Bucknam, A.J. Eds. 2007. *Participatory action research for educational leadership: using data-driven decision making to improve schools*. London: Sage. pp. 65–85.
- Jonsson, A. & Svingby, G. 2007. The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2):130–144.
- Kohn, A. 2006. Speaking My Mind: The Trouble with Rubrics. *English Journal*, 95(4):12–15.
- Louw, H. 2006. *Standardizing written feedback on L2 student writing*. Masters Dissertation. North-West University.
- Louw, H. 2009. Moving to more than editing: A checklist for effective feedback. *Journal for Language Teaching*, 43(2):86–100.
- Lockett, K. & Sutherland, L. 2000. *Assessment practices that improve teaching and learning*. Johannesburg: Witswaterstrand University Press.
- Lumley, T. 2002. Assessment criteria in a large-scale writing test: What do they really mean to the raters? *Language Testing*, 19(4):246–276.

- Marzano, R.J. 2002. A comparison of selected methods of scoring classroom assessments. *Applied Measurement in Education*, 15(3):249–267.
- McKenna, S. 2007. Assessment rubrics as a means of making academic expectations explicit. *The Journal of Independent Teaching and Learning*, 2(1):22–30.
- McNamara, T. 2000. *Language Testing*. New York: Oxford University Press.
- Partington, J. 1994. Double-marking students' work. *Assessment & Evaluation in Higher Education*, 19(1):57–60.
- Shaw, S.D. & Weir, C. 2007. *Examining writing: research and practice in assessing second language writing*. Vol 26. Cambridge: Cambridge University Press.
- Shaw, S.D. 2002. The effect of training and standardization on rater judgement and inter-rater reliability. *Research Notes*, 8(1):13–17.
- Sibeko, J. 2015. The use of rubrics and correction codes in the marking of Grade 10 Sesotho home language creative writing essays. Masters Dissertation. North-West University.
- Sibeko, J. 2017. A bilingual English-Sesotho rubric explanation guide for the marking of Sesotho home language creative writing essays. *Journal of Language Teaching*, 51(1):69–93.
- Soiferman, L.K. 2017. *Teaching High School Students How to Write: The Importance of Direct Explicit Instruction and Teacher Training*. Available: <https://eric.ed.gov/?id=ED572369> [Accessed: 15 January 2020].
- South Africa. Department of Basic Education. 2008b. *National Curriculum Statement Grades 10-12 (General): Subject Assessment Guidelines*. Government printing works. Pretoria.
- South Africa. Department of Basic Education. 2011a. *English first additional language. Further Education and Training Phase Grades 10–12: Curriculum and Assessment Policy Statement (CAPS)*. Government printing works: Pretoria.
- South Africa. Department of Basic Education. 2011b. *English home language. Further Education and Training Phase Grades 10–12: Curriculum and Assessment Policy Statement (CAPS)*. Government printing works: Pretoria.
- South Africa. Department of Basic Education. 2011c. *SeSotho puo ya lapeng. Mokgahlelo wa Thuto le Thupelo eTswelang Pele Dikereiti tsa 10–12: Setatementsa sa Leano la Kharikhulamu le Tekanyetso (SLKT)*. Government printing works: Pretoria.

- South Africa. Department of Basic Education. 2017. Grade 10 Exams. Available: <https://www.education.gov.za/LinkClick.aspx?fileticket=MHlurzCsx5w%3d&tabid=1853&portalid=0&mid=7376&forcedownload=true> [Accessed: 30 May 2020].
- South Africa. Department of Basic Education. 2018. November Past Papers. Available: <https://www.education.gov.za/LinkClick.aspx?fileticket=DZxgfB37GVM%3d&tabid=2268&portalid=0&mid=8373&forcedownload=true> [Accessed: 17 August 2019].
- Spencer, B. & Louw, H. 2008. A practice-based evaluation of an on-line writing evaluation system: First-World technology in a Third-World teaching context. *Language Matters*, 39(1):111–125.
- Spencer, B. 1998. Responding to student writing: strategies for a distance-teaching context. PhD thesis. University of South Africa. (Unpublished).
- Spencer, B. 2009. A critique of response strategies: measures to induce a paradigmatic shift in response to student writing. *Per Linguam*, 25(2):17–34.
- Spencer, B. 2012. Responding to student writing: the students' perspective. *Southern African journal of Applied Language studies*, 6(1):27–42.
- Thompson, T.C. 1995. Understanding attitudes toward assessment: The personality factor. *Assessing Writing*, 2(2):191–206.
- Van der Walt, C. 2010. Of shoes-and ships-and sealing-wax: A dynamic systems approach to language curriculum orientation. *Southern African Linguistics and Applied Language Studies*, 28(4):323–337.
- Weigle, S.C. 2002. *Assessing writing*. Cambridge: Cambridge University Press.
- Weigle, S.C. 2007. Teaching writing teachers about assessment. *Journal of Second Language Writing*, 16(3):194–209.
- Weinberger, A., Dreher, H., Al-Smadi, M. & Guetl, C. 2011. Analytical assessment rubrics to facilitate semi-automated Essay grading and feedback provision. In Yorke, J.D. Ed. 2011. *Meeting the Challenges. Proceedings of the ATN Assessment Conference 2011, 20-21 October*. Perth, Western Australia: Curtin University. pp. 170–177.
- Weir, C.J. 2005. *Language testing and validation: an evidence-based approach*. Oxford: Palgrave Macmillan.
- Whipp, P. 2011. University assessment practices at level 1: Exploring student perceptions of fairness, transparency and authenticity. In Yorke, J.D. Ed. 2011. *Meeting the Challenges. Proceedings of the ATN Assessment Conference 2011, 20-21 October*. Perth, Western Australia: Curtin University. pp. 161–169.

- Whithaus, C. 2005. Teaching and evaluating writing in the age of computers and high-stakes testing. London: Taylor and Francis.
- Withers, G. 1987. From marking strategy to assessment procedure: A review of recent Australian practices. *Studies in educational evaluation*, 13(1):7–19.
- Wolfe, E., Kaokao, C.W. & Ranney, M. 1998. Cognitive difference sin proficient and non-proficient essay scorers. *Written Communication*, 15(4):469–482.
- Yürekli, A. & Üstünlüoğlu, E. 2007. Towards Student Involvement in Essay Assessment. *Essays in Education*, 22(2):55–64.

ABOUT THE AUTHOR

Sibeko Johannes

Nelson Mandela University, North Campus, University Way, Summerstrand,
Port Elizabeth, 6019
Email: johannes.sibeko@mandela.ac.za

Johannes Sibeko is lecturer of Digital Humanities at Nelson Mandela University. He worked as a high school teacher for three years before joining higher education. His special interests as a teacher of English first additional language and Sesotho home language were creative writing and translanguaging. His previous research has thus been on creative essay writing in high schools. He is currently completing his PhD with North-West University Vaal Tringle Campus.