Tydskrif vir Taalonderrig - Journal for Language Teaching
- Ijenali yokuFundisa iLimi - IJenali yokuFundisa iiLwimi -
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta Polelo
- Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo - Ijenali
Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -
Journal for Language Teaching - Ijenali yokuFundisa iLimi
- IJenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa
Ririmi - Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali yokuFundisa
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo
- Jenala yo Dyondzisa Ririmi
- Tydskrif vir Taalonderrig
- Journal for Language
Teaching - Ijenali
yokuFundisa iLimi -
IJenali yokuFundisa
iiLwimi - Ibhuku
Lokufundisa Ulimi
- Tšenale ya tša
Go ruta Polelo -
Buka ya Thuto
ya Puo - Jenale
ya Thuto ya Dipuo
- Ijenali Yekufundzisa
Lulwimi - Jena?a ya u
Gudisa Nyambo - Jenala   yo
Dyondzisa Ririmi - Tydskrif vir Taalonderrig
- Journal for Language Teaching - Ijenali
yokuFundisa iLimi - IJenali yokuFundisa iiLwimi -
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya
Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a ya
u Gudisa Nyambo - Jenala yo Dyondzisa Ririmi
- Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali
yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi -
Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa
Ririmi - Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali yokuFundisa
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -
Journal for Language Teaching - Ijenali yokuFundisa iLimi
- IJenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa
Ririmi - - Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali yokuFundisa
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta

Sharita Bharuthram
University of the Western Cape

Evaluation of assessment skills using essay rubrics in student self-grading at first year level in higher education: a case study

Abstract

This paper reports on a study in which students self-graded an assessment task with the aid of an assessment rubric. On comparing student self-grades with those of the tutor it was found that majority (72.6%) of the students failed to demonstrate good self-assessment skills with student grades ranging from 25 less than the tutor grade to an overestimation of 36 above the tutor grade. Consistent with other research studies the results further show that weaker students graded themselves higher than the better performing students. However, there was a tendency for male students to grade themselves higher than female students. Analysis of the qualitative data reveals that students' feelings about self-grading may play a role in the grades they assign themselves. The author therefore recommends that affective factors need to considered and addressed prior to the self-grading task. Furthermore, it is evident that rubrics alone may not necessarily improve self-grading and that internalisation of the rubric criteria and standards, as well as practice is crucial. Finally, in order to produce graduates who are able to appraise their performance, self-assessment should be embedded early in the students degree programme and be sustained throughout the degree.

Key words: grading, self-assessment, learner-centred assessment, higher education, assessment rubric

1. Introduction

Assessment plays a crucial role in education as it is through assessment methods, be it tests or examinations that students are allowed to progress from one level to the next. Although the role of assessment and assessment methods have been the subject of debate for a number of years, the consensus among researchers is the move towards learner-centred assessment methods coupled with the use of explicit assessment criteria for marking, both of which should be clearly communicated to students with the aim of enabling students to develop competencies in the real world (Pereira, Flores & Nicklasson, 2016). Self and peer-assessment are methods that promote student-centred learning and have the potential to help develop students into independent and lifelong learners, and as practitioners who would be able to reflect critically on their own professional domains – characteristics that are key goals of higher education (Sambell & McDowell, 1997; Dochy, Segers & Sluijsmans, 1999). Yet, most often academics are reluctant to engage students in self-assessment processes especially at first year level. The general view seems to be that first year students are not able to make fair and appropriate judgements on their own work (Nulty, 2011), the assumption being that asking these students to self-assess would be a wasteful exercise. Contrary to this view, research studies suggest that students should receive practice in self-assessing from the very start of their degree programmes (Boud, 1995; Nulty, 2011).

Despite the controversy surrounding the grading[1] of assessment tasks, it still remains a common practice and in most instances the only way in which a student's advancement in an academic programme is determined. Hence, it makes sense that students be taught how to grade their own work so that they can use their self-assessment to improve their work which may result in better quality output thereby enhancing their chances of advancement in their degree programmes. Consequently, this article discusses a self-assessment practice implemented at the first year university level whereby students graded their own work with the aid of an assessment rubric and rubric checklist. While much has been written globally on the use of self and peer-assessment to assign grades to students' work (see for example, Langan, Shuker, Cullen, Penney, Preziosi & Wheater, 2008; Lew, Alwis & Schmidt, 2010; Alias, Masek & Salleh, 2015), published related research within the South African context is scarce and almost non-existent. Furthermore, results of studies that correlate students' assessment grades (self and peer) and teacher assessments have been inconclusive (Alias *et al*., 2015). The current study is therefore an attempt to fill the gap in the South African literature while also contributing to the international discussion on the correlation between tutor and student self-assessment. The study is extended further by exploring the role, if any, that gender plays in self-grading, and by correlating the student self-grade with the student continuous assessment (CA) grade, in an attempt to understand whether those students who made judgements of their work that differed significantly from the grading of the tutor can be classified as poor academic performers. Additionally, students' feelings about grading themselves are explored.

---

1   In this article grading refers to the assigning of a mark i.e. a numeric value.

2.   Literature review

Self-assessment is a learner-centred assessment practice that cannot be separated from any assessment that is aimed at improving learning (Wiggins, 1998).   Although self-assessment is typically combined with peer-assessment, the focus in the current article is only on self-assessment as it is based on the premise that students should first learn how to assess themselves and become confident in making judgements about their own work before attempting to assess others. Many research studies have discussed the benefits from involving students in self-assessment practices (Anderson & Freiberg, 1995; Longhurst & Norton, 1997; Dochy *et al*., 1999; Sivan, 2000; Orsmond & Merry, 2012).   Some studies have shown that self-assessment promotes deep learning (Cowan, 1988; Mok *et al*., 2006; Kirby & Downs, 2007; Brew, Riley & Walta, 2009), fosters a sense of responsibility in students, enhances a reflective practice, and develops students' self-confidence and autonomy (Anderson & Freiberg, 1995; Sambell & McDowell, 1997; Sivan, 2000).   Self-assessment is also known to foster the development of a variety of skills, such as, listening skills, writing skills and lifelong learning skills (Falchikov, 1995; Sambell & McDowell, 1997; Challis, 1999). As such, it is important that self-assessment be embedded as an integral part of the teaching and learning process throughout the degree programme.

According to Boud and Falchikov (1989) self-assessment may be formative or summative and should encompass two key elements, namely, a set of standards or criteria that are applied to the assessment task, and making judgements about the extent to which these standards have been met.   In line with these key elements Boud (1991:5) defines self-assessment as "the involvement of students in identifying standards and/or criteria to apply to their work, and making judgements about the extent to which they have met these criteria and standards". Many researchers studies (Dochy *et al*., 1999; Panadero, Alonso-Tapia & Reche, 2013; Jonsson, 2014) advocate towards the use of rubrics to identify standards and criteria and in guiding the assessment process while also strongly making a case for the rubric as a formative self-assessment tool which should be used throughout the completion of a particular task. The use of self-assessment, however, as a summative process which involves students grading their own work is a very contentious issue and has been less favourably received (see for example Andrade, 2007/2008; Kohn, 2011). Nonetheless, Sadler (2010) argues that self-assessment should also been seen as a strategy that develops students' evaluative skills. In addition, he suggests that students should be provided with appraisal experiences similar to those of their teachers. Taking cognisance of the importance of self-assessment for both formative and summative purposes (as discussed above), the design of the rubric and rubric checklist used in the current study catered for both. Furthermore, in lieu of the fact that grading still remains central in the assessment process it is important for students to be taught how to grade their own work.

2.1. Self-assessment and grading

Despite grading being a widely contentious issue it has been "the mainstay and bed-rock of education" (Taras, 2015:5) representing achievement for academic success (Broadfoot & Black, 2004). In making a case against grades, Kohn (2011:28-29) argued that "[g]rades tend to diminish students' interest in whatever they're learning"; "grades create a preference for the easiest possible task"; and "grades tend to reduce the quality of students' thinking". Other studies have also reported negatively on the use of grades (see for example, Butler, 1987; Anderman & Murdock, 2007; Pulfrey, Buch & Butera, 2011). Such studies have resulted in the call for grades to be excluded from the teaching and learning process. However, it is argued in this paper that since grades form the "foundation for the certification of learning" (Sadler, 2009:159) and still remains the norm in the South African higher education system, it would be irresponsible not to include it as part of the self-assessment process. Furthermore, grades may have a role in helping students understand the expected standards of a task (Sadler, 1989). Both Sadler (1989) and Taras (2015) concur that the timing of the grade is important. Taras (2015:5) suggests that "in order to prevent interference with understanding and take up of feedback students should receive their grade only after the pedagogic cycle of discussion with peers and tutors". By implication, students should self-grade only after going through the formative process.

Research studies that compared tutor and student grades have been inconclusive with some studies reporting a high correlation between tutor and student self-grades and others reporting little or no correlation. For example, in a study with seventh grade students, Sadler and Good (2006) compared grades given by the tutors with the grades students gave themselves and their peers. The authors reported a high correlation between the tutor and the student self-grade. Similar findings were reported with students in the higher education context (see for example, Boud, 1989; Longhurst & Norton, 1997). In contrast, studies conducted by Cassidy (2007) and Lew, Alwis and Schmidt (2010) found that students tend to underestimate their performance in comparison to their teachers and peers. In yet another study which was conducted by Alias *et al.*, (2015) at a Technical and Vocational Institute, the authors reported a correlation between self and peer-assessment scores but there was no correlation between teacher's and students' assessment scores. The authors found that students scored themselves and their peers higher than the teacher's score.   Other studies report that it is usually the lower performing students who give themselves higher grades (Longhurst & Norton, 1997; Sadler & Good, 2006; Boud, Lawson & Thompson, 2015).

Despite the contradictory findings presented above there is general agreement among the researchers of the value of student self-assessment and more specifically self-grading. For example, Sadler and Good (2006) reported that when the students who graded themselves retook the same test a week later their performance improved dramatically in comparison to students who graded their peers and those who were not involved in any grading. They concluded that self-grading resulted in increased student learning whereas peer grading does not. In a similar vein, studies have also reported that low performing students show "the greatest improvement in performance through

self-assessment" (Brown & Harris, 2013: 387) although slightly different findings were presented in a more recent study conducted by Boud, Lawson & Thompson (2015:9) who examined students according to their ability levels i.e. high, low and mid ability. The authors found that while the high ability group had significantly underestimated their grades in all of the assessment tasks taken, by the end task the gap had narrowed. Similarly, while the mid-range group results were significantly higher than that of tutors at the beginning task, by the end task there was no significant difference between themselves and the tutors. The results were very different for the low ability students who had significantly overestimated their ability in all their assessment tasks. By the end task, these students had shown no improvement in their ability to make judgements. The authors conclude that these results show that ability level has an effect on students' accuracy of judgement with the low performing students being "at risk in terms of both their academic performance and their competency to self-assess" (Boud *et al.*, 2015:52). However, other research studies have shown that accuracy in self-grading improves with practice especially in poor performing students (Syed, 2011; Brown & Harris, 2013). Nonetheless, some authors argue that accuracy is secondary – more important is the fact that the self-assessment supports learning.

3. Research methodology

The overall aim of this research is to help students develop the capacity to make judgements of their own work thereby enabling them to ultimately be in a position to reflect critically on their performance outside of academia and in becoming independent lifelong learners and practitioners. Consequently, students made use of an assessment rubric and a rubric checklist to assign a grade to their essay which was then compared to the tutor grade and to their overall course work grade.

As such, this research attempts to address the following questions:

1. Is there a correlation between the student grade and the tutor grade?

2. Do the grades of males or females better align with that of the tutor?

3. Is there a correlation between the student grade and their overall course work grade?

4. How do students feel about grading themselves?

In order to answer the above questions both quantitative and qualitative research designs were used. While a quantitative research design is used to determine the relationship between variables, researchers (see for example Denzin & Lincoln, 1998) are in agreement that qualitative research allows for the examination of opinions, beliefs and emotions of people in a particular setting. Therefore, the study relied on the quantitative research design to address questions 1 to 3 above and the qualitative research design to understand students' feelings in relation to the self-assessment task.

3.1. Context

This research was conducted at a university in South Africa and in a module called English for Educational Development (EED) which is offered to students from the Faculty of Community and Health Sciences (CHS). The EED-CHS module is an academic literacies module which focuses on developing the students' disciplinary literacy practices. The process approach to writing is adopted whereby students go through the various stages of drafts and re-drafts of their essays and on which they get constructive and developmental feedback before the submission of their final essay. In addition, students are given assessment rubrics for all 4 of their major assessment tasks which make up the students' continuous assessment grade. These rubrics are explained to students in detail before the commencement of a particular assignment. In particular, students are given an explanation of the different criteria on the rubric and are shown how to use these criteria to assess their work.

One of the 4 assessment tasks which together make up the students continuous assessment (CA) grade for the semester, is the writing of an argumentative essay. For this essay, in addition to the rubric students are given a rubric checklist which they use to assess themselves prior to the submission of their final essay i.e. at the end of the essay writing process cycle. While serving as a tool to reinforce the rubric, the rubric checklist is different from the rubric in that it requires students to indicate whether they have met the criteria listed in the rubric and show evidence of where in their essay they have met the criteria (for further explanation of the original design of the rubric and rubric checklist see Bharuthram and Patel, 2017[2]). The checklist also makes provision for students to grade their essay out of 100. Students are required to attach both the rubric and the rubric checklist to their final essay submission. The tutors provide their grade and feedback on the rubric itself.

Prior to commencement of the academic semester tutors undergo a training session which includes amongst others a session on providing feedback on students essays. In addition, every assessment task involves a committee marking session which is led by the co-ordinator of the course who discusses the rubric. Thereafter, a student essay is assessed by the tutors and the co-ordinator in the form of written feedback and a grade. This is followed by a discussion on the type of feedback given, the grade allocated, and the reasons for allocating the grade. These committee meetings are very useful for a number of reasons: tutors get a better understanding of the assessment criteria and how to use the rubric to assess students, the discussions help develop tutors expertise further in providing constructive and developmental feedback on students essays so as to enhance students' higher-order thinking, and it also ensures that the tutors and the co-ordinator are on par with the allocation of the grades. To ensure standardisation of grades across the different tutorial groups all assessment tasks are moderated by the co-ordinator before they are returned to students.

---

2   Of note, in the current study the students did not participate in the design of the rubric or the rubric checklist. These were designed by a different cohort of students. These instruments are not the main focus of discussion in this article.

3.2.   Participants

The participants in this research were a total of 164 students who were registered in 2016 for a compulsory one semester EED course. These students came from various Community and Health Sciences disciplines. Their ages ranged from 18 to 22 years and most of them were English Additional Language (EAL) speakers. In addition to the students, 5 tutors who are post-graduate students participated in this research. With the exception of one tutor, all the others were experienced tutors in that they were tutoring in the programme for two years and more.

3.3.   Data collection instruments

The rubric and rubric checklist formed the primary source of data collection.   The quantitative data was in the form of the grade given on the assignment by the tutor on the rubric and by the grade given by the student on the rubric checklist. Qualitative data was collected from a subgroup of the 164 students i.e. from three different tutorial groups comprising a total of 48 students. They were asked to write a paragraph in response to the following open-ended question: How do you feel about grading your own work? Students completed this task anonymously.

3.4.   Data analysis

The data was analysed and interpreted in keeping with the aims of the research. First, in an attempt to answer the first 3 questions posed above, descriptive statistics for the cohort of 164 students were calculated to establish means, standard deviations (SD) and range of the student, tutor and the continuous assessment grades. Correlations between the student grade and the tutor grade, and student grade and the continuous assessment grade were examined. Differences between the student grade and the tutor grade for each student were calculated to see how student grades varied from tutor grades, and were also used to assess whether gender influenced self-assessment using the non-parametric Mann-Whitney U Test. To investigate performance, students were stratified based on whether their continuous assessment grade fell below/on/ above the class average. All statistical analyses were performed using GraphPad Prism version 5.03 for Windows (GraphPad Software, San Diego California USA). Statistical significance levels were set at $p=0.05$.

Next, the reasons provided by students on how they felt about grading their work were analysed. This involved an initial reading of all the reasons provided by students to get an overall sense of what students were saying. In a second reading an attempt was made to categorise student responses according to negative and positive responses. However, in many cases a clear distinction could not be made as there were many overlaps (i.e. students reported both positive and negative feelings).   Thereafter, it was decided to simply highlight and record key words and or phrases to ascertain the emergence of any significant patterns in student responses in relation to the research

question. Similar methods of qualitative analysis were used by other researcher (see for example, Crossman, 2007).

4. Findings and discussion

4.1. Is there a correlation between student grade and the tutor grade?

The overall difference between self-assessed student grades and tutor grades out of a 100, in a cohort of 164 students was strongly significant (Figure 1, $p<0.001$). The student grades were higher than the tutor assigned grades, with average values of 69 (SD = 9) and 61 (SD = 9) respectively. Although the range of the student and tutor grades were comparative, 40 and 39 respectively, the minimum student grade was 50, which was 12 grades higher than the minimum tutor grade of 38. Interestingly, none of the students assigned themselves a grade below 50%, which would constitute a fail. The maximum student grade and tutor grades were 90 and 77 respectively.
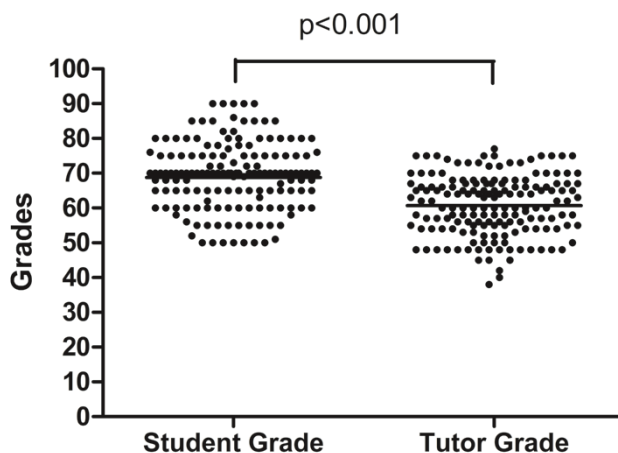
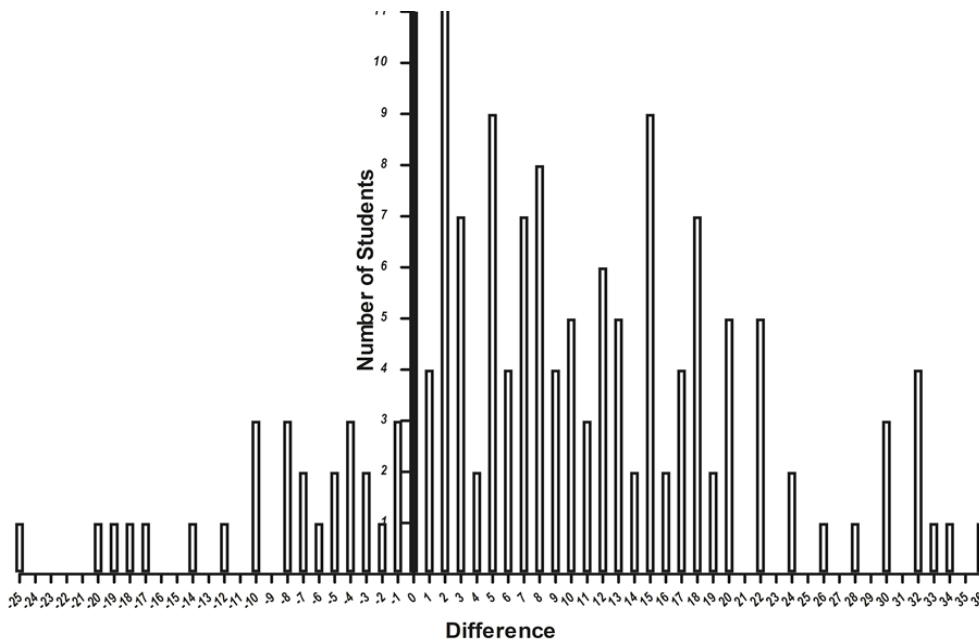*Figure 1. Difference between self-assessed student grades and tutor grades*

In Figure 2 below, we note that although the student grades were higher than the tutor grades, there was a weak positive correlation between the student grade and the tutor grade (Figure 2, $p = 0.008$, correlation coefficient $r = 0.21$). These results tend to concur with the findings of other researchers. For instance, Cassidy (2007) reported a significant difference in the way teachers and students graded themselves on the same task; finding a significant positive but fairly low correlation of 0.25 between tutor and student grades.
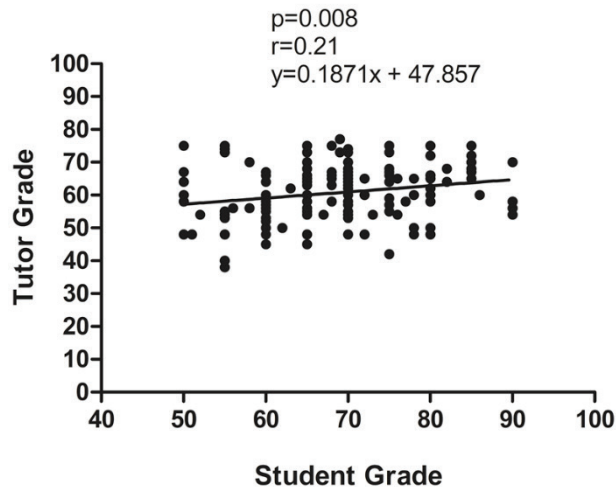
Figure 2. Correlation between self-assessed student grades with tutor grades.

P and r values, as well as the best fit line equation are indicated.

While the results from the current study indicate that there was a relationship between the student grades and tutor grades, these results do not provide a detailed picture of the similarities and differences between tutor grades and student-self grades. Figure 3 represents the range of differences for each student as well as the frequency of the difference values.

Figure 3. *Range of difference between self-assessed student and tutor grades for the same assignment and the frequency of students per calculated difference value. Point zero is where the student grade equalled the tutor grade.*

It can be seen from the figure that student grades ranged from 25 less than the tutor grade to an overestimation of 36 above the tutor assigned grade. Eleven students out of 164 were able to match the tutor grade. As reported above, it can be clearly seen that student grades were generally higher than the tutor grades suggesting that students were more likely to overestimate the quality of their work rather than underestimate. It is seen that 77% (127) of the students overestimated their grades while 16% (26) underestimated. These results contradict the results of Cassidy (2007) who found that 56% of the students underestimated their grades and 40% overestimated them.

Table 1 below provides a further breakdown into student grades that are higher and those lower than the tutor grades. A total of 45 students (27.4%) had a grade difference of less than 5 in comparison to the tutor grade. Since this was their first attempt at self-grading in the EED class a difference of less than 5 was considered as a negligible difference. This means that 72.6% of the class were not able to adequately self-assess. Hence, one can conclude that the majority of students failed to demonstrate good self-assessment skills despite the use of the rubric and the rubric checklist that were provided as guiding tools. It is possible that students at that time may not have internalised the criteria and standards set in the rubric/rubric checklist. Students' active participation in the rubric

11

design may have yielded different results. It could also be conjectured that for many of the participants this was their first experience at self-grading - with or without the use of a rubric - and there is a possibility that given more practice in making self-judgements, improvement may be noticeable. Research studies (see for example Boud *et al*., 2015) do show that while initially students struggle to accurately self-assess, their accuracy improves over time. However, Boud *et al*., (2015:17) also report that the assessment type and the assessment criteria have a role to play in the convergence of student and tutor grades.

Table 1. Student grades higher and lower than the tutor grades

Student grade more than tutor grade		Student grade lower than tutor grade	
Difference	Number of students	Difference	Number of students
30 – 36	10	-	-
20 – 29	14	20 – 25	2
10 – 19	35	10 – 19	8
6 - 9	23	6 - 9	6
1 - 5	34	1 - 5	11
Total	126	Total	27

To further explore the factors that may contribute to the over or underestimation of the grades allocated by the students, the cohort was divided into males and females.

4.2.   Do the grades of males or females better align with that of the tutor grade?

Of the 164 students, 101 (61.6%) were females and 63 (38.4%) were males. On initial analysis, no significant differences were observed in the calculated difference values between males and females, that is, gender did not influence the self-assessment grade relative to the grades calculated by tutors (Figure 4 below, panel A, $p = 0.052$). However, in the male cohort it was observed that there was a single data point that seemed to be skewing results; this outlier is marked in figure 4, panel A. Upon reanalysis with the removal of the outlier (Figure 4, panel B, $p = 0.033$), the calculated difference scores were modified and found to be significantly higher in males compared to females, meaning that male students rated themselves more highly than the tutor, compared to female students. While these results are interesting, no comparisons could be made to other research studies, since to the best of my knowledge, none were conducted especially in a higher education context with a focus on the role of gender in the accuracy of student self-grades in comparison to the tutor grades.
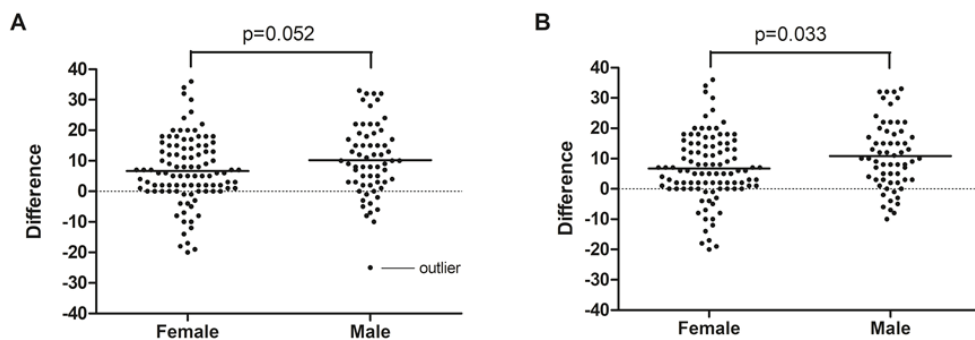
Figure 4. **Differences between self-assessed student and tutor grades in females (n=101) and males (n=63) (panel A).** *In panel B the outlier indicated in panel A is removed and the difference now becomes significant.*

On further exploration of gender differences, it was found that the tutor assessment averages for the females was 62 (SD 8) and males 60 (SD 9) and the average of the CA grades for females was 62 (SD 8) and males 59 (SD 8). Hence, once could conclude that in terms of the actual student performance on the argumentative essay as assessed by the tutor and students overall performance in the course on the basis of their CA grade, both males and females performed almost equally.

4.3.   Is there a correlation between student grades and their CA grades?

Since in the current study the students' ability levels were not assessed upon entry into university, it was decided that their CA grades (which would give one a sense of the ability level of the student) would be used to determine if there was a correlation between the student self-grade and overall performance in the course.

The correlation analysis of the student grades and their CA grades gave a significance level of $p = 0.001$ and a correlation coefficient $r = 0.240$, which indicates weak positive correlation (see Figure 5 below). As such, evidence from this analysis seems to indicate that there is a weak but significant relationship between the grades that the students gave themselves and the overall performance. To better understand the link between overall performance on self-assessment, the students CA grades were divided into below and above the class average as presented in Figure 6.
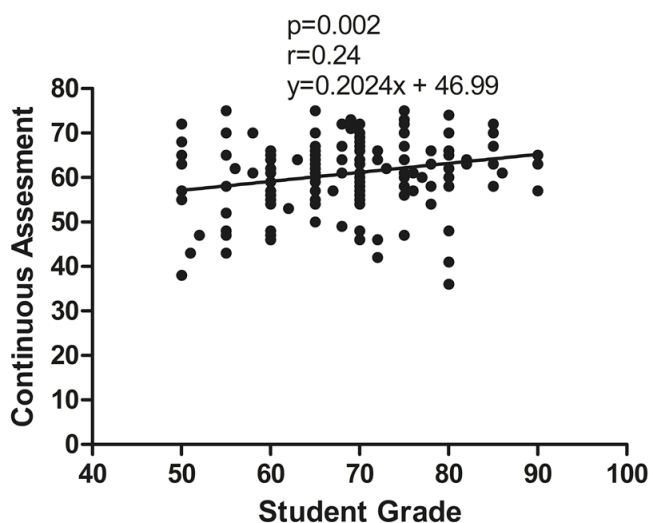
Figure 5. Correlation between self-assessed students' grades. P and r values, as well as the best fit line equation are indicated.

Following observations made by Lew et al., (2010) and Longhurst and Norton (1997), the cohort of students was stratified based on performance using the continuous assessment mean of 61% (SD = 8), with 97 students having a CA grade above or equal to the average and 67 with a CA grade below average. The difference between student grade and tutor grade for these two cohorts is shown in Figure 6 below. There was a strong significant difference in the calculated grade difference values of students who scored above and equal to the class average compared to those who scored below the class average, with those below the class average assigning higher grades to themselves compared to their counterparts (Figure 6, $p < 0.001$). From these results it can be concluded that the weaker students overestimated their performance and gave themselves higher grades thereby concurring with the results of other researchers (Falchikov, 1989; Orsmond, Merry & Reiling, 1997; Lew et al., 2010; Boud et al., 2013) who found that the student who were judged as being more academically competent were able to self-assess with higher accuracy than less competent students who had the tendency to over rate their own performance. It could be postulated here that the 'weaker' students put in a lot of effort into the task and therefore rated themselves accordingly. For these students then there was no synergy between their final products and the grades they had given themselves.

Interestingly, of the top 10 performing students in the class based on the continuous assessment grade, only one was male and was the student who was the outlier mentioned in the section reporting on gender differences above.

14
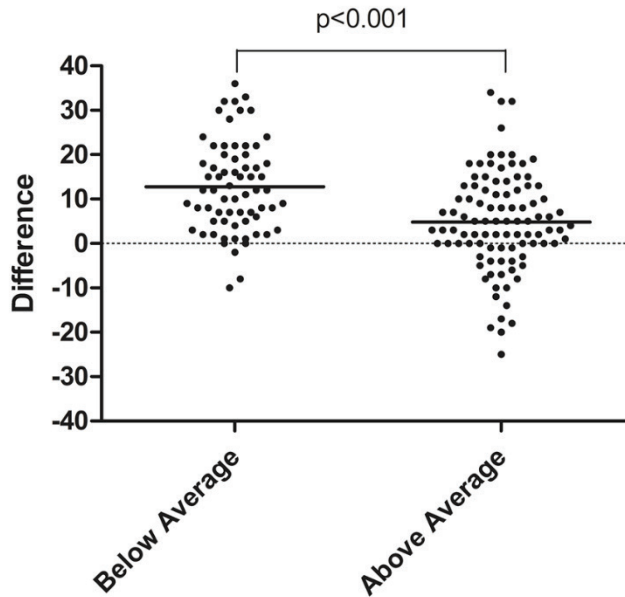
Figure 6. *Comparison of the differences in student and tutor grades in above and below average performing students. Below average performers obtained a continuous assessment grade below the class average (n=67), while above average students obtained a continuous assessment grade above or equal to the class average (n=97).*

4.4. Students' feelings about self-grading

Mixed responses were received from students on grading themselves. Some welcomed the idea: *'It made me to think of the quality I produced'; 'I felt that it was a good thing that I could grade my own work as I have pride in the work I do'; 'I felt that this is a great opportunity to express our opinion on our own worth based on the efforts we know we put into the essay'.* What emerged from these positive responses was that the task of grading themselves - in a way – made students feel included in the assessment process and this in itself could result in enhanced motivation. It also forced students to reflect on their own performance and then assess their performance. To this end, most students made reference to the effort that they put into their work. As such, it appears that students correlated performance with effort holding the view that extra effort would be deserving of a good grade. This could account for the elevated marks reported earlier. Unfortunately, students written responses to how they felt about self-grading could not be correlated with their actual grades as they responded to this question anonymously. Some students, albeit a few, alluded to making use of the rubric/rubric checklist: *'It was a way to assess our work from a different perspective. I enjoyed this'; '…we could use the guidelines to give a mark'.* The use of rubrics as a tool to support self-assessment

has been supported by researchers (see for example Johnsson, 2014) as the rubric could assist students in obtaining a better understanding of the criteria which in turn may lead to reinforcement of their self-assessment practices.

Approximately 45% of the students said that they did not enjoy grading their own work as they found the task *'unnerving'* and they felt '*unsure*', '*uncomfortable*' and '*awkward*' having to give themselves a grade. Similar findings were reported by Sher and Twigg (1991) who found that students felt apprehensive because they felt that they were not adequately trained to self or peer-assess. More specific statements received from students included 'I didn't really enjoy it - *I felt I had to do it. It was pointless'; '…my mark is not final'; 'My mark does not count...'; and '…I felt exposed'*. Of note here is that the grade provided by the student was not taken into account. Therefore, students saw it as a futile exercise and felt forced to complete the task. It also appears that some students felt that they would be adversely affected once the marker saw their grade which could have resulted in them elevating their grades. It must be noted that this observation could not be confirmed for individual students. However, it ties in with the data from the statistical analysis presented earlier in this paper which showed that students over-estimated the quality of their work.

A few students reported self-grading as *'intimidating'* as indicative by the following responses: *'I felt intimidated because I never knew exactly what I was capable of producing…I did not want to underestimate myself'*, and *'It was weird because you don't want to be too harsh but also not too lenient or else it feels like you think it's great when it might not be great. It's hard'.* It appears that these students did not feel confident in grading themselves which could possibly be as a result of their inexperience or inadequate understanding of the assessment criteria. Research (Boud, 1986) shows that novice students do struggle to rate their work in comparison to advanced students. Presumably students' confidence and accuracy will improve with practice. Taras (2015:6) holds the view that "accuracy of student grading is secondary to the learning benefits of involving students within the assessment process" however Sadler (1989) points out that accuracy could be a way to confirm students understanding of the standards and criteria laid down in the rubric.

5.   Conclusion

The purpose of the study was to gain an understanding of students ability to make judgements of their own work with the aid of a rubric and rubric checklist. Consequently, students graded themselves on an assessment task and these self-grades were correlated with the tutor grades. The data from this study supports the conclusion by other researchers of a weak correlation between the student grade and the tutor grade and a tendency for weaker students to elevate their grades. Closer analysis reveals that of the cohort of 164 students only 11 students scored the same grade as the tutor, with the majority of the students either overestimating or underestimating their grades. The conclusion drawn from the data is that majority of the students were not able to make

accurate judgements of their own work despite being given a rubric and a rubric checklist, both of which were explained to them in detail. Students were also shown how to use these tools to self-assess. Hence, the expectation was for greater alignment between student and tutor grades. This then begs the question: why did so many students still make inaccurate judgements?

The obvious conclusion is that these are novice students and one can conjecture that for many of them this was their first self-assessment experience. They therefore associated the grade they allocated themselves to the amount of effort put in, rather than the quality of the work. Research shows that novice students do have a tendency to overrate themselves compared to advanced students but they do improve with practice. It could also be that some students did not fully understand the standards and criteria laid down in the rubric and rubric checklist. To facilitate the internalisation of these tools a possible consideration could be to provide room for students to participate as co-designers. Of importance here is the point raised by Andrade and Valtcheva (2009) who argue that students need to be taught how to self-assess/evaluate which could lead to greater engagement with the rubric, which in turn, could result in deeper learning and better academic performance (Reddy & Andrade, 2010).

The qualitative data points to a possible consideration of affective factors, more specifically, students' feelings towards the task may affect their self-judgment. However, this finding should be used with caution since the results could not be explored further as students completed their responses to the open-ended question anonymously and therefore their self-grades could not be compared to their individual responses. Furthermore, the qualitative data collection involved a small cohort of students. It is possible that discussions that specifically focused on the purpose(s) of self-assessment prior to the self-grading task thereby making the task more purposeful to students, may have resulted in more positive feelings and therefore better self-judgement. Such discussions may increase students' motivation to self-assess and increased motivation levels may influence accuracy (Longhurst & Norton, 1997). Nonetheless, the affective dimension is worthy of further exploration.

This study further contributes to the research on self-assessment by exploring the role of gender in self-assessment. While it was found that the males graded themselves slightly higher than the tutor in comparison to the females, no differences were found in their performance of the essay as assessed by the tutor and their overall performance in the course on the basis of their CA grades. A study of male and females differences in self-grading over a number of tasks could yield different results and is an area that could be explored further.

Finally, to achieve the full benefits of self-assessment it should be embedded early in the students degree programme and be sustained throughout the degree.

References

Alias, M., Masek, A. & Salleh, H.H.N. 2015. Self, peer and teacher assessments in problem based learning: are they in agreements? *Procedia – Social and Behavioural Sciences*, 204 (2015): 309-317.

Andrade, H. & Valtcheva, A. 2009. Promoting learning and achievement through self-assessment. *Theory into Practice*, 48 (1): 12-19.

Anderman, E.M. & Murdock, T.B. 2007. *Psychology of academic cheating*. Burlington, MA: Elsevier Academic Press.

Anderson, J.B. & Freiberg, H.J. 1995. Using self-assessment as a reflective tool to enhance student teaching experience. *Teacher Education Quarterly*, Winter 22: 77-91.

Andrade, H. 2007/2008. Self-assessment through rubrics. *Educational Leadership*, 65 (4): 60-63.

Bharuthram, S. & Patel, M. 2017. Co-constructing a rubric checklist with first year university students: a self-assessment tool. *Journal of Applied Language Studies*, 11 (4): 35-55.

Boud, D. 1986. *Implementing student self-assessment,* Green Guide No. 5, Sydney: Higher Education Research and Development Society of Australasia.

Boud, D. 1989. The role of self-assessment in student grading. *Assessment and Evaluation in Higher Education,* 14 (1): 101-111.

Boud, D. 1999. *Implementing student self-assessment.* 3rd ed. Campbell Town: The Higher Education Research and Development Society of Australasia (HERDSA).

Boud, D. 1995. *Enhancing learning through self-assessment*. London: Kogan Page.

Boud, D. & Falchikov, N. 1989. Quantitative studies of self-assessment in higher education: a critical analysis of findings. *Higher Education*, 18 (5): 529-549.

Boud, D., Lawson, R. & Thompson, D. 2013. Does student engagement in self-assessment calibrate their judgement over time? *Assessment and Evaluation in Higher Education*, 38 (8): 941-956.

Boud, D., Lawson, R. & Thompson, D. 2015. The calibration of student judgement through self-assessment: disruptive effects of assessment patterns. *Higher Education Research and Development*, 34 (1): 45-59.

Brew, C., Riley, P. & Walta, C. 2009. Education students and their teachers: comparing views on participative assessment practices. *Assessment and Evaluation in Higher Education,* 34 (6): 641-657.

Broadfoot, P. & Black, P. 2004. Redefining assessment? The first ten years. *Assessment in Education: Principles, Policy and Practice,* 11 (1): 7-27.

Brown, G.T.L. & Harris, L.R. 2013. Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks, CA; SAGE.

Butler, R. 1987. Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest and performance. *Journal of Educational Psychology,* 79 (4): 474-482.

Cassidy, S. 2007. Assessing 'inexperienced' students' ability to self-assess: exploring links with learning style and academic personal control. *Assessment and Evaluation in Higher Education,* 32 (3): 313-330.

Challis, M. 1999. AMEE Medical education guide no 11 (revised): portfolio-based learning and assessment in medical education. *Medical Teacher,* 21 (4): 370-386.

Cowan, J. 1988. Struggling with student self-assessment, in D. Boud (Ed.) *Developing student autonomy in Learning* (2nd Edition), London: Kogan Page.

Crossman, J. 2007. The role of relationships and emotions in student perceptions of learning and assessment. *Higher Education Research and Development,* 26 (3): 313-327.

Denzin, N.K. & Lincoln, Y.S. 1998. Introduction: entering the field of qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), *The landscape of qualitative research: theories and issues* (pp. 1-34). Thousand Oaks, CA: Sage.

Dochy, F., Segers, M. & Sluijsmans, D. 1999. The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education,* 24 (3): 331-350.

Falchikov, N. 1995. Peer feedback marking: developing peer assessment. Innovations in *Education and Training International,* 32 (2): 175-187.

Jonsson, A. 2014. Rubrics as a way of providing transparency in assessment. *Assessment and Evaluation in Higher Education,* 39 (7): 840-852.

Kirby, N.F. & Downs, C.T. 2007. Self-assessment and the disadvantaged student:

potential for encouraging self-regulated learning? *Assessment and Evaluation in Higher Education,* 32 (4): 475-495.

Kohn, A. 2011. The case against grades. *Educational Leadership,* 69 (3): 28-33.

Langan, A.M., Shuker, D.M., Cullen, W.R., Penney, D. Preziosi, R.F. & Wheater, C.P. 2008. Relationship between student characteristics and self-, peer and tutor evaluations of oral presentations. *Assessment and Evaluation in Higher Education,* 33 (2): 179-190.

Lew, M.D.N., Alwis, W.A.M. & Schmidt, H.G. 2010. Accuracy of students' self-assessment and their beliefs about its utility. *Assessment and evaluation in Higher Education,* 35 (2): 135-156.

Longhurst, N. & Norton, L.S. 1997. The self-assessment in course work essays. *Studies in Educational Evaluation,* 23 (4): 319-330.

Mok, M.M.C., Lung, C.L., Cheng, D.P.W., Cheung, H.P. & Ng. M.L. 2006. Self-assessment in higher education: experience in using metacognitive approach in five case studies. *Assessment and Evaluation in Higher Education,* 312 (4): 415-433.

Nulty, D.D. 2011. Peer and self-assessment in first year of university. *Assessment and Evaluation in Higher Education,* 36 (5): 493-507.

Orsmond, P. & Merry, S. 2012. The importance of self-assessment in students' use of tutors' feedback: a qualitative study of high and non-high achieving biology undergraduates. *Assessment and Evaluation in Higher Education,* 38 (6): 737-753.

Orsmond, P., Merry, S. & Reiling, K. 2000. The use of student derived marking criteria and peer and self-assessment. *Assessment and Evaluation in Higher Education,* 25 (1): 21-38.

Panadero, E., Alonso-Tapia, J. & Reche, E. 2013. Rubrics vs self-assessment scripts effect on self-regulation performance and self-efficacy in pre-service teachers. *Studies in Educational Management,* 39 (3): 125-132.

Pereira, D., Flores, A. & Nicklasson, L. 2016. Assessment revisited: a review of research in Assessment and Evaluation in Higher Education, *Assessment and Evaluation in Higher Education,* 41 (7): 1008-1032.

Pulfrey, C., Buch, C. & Butera, F. 2011. Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology,* 103 (3): 683-700.

Reddy, Y.M. & Andrade, H. 2010. A review of rubric use in higher education. *Assessmentand Evaluation in Higher Education,* 35 (4): 435-448.

Sadler, D.R. 1989. Formative assessment and the design of instructional systems. *Instructional Science,* 18: 119-144.

Sadler, D.R. 2009. Indeterminacy in the use of preset criteria for assessment and grading. *Assessment and Evaluation in Higher Education,* 34 (2): 159-179.

Sadler, D.R. 2010. Beyond feedback: developing student capability in complex appraisal.*Assessment and Evaluation in Higher Education,* 35 (5): 535-550.

Sadler, P.M. & Good, E. 2006. The impact of self- and peer-grading on student learning.*Educational Assessment,* 11 (1): 1-31.

Sambell, K. & McDowell, L. 1997. The value of self- and peer assessment to the developing lifelong learner. In C. Rust (Ed.), *Improving student learning – improving students as learners* (pp. 56-66). Oxford, Oxford Centre for Staff and Learning Development.

Sher, W.D. & Twigg, D.R. 1991. Peer assessment – a construction 'Tool'? In S. Brown (Ed.), *Peer Assessment in Practice* (pp. 87-105). Birmingham: Staff & Educational Development Association.

Sivan, A. 2000. The implementation of peer assessment: an action research approach.*Assessment in Education,* 7 (2): 193-21.

Syed, M. 2011. Bounce. *The myth of talent and the power of practice.* London: FourthEstate.

Taras, M. 2015. Student self-assessment: what have we learned and what are the challenges. *RELIEVE,* 21 (1), art. ME8. DOI: 10.7203/relieve.21.1.6394.

Wiggins, G. 1998. *Educative assessment.* San Francisco: Jossey-Bass.

ABOUT THE AUTHOR

Sharita Bharuthram

Senior Lecturer, English for Educational Development, University of the Western Cape, Private Bag X17, Bellville, 7535.

Dr Sharita Bharuthram is a Senior Lecturer in the English for Educational Development (EED) Programmes at the University of the Western Cape, South Africa. Dr Bharuthram co-ordinates and lectures an EED module which is offered to students from the Faculty of Community and Health Sciences. Her PhD was on the reading development of students in higher education which she obtained in 2007 through the University of KwaZulu-Natal. Since then her research focus has been in the area of academic literacies and integrating these into the students' disciplines. Her research interests also include assessment practices, in particular, self and peer assessment.

Tydskrif vir Taalonderrig - Journal for Language Teaching
- Ijenali yokuFundisa iLimi - IJenali yokuFundisa iiLwimi -
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta Polelo
- Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo - Ijenali
Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -
Journal for Language Teaching - Ijenali yokuFundisa iLimi
- IJenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa
Ririmi - Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali yokuFundisa
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo
- Jenala yo Dyondzisa Ririmi
- Tydskrif vir Taalonderrig
- Journal for Language
Teaching - Ijenali
yokuFundisa iLimi -
IJenali yokuFundisa
iiLwimi - Ibhuku
Lokufundisa Ulimi
- Tšenale ya tša
Go ruta Polelo -
Buka ya Thuto
ya Puo - Jenale
ya Thuto ya Dipuo
- Ijenali Yekufundzisa
Lulwimi - Jena?a ya u
Gudisa Nyambo - Jenala   yo
Dyondzisa Ririmi - Tydskrif vir Taalonderrig
- Journal for Language Teaching - Ijenali
yokuFundisa iLimi - IJenali yokuFundisa iiLwimi -
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya
Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a ya ya
u Gudisa Nyambo - Jenala yo Dyondzisa Ririmi
- Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali
yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi -
Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa
Ririmi - Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali yokuFundisa
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -
Journal for Language Teaching - Ijenali yokuFundisa iLimi
- IJenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa
Ririmi - - Tydskrif vir Taalonderrig - Journal for Language
Teaching - Ijenali yokuFundisa iLimi - IJenali yokuFundisa
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta