AOSIS

Perceptions of writing prowess in honours economic students

CrossMark

Authors:
Jean du Toit[1] ®
Anmar Pretorius[2] ®
Henk Louw[3] ®
Magdaleen Grundlingh[4] ®

Affiliations:
[1]Research Focus Area for Chemical Resource Beneficiation, Faculty of Natural and Agricultural Sciences, North-West University, Potchefstroom, South Africa

[2]School of Economic Sciences, Faculty of Economic and Management Sciences, North-West University, Potchefstroom, South Africa

[3]School of Languages, Faculty of Humanities, North-West University, Potchefstroom, South Africa

[4]Writing Centre, Faculty of Humanities, North-West University, Potchefstroom, South Africa

Corresponding author:
Anmar Pretorius,
anmar.pretorius@nwu.ac.za

Read online:
Scan this QR code with your smart phone or mobile device to read online.

Background: Economists are often asked to explain or foresee the economic impact of certain events. Except for theoretical and practical knowledge, clear communication of views is therefore required. However, post-graduate training in Economics mostly focuses on technical modules. Furthermore, students often overestimate their (writing) abilities – as described in the Dunning-Kruger effect.

Aim: This article aims to establish if, and to what extent, perceptions of writing quality differ between students, subject-specific lecturers and writing consultants.

Setting: Honours students at a South African university wrote an argumentative essay on a specific macroeconomic policy intervention.

Methods: In this study qualitative samples (an evaluation rubric) were quantified for an in-depth analysis of the phenomenon, which allowed for a mixed-methods research design. The essays were evaluated by fellow students, the Economics lecturer, Academic Literacy lecturers and Writing Centre consultants and then their evaluations were compared. The evaluation form contained 83 statements relating to various aspects of writing quality.

Results: Student evaluators in the peer review were much more positive than the other evaluators – in a potential confirmation of the Dunning-Kruger effect. However, despite the more generous evaluations, students were still able to distinguish between varying skills levels, that is, good and bad writing. Discrepancies in evaluations between the subject specialists were also observed.

Conclusion: More conscious effort needs to be put into teaching economics students the importance and value of effective writing, with clear identification of the requirements and qualities of what is considered to be effective writing.

Keywords: perceptions; writing skills; economics students; Dunning-Kruger effect; peer review; writing assessment

Introduction

Background: The importance of writing for economists

Economists are often asked for opinions about the economic consequences of certain events. Examples include the potential impact of Brexit on South African exports to the United Kingdom (UK) and the potential impact of the US-China trade war on South African trade. To respond to such requests, economists have to analyse the situation, using subject knowledge and existing models, but they are then also required to clearly communicate their message, either in written format or through oral communication (interviews).

Curriculums for post-graduate studies in Economics mostly include technical modules and other applied fields of economic theory and although the field of economics is quite logically associated more with numbers than with writing ability, some scholars on economics make a strong case for honing writing skills. To be an economist, is to be a writer, or in the words of McCloskey (1985):

> Economics depends much more on the mastery of speaking and writing than on the mastery of engineering mathematics and biological statistics usually touted as the master skill of the trade. (p. 188)

McCloskey (1985:188) further states that writing '… is the Economist's trade'. He also identifies the problem: Most people who write a lot, as do economists, have an amateurish attitude towards writing' (McCloskey 1985:187). Therefore, the question: 'How well do you think you write and

how well do other people think you write?' at the heart of this article, is a very relevant question for post-graduate students training to be economists.

McCloskey is not alone in his opinion; therefore, the graduate attributes ascribed to by the Faculty of Economic and Management Sciences at the North-West University, include the following, very specific references to written communication:

> In addition to the general outcomes of the B.Com Hons degree, the content of this qualification is structured in such a manner that specific exit levels (including the critical outcomes) will enable students to demonstrate the ability to **formulate**, **present** and **communicate** insightful and creative academic and professional ideas and arguments effectively – verbally and **in writing**, using appropriate media and communication technology and suitable economics methods. (NWU 2015: 37 [Emphasis added])

Specific module outcomes for honours in Economics also refer to the ability to analyse and disseminate information and present it in an 'honours dissertation' (written research report) and an oral presentation (NWU 2015).

Problem

Despite these explicit requirements regarding writing contained in outcomes, lecturers often experience that, students, when required to write longer texts in their final and honours years, illustrate a half-hearted outlook regarding the necessity of effective writing. Many fail to see the relevance of effective written communication and display an apathetic attitude towards acquiring or honing these skills. This may in part be due to large class sizes at undergraduate level, which necessarily limit written assignments as options of assessment. As a result, extended writing is seldom done at undergraduate level, which sends the unintended message to students that writing as a skill is not required. In addition, writing may be considered as the educational outcome of the so-called soft sciences (humanities); and therefore not considered to be necessary for the harder, numerical sciences.

Dunning (2011:248) touched on this in his explanation of the Dunning-Kruger effect, with specific reference to economists. He states that economists have reasoned that most ignorance is rational, seeing that in many cases the effort to gain expertise on several topics would not provide a person with enough perceptible benefit to justify the expended learning effort. This seems to be aggravated by students' unfounded beliefs that their writing abilities are of higher standard than is the actual case. When their texts are measured objectively or externally (by lecturers), the perceived quality is in reality much lower than students themselves, and by implication peers, evaluate it to be – a frustrating situation for lecturers.

The current study was therefore prompted in part by perceptions on part of the lecturers who experienced that students in economics had erroneous beliefs regarding their writing abilities; the lecturers wished to dispel these mythical self-evaluations. In addition, at face value, students with

overall better performance in the subject seemed to be able to communicate better in writing about the subject content. This is already a classic manifestation of the so-called Dunning-Kruger effect (Kruger & Dunning 1999). Although not reported as part of the Dunning-Kruger effect, such a disparity between actual and perceived performance may lead to a sense of disillusionment and avoidance (Munchie 2000:49) when students receive results for written assessments. Numerous articles on writing pedagogy (Hyland 1990; Spencer 1998) report, for example, that students discard feedback due to not possessing the skill to interpret and understand the feedback they receive; failing to use it to improve future writing. It may also lead to other non-desirable consequences such as withdrawal, or a lack of trying when constructing follow-up texts.

The activity reported on in this article can also be seen as an exercise in awareness raising (or 'consciousness raising' as in James 1997:257), which was aimed at establishing if and to what extent the perceptions of writing quality differ among students, lecturers (subject-specific and Academic Literacy lecturers), and also Writing Centre consultants. The ultimate goal with this exercise was to raise the awareness of students as to the quality of writing expected of them, with the hope that an illustration of the mismatch between peer evaluation and actual performance may stimulate greater attention to detail and metacognitive evaluation of activities when working on written texts.

Method preview

For this exercise a written assignment was evaluated on four levels: peer to peer, subject lecturer to student, Writing Centre consultant to student, and Academic Literacy lecturer to student (the full method is explained under methodology). Evaluations were done via a Google Forms interface, which yielded quantifiable data. The data were analysed and discussed with students to identify the most salient discrepancies in evaluation, and these discrepancies are also identified and discussed in this section. For the purposes of this article, an in-depth analysis of six of these texts is also included.

Findings preview

Although there are certain areas of agreement in evaluation, some specific areas find great discrepancies in evaluation among students, subject lecturers, Academic Literacy lecturers and writing consultants. These areas can be associated with the finesse of writing.

Preview

In the rest of this article an overview of the Dunning-Kruger effect is first presented, then what was defined as quality writing in this specific context is discussed, whereafter the full methodology is explained, followed by a discussion of the findings and their implications for the teaching of writing to students of economics.

Literature overview
Dunning-Kruger effect

The Dunning-Kruger effect explains that a person's perception of his/her ability is rarely accurate (Kruger & Dunning 1999). While it falls outside the scope of this article to provide a full overview of the Dunning-Kruger effect, a few key points relevant to this article will be raised. Dunning and Ehrlinger (2003:5) explain that a person's views of their own abilities in general and during specific instances, usually do not correlate well with the actual quality of the performance and in some instances do not correlate whatsoever. Dunning and Ehrlinger (2003:5–6) then list a range of published research articles in which the findings repeatedly illustrate that actual performance rarely, if ever, matches perceptions of performance and note that people generally hold perceptions of their overall ability or specific performances that only correlate modestly with the reality of the actual output. The popular maxim of 'believe it and you can do it' is, simply put, false.

Peer evaluations, by contrast, tend to be much more accurate with regards to aspects such as leadership qualities, but as for an illustration of subject knowledge, an incompetent peer suffers from the same 'unknown unknowns' (Dunning 2011) as the one whose work they are evaluating. This statement is also supported by the study of Brindley and Scoffield (1998:85) in which they state that students felt peer evaluations to be unfair due to a lack of objectivity.

Grammar and argumentation formed the basis of a few of the experiments central to the Dunning-Kruger theory and in Dunning (2011:261) it is explained that the ability to judge whether you or someone else has written a sentence that is grammatically correct, relies on the same skillset required to produce a grammatically correct sentence in the first place. The same holds true for the evaluation of an argument. In order to evaluate if an argument is sound, the same know-how is needed as for constructing one. A poor performer has limitations which cause them to fail when attempting to form correct responses. Those same limitations also fail them when they have to judge the accuracy and correctness of their own responses and those of others. They would be unable to know when their own responses are incorrect and will lack the ability to know when somebody else formed a better response.

Performance estimates are also partly driven by perceived task difficulty. If something is perceived to be easier, a person tends to overestimate their ability and if something is perceived to be hard, they underestimate their ability (Dunning 2011). To complicate matters even further, the more skilled a person is in a task, the more accurate they can judge the level of difficulty, while the more unskilled a person, the less chance they have of accurately judging the difficulty of the task at hand (Kruger & Dunning 1999). This leads to Dunning (2011) explaining that the Dunning-Kruger effect is not a single misjudgement, but a range of activities all dealing with perceptions and judgements of abilities and performance.

Probably the most well-known findings from the research done on the Dunning-Kruger effect, can be summarised in the quote that poor performers are:

> … doubly cursed: their lack of skill deprives them not only of the ability to produce correct responses, but also of the expertise necessary to surmise that they are not producing them. (Dunning et al. 2003:83)

This leads to incompetent students failing to estimate the level of their incompetence (those performing in the lower 25% extensively overestimate their performance [Dunning 2011:264]), while those who are competent, underestimate the level of their competence. Note that two estimates are applicable here: the first is judging the task difficulty and the second is judging performance. Poor performers are worse in judging performance, regardless of whether it is their own performance or that of someone else – as confirmed in study three of the four Kruger and Dunning (1999) studies.

Although Dunning (2011) reports that motivation plays no statistically significant role in the Dunning-Kruger effect, one may suspect that the perceived usefulness of the task influences the amount and level of care and conscious effort a student will spend on an activity, such as producing and evaluating a written assignment. In this context, when students consider writing to be of no importance for their academic studies, one can assume they do not see the relevance of improving it, and if they think the task of writing is easy, they will overestimate their and others' ability to write well.

Previous South African studies

Studies on the performance of South African Economics students are limited to the overall performance in specific modules like first year (Edwards 2000; Horn & Jansen 2009; Parker 2006; Pretorius, Prinsloo & Uys 2009; Smith 2009; Smith & Edwards 2007; Van der Merwe 2006; Van Walbeek 2004) and second year (Horn, Jansen & Yu 2011; Smith & Edwards 2007). No empirical study could be found that focuses on performance in post-graduate modules in Economics. Broadly linked to the theme of this article, Matoti and Shuma (2011) assessed the writing efficacy of post-graduate students at the Central University of Technology. Their sample included students with three post-graduate qualifications, but none of these were in Economics. Students were found to exhibit low levels of writing efficacy through-out the writing process – including 'sentence construction, lack of understanding of concepts related to the discipline, problems in interpreting questions, spelling, grammar, tenses, reference techniques, punctuation, writing coherently and logically, synthesising information'. No self- or peer-evaluation was present in this study.

Only two studies were found investigating the Dunning-Kruger effect in a South African context. In Dodd and Snelgar (2013) students' self-evaluation was compared to a formal assessment of learning. The findings indicate that students with lower levels of learning achievements had higher levels

of self-evaluation than students with moderate learning achievements. As was to be expected, students with the highest learning achievements also demonstrated the highest levels of self-evaluation, in keeping with the international findings. The study did not, however, focus specifically on writing abilities, leaving a large gap in the research. The only other South African study that used the Dunning-Kruger effect, was by Yani, Harding and Engelbrecht (2019) which attempted to explain the development of students' perceptions regarding their academic maturity.

Defining quality writing

While 'quality writing' is the topic of numerous books on its own (see McCloskey 1985; Wentzel 2017), it is necessary for the purposes of this article to explain what was defined as quality writing in this specific exercise. This is of course the first place where the Dunning-Kruger effect manifests itself: in essence the question is if a person knows what (s)he experiences to be quality writing, or if it falls into the categories of 'unknown unknowns' or 'general world knowledge', or 'reach-around knowledge' (to use the terminology in Dunning 2011:252). Dunning (2011) explains 'reach-around knowledge' as something which appears to be 'relevant and reasonable' but is actually make-belief standards.

The definition for a quality text used for the context of this article, is a few years in the making for the learners. All learners were required to take a compulsory course in Academic Literacy in their first year of study (about three years prior to the experiment for most learners in this study). In the Academic Literacy module, students were introduced to certain universally accepted qualities of effective academic writing (Adrianatos et al. 2018; Van der Walt 2013). In their current honours year, workshops were presented to the students on elements of good writing covering topics on the structure of a dissertation (introduction, literature review, conclusion, etc), as well as general writing skills (sentences, paragraphing, logical order, linking devices, source use, etc). Their written assignments were also assessed using a marking scheme (see the method section) which was created with reference to earlier work covered by the students in their first year of study and again revisited in said workshops.

The marking scheme (evaluation rubric) for their written assessment was divided into five sections. The complete evaluation rubric is available as addendum, but an overview is provided below. Note that it falls outside the scope of this article to provide a full academic justification for the categories and also note that the concept of 'quality' with regards to written texts admittedly contains many more categories than those selected for this exercise. For obvious practical reasons, some limits had to be imposed on the number of qualities attended to.

The five categories of text quality used for this exercise

Five categories of text quality were used in the marking rubric based on expert opinion for the inclusion of standard writing elements and the specific goals for the written assignment: title, language, paragraphing, macrostructure, and academic sources, each of which is elaborated on very briefly below.

The title of the text should be fit for purpose, clear and descriptive. Titles like 'assignment 1' or 'essay' are of insufficient quality. The ability to write a clear concise title illustrates the ability to state the main idea of a text succinctly.

The category of Language refers to objective measures (correctness), as well as subjective measures of quality or standard. Punctuation, spelling and grammar are in most cases objectively measured – a matter of right or wrong. Sentence length and structure are more subjective categories, referring to the readers' expectations of and experience with the sentences – overly long and complex sentences are usually more difficult to read, while an overuse of very short sentences stunt reading and fluency and may create the impression that the writer is unable to 'string together' larger sections of meaning, which may imply a lack of subject knowledge or a lack of linguistic control.

Paragraphing is a factor of writing that seldom falls into place automatically, but takes conscious editing efforts. Well-structured paragraphs should focus on one main idea per paragraph and contain a clear topic sentence. Support for the main idea (examples, elaboration, explanations) should be present in the paragraph and these support statements should be clearly linked to the main idea and to each other. For this reason, the use of so-called linking devices are vital (linking devices are known by a number of synonyms, like cohesive devices, discourse markers, or transition markers).

The Macrostructure category refers to the organisation of information in the text. For this, each section has certain characteristics which usually occur in academic texts. For example, the introduction and conclusion each has specific qualities which need to be present. These qualities could be measured objectively (are they there?), as well as subjectively (are they good?). The introduction should at the very least contain a background which leads to a thesis. It should also contain an overview of the discussion to follow, as well as an overview of findings. The conclusion should in turn present a summative review of the text with a final verdict on the question investigated. The conclusion should not present new information.

The most subject-specific part of the text (still part of the macrostructure), is the text body in which the main argument is presented. For this, students were required to provide additional background, as well as an explanation of the concept under discussion.

The final evaluation category is the use of academic sources, once again subject to both objective evaluation (the correctness of referencing and bibliographical formats) and subjective evaluation, which refers to how well academic sources were used to support statements in the text and the argument in general.

The evaluation scheme was set up to allow the reader the 'ease into' the text and gradually build up to the more cognitively demanding aspects of textual evaluation. For example, while it is relatively easy to spot a typing error or identify a sentence with unclear meaning, it takes a lot more effort and subject knowledge to identify errors in the argument. In the terminology of the Dunning-Kruger effect, evaluation was started at supposedly 'known-knowns' (spelling) and moved to 'known-unknowns' (where a learner can see that, for example, a sentence is bothersome, but is unable to identify how to improve it) to 'unknown-unknowns' (a complete inability to identify an issue, which will become apparent when a student evaluates an argument as effective, while it is in fact incomplete). Awareness raising occurs when known-unknowns become clearer and unknown-unknowns are taught. This systematic approach in evaluation also mirrors to an extent the study by Williams, Dunning and Kruger (2013) who found that a systematic approach to a problem gives learners greater confidence in their performance and evaluation of own performance.

Method

A mixed method exploratory sequential design (qualitative followed by quantitative) was used to investigate the observed trend in-depth and to measure its occurrence (Creswell & Clark 2017; Creswell et al. 2003). A similar approach was employed by Joubert and Feldman (2017). A rubric, a qualitative grading method in the form of an evaluation form, was utilised to record the evaluators' perceptions of the written assignments. The responses were then quantified for an in-depth analysis.

Experiment setup: Data generation system

Students wrote a long essay of 2000 words as part of their preparation for entering the annual National Budget Speech competition. The essay also formed part of their assessment plan for the semester. The requirements for the essays are outlined on the competition website, but can be summarised that students should argue for or against a specific economic policy intervention; in this particular case, quantitative easing (abbreviated as QE).

The completed essays were then evaluated by four different stakeholders using an evaluation form in question format on Google Forms. The content of the evaluation form is discussed above. The form was set up in multiple choice format, meaning that it contained a number of statements on which evaluators could respond by picking only one option (although additional comments could be added as general comments at the end of a section). This format was chosen as clearly expressed answer statements reduce marker uncertainty. Additionally, a feedback letter containing the selected answers was automatically generated giving the students immediate guidance on the issues in their manuscript. In total, 83 statements had to be evaluated. Figure 1 shows an extract of the evaluation form.

This technique allows for focused peer review as well as having the benefit of producing a data set which is easily quantifiable. For pedagogical purposes it was also connected to a form-fill letter which provided a feedback letter of approximately three pages in length to the student, with the letter containing the chosen comments. Students would therefore have been able to use this feedback to revise their work towards their second draft, but also to use the feedback form as general guidelines for future writing. The form was created by staff at the North-West University Writing Centre.

The evaluators

All submitted essays were first evaluated by Writing Centre consultants, which yielded 66 responses on the Google Form datasheet. The next evaluation was a peer review exercise using the same form. For the peer review, all students were assigned to evaluate the essays written by two peers and they received access to these files via electronic distribution. This was done with the intention of raising their awareness of the writing standard as well as the subject content in the essays of the peers. This part of the exercise provided 103 responses. Students were not randomly assigned to each other's work, but were matched according to subject field, with students from the different fields (economics, international trade, and risk management) evaluating students taking similar subject combinations.

The next evaluator was a subject specialist – the lecturer for the subject. The lecturer was assigned to evaluate a selection of six essays, which were randomly selected from three bands of results – high, average and low academic achievement scores.

FIGURE 1: Extract from digital evaluation form.

The final evaluators were lecturers from the subject group, Academic Literacy who also evaluated the selection of six essays. In this instance the Academic Literacy lecturers and the Writing Centre consultants act as the writing specialists, or non-subject specialist evaluators, with the aim of focusing on writing structure and style. Their evaluations therefore serve as the external standard.

Students were not required to do a self-evaluation on the Google Form, which would have provided interesting additional data. However, the expectation was that students would have utilised the form as a marking rubric before submission to evaluate their own work and, based on their self-evaluation, submitted an essay they deemed to be of good quality.

Ethical considerations

All participating students signed the standard consent form of the Writing Centre which allows the university to use submissions at the Writing Centre in anonymised research. The consent form also informed students on confidentiality and withdrawal procedures.

Analysis method

Two sets of analyses were done on the data.

In the first place, the data were observed holistically. In this instance, data for the evaluation for all 66 student essays are discussed. The data obtained for this set includes peer review and writing consultant review answers to the form questions. The second analysis is an in-depth analysis on just six student essays. These six essays were also evaluated with the same form, by the subject lecturer as well as a writing specialist.

Data set A: Holistic analysis

All students were bona fide students who had moved directly from undergraduate studies to their honours degree. For the majority, English is their second language.

Table 1 summarises the overall evaluations of the writing consultants and the students (peer-reviewing other students' essays). The third column contains the specific statement that was used to evaluate the different aspects of the essays. From the total number of 83 questions only the answers that had a percentage difference of greater than 20% points between the consultants and the students were further analysed (22 answers in total). This percentage difference was selected according to the Pareto Principle as utilised by O'Neill (2018) to identify the 'critical few' errors that could improve 80% of the students' writing ability. Because 22 responses (very close to 17 out of 83) had a difference of greater than 20% this percentage was therefore chosen as the cut-off for the analysis. The fourth and fifth columns indicate the percentage of cases for which the consultants and students agreed with the statement. For example, the consultants felt that 30.3% of the essays had a descriptive title, while the students in their peer review felt that 52.4% of the essays had a descriptive title. The last column of Table 1 indicates the percentage point difference between the two scores. The table is organised

TABLE 1: Comparing the evaluation between writing consultants and students.

Category of quality evaluated		Statement evaluated and number	Consultants	Students	Differ
Configuration		The title is descriptive (A1)	30.3	52.4	22.1
Language	Punctuation	You had correct punctuation and spaces (A4)	21.2	62.1	40.9
Language	Sentence structure	In general your sentence structure was good (A12)	53.0	82.5	29.5
Language	Word choice	In general your word choice was formal (A14)	13.6	37.9	24.3
Language	Spelling	Your essay had no spelling mistakes (A17)	27.3	52.4	25.1
Language	Grammar	Your essay had good grammar (A20)	33.3	83.5	50.2
Paragraphing	One idea	Your paragraphs had single clear main ideas (A22)	60.9	84.3	23.4
Paragraphing	Clear topic sentence	Most of your paragraphs had a clear topic sentence (A24)	48.5	70.9	22.4
Paragraphing	Logical linking of words	You used linking words or phrases effectively to guide the reader to follow your train of thought (A26)	30.3	67.0	36.7
Paragraphing	Logical linking of ideas	The main ideas linked logically and obviously (A28)	45.5	87.4	41.9
Macro structure					
Introduction	Thesis statement	Your thesis statement is clear, indicating the position that you will be taking (A32)	19.7	66.0	46.3
Introduction	Overview	It is clear what main parts your discussion will follow (A34)	48.5	83.5	35.0
Introduction	Preview of conclusion	It is clear what your conclusion to the problem is (A36)	24.2	66.0	41.8
Specific position	Reason for specific policy	The reasons why you took the position is explained (A52)	31.8	58.3	26.5
Specific position	Superiority of your position explained	The superiority of your position is explained showing why the alternatives should not be used (A56)	16.7	51.5	34.8
Conclusion	Summary	The conclusion in your essay had a clear recapping or summary of the main parts of the argument (A60)	33.3	72.8	39.5
Resources	Sources	Your use of sources was good (A66)	45.5	73.8	28.3
Resources	Referencing style	You used one style consistently and correctly (A69)	60.6	84.5	23.9
Resources	In-text referencing correctness	In your text, you used the referencing style correctly (A72)	21.5	70.7	49.2
Resources	Referencing of facts	In your use of facts you have a good mix of sources (A79)	43.9	68.9	25.0
Resources	Bibliography – matching sources	Your bibliography corresponds correctly with the in-text references (A81)	27.3	80.6	53.3
Resources	Bibliography correctness	Your bibliography was correct (A83)	1.5	59.2	57.7

according to the five categories of text quality which is also the order in which they appeared in the marking rubric.

For all of the statements, without exception, the students' peer review was much more positive than the evaluation of the consultants. This is an overwhelming indication of students' over estimation of their peer's performance. Figure 2 provides a graphical representation of the values in Table 1. Panel A indicates the percentage awarded by the students and consultants for each statement number, while Panel B illustrates the percentage point difference for each statement (similar to the last column of Table 1). The smallest difference (of 22.1) was observed for the statement relating to a descriptive title (number A1) and the largest (57.7) for statement A83 about a correct bibliography.

A closer look at Panel B indicates that some of the largest differences are observed in the later statements (81–83) relating to the handling of resources. Students have a much more skewed perception of their abilities and accuracy in handling sources. Statements dealing with paragraphing received scores closest to each other (discussed later in the article).

Table 2 summarises the average scores for the five categories of text qualities as explained above (see five categories of text quality used for this exercise section). Table 2 also ranks the scores per category for each group of evaluators. Rank 5 refers to the ranking of the five main categories and Rank All to the ranking of all seven categories – where macrostructure is represented by three sub-categories. Paragraphing received the highest scores from both groups of evaluators, while the use of resources received the second highest scores – also from both groups of evaluators. Indication of a relevant title, language ability and macrostructure received the lowest scores.

Apart from the evident overoptimistic view of their peers' writing ability, Table 2 conveys two additional messages or underlying trends. Despite the higher scores awarded by the students, the students' overall ranking of the various categories is similar to those of the consultants. The students agreed with the consultants that paragraphing and the use of resources were the best qualities of the essays. The students thus display the ability to satisfactorily distinguish varying skill levels between the various elements of the writing exercise. The comparison in Table 2 also indicates that, although not at a
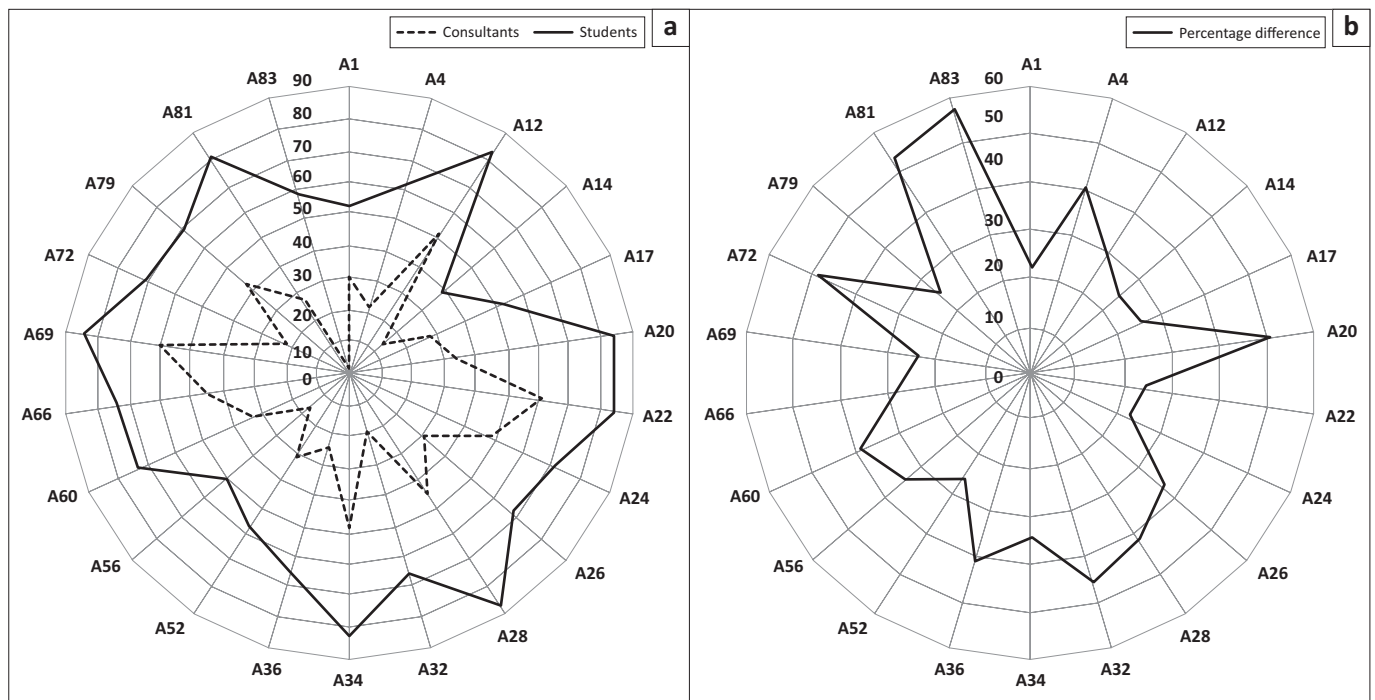
FIGURE 2: A graphical representation (a and b) of the evaluation between writing consultants and students (Table 1).

TABLE 2: Comparing average scores between writing consultants and students.

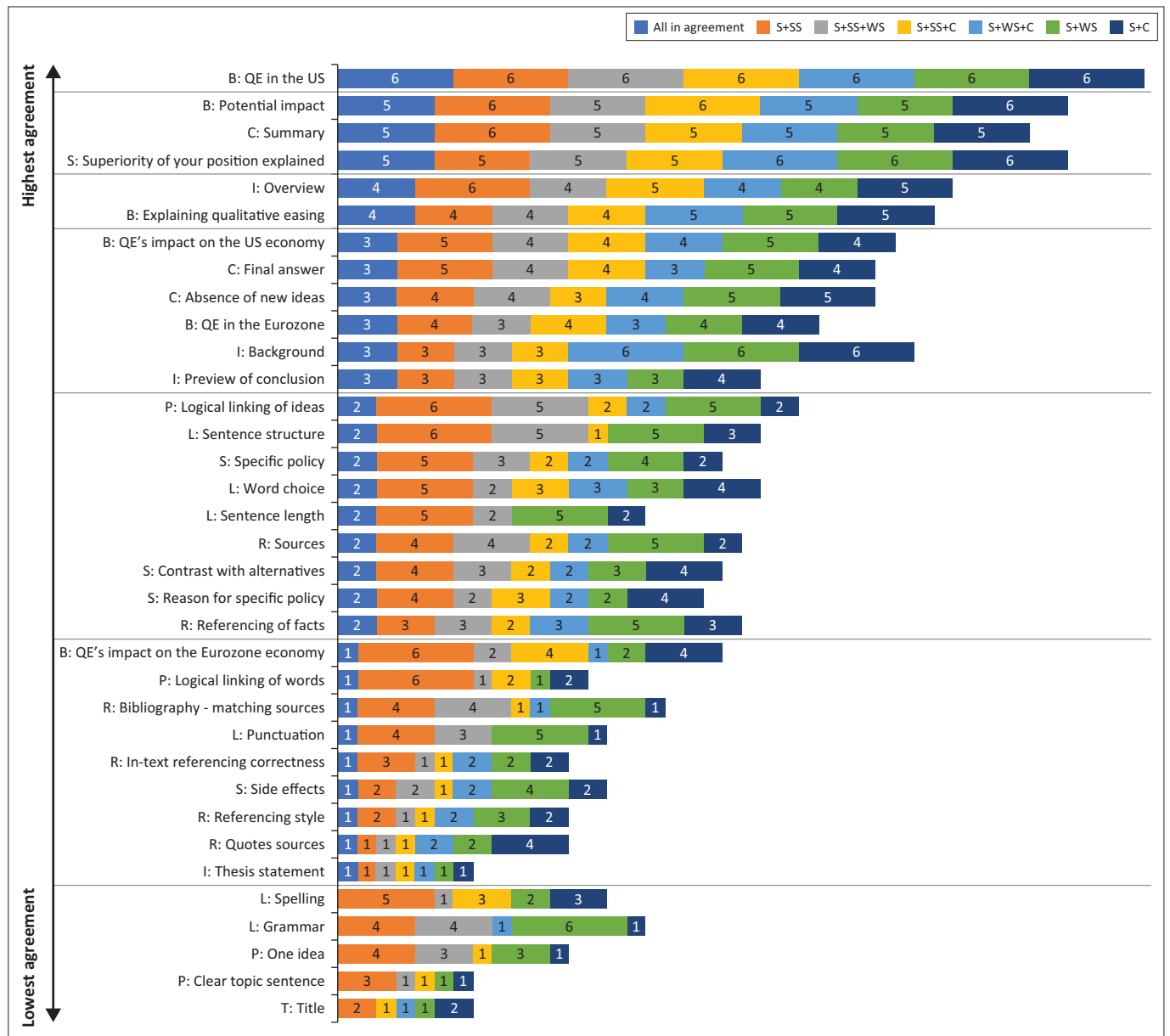| Category	Consultants			Students			Differ		
	Score	Rank 5	Rank All	Score	Rank 5	Rank All	Score	Rank 5	Rank All
Configuration	30.30	3	5	52.40	5	7	22.10	5	7
Language	29.68	4	6	63.68	4	5	34.00	3	4
Paragraphing	46.30	1	1	77.40	1	1	31.10	4	5
Macrostructure	29.03	5	-	66.35	3	-	37.32	2	-
Introduction	30.80	-	4	71.83	-	4	41.03	-	1
Specific position	24.25	-	7	54.90	-	6	30.65	-	6
Conclusion	33.30	-	3	72.80	-	3	39.50	-	3
Resources	33.38	2	2	72.95	2	2	39.57	1	2

satisfactory level yet, the students' ability to write coherent paragraphs received the highest score. This is a core competency and very much needed in building convincing arguments and communicating successfully (McCloskey 1985). It also serves as a solid foundation to improve on their ability to combine all the elements into a meaningful macrostructure. As could be expected from a sample of predominantly non-English first language speakers, some language and grammar skills are lacking and students are aware of it. Recognising the need to improve a specific skill will hopefully encourage the students to pay attention to it in future.

Data set B: In-depth analysis of six student essays

Data set B is also analysed in two ways. The first, easiest analysis is simply to compare the accuracy of the students' peer evaluation with all evaluators' responses.

Six student documents were assessed by four different evaluators. These evaluators comprised fellow students (peer evaluation) (S), the subject specialist (the Economics lecturer) (SS), a writing specialist (first-year academic writing lecturer) (WS), and a Writing Centre consultant (C). It must be noted, however, that the fellow student evaluator was not the same person for all six documents. This is also the case with the Writing Centre consultant. The documents were firstly assessed based on the number of times all evaluators agreed on the given criteria for each topic (as shown in Figure 3). This agreement is irrespective of the correctness of the writing, only the question whether all evaluators have the same opinion of the criteria is investigated.

The second analysis is an attempt at finding agreement or disagreement in evaluations based on the understanding of

Note: (students [S], subject specialist [SS], writing specialist [WS], and consultants [C]) are plotted against the topics that were evaluated in the eight different categories (Title [T], language [L], paragraphing [P], introduction [I], background [B], specific position [S], conclusion [C], and resources [R]).

FIGURE 3: The count of cases where there is an agreement on each topic in all six student documents among the evaluators.

the topics, between two 'camps' – the student and subject expert in the one camp and the Writing Centre consultant and Academic Literacy lecturer on the other hand. The student's own evaluation is compared to the evaluation of the three other evaluators. This kind of analysis will indicate where the student and subject specialist understand the definitions or concepts of writing elements differently from the writing experts. This illuminates areas of shortcoming where both the subject specialist and the student act in the area of unknown-unknowns and subsequently should be assisted by agents trained in writing, in order to attain exit level graduate attributes, but it could also indicate areas where the writing experts may need to ascertain additional clarity from subject experts in an activity like this.

Data set B: Analysis of peer evaluation accuracy

While the above discussion is drawn from a comparison between the evaluation of the students and Writing Centre consultants, the last part of the analysis compares the evaluations of the students, academics of the Writing Centre and the Economics lecturer. As discussed earlier, six specific essays were chosen and the comparison includes all the statements and aspects covered in the evaluation form and not only the 22 mentioned in the analysis method section.

Comparing the responses of the academic literacy centre (as language/writing specialist), the Economics lecturer (as subject specialist) and the students, confirms the earlier observation that students judge their fellow students' work to be of a higher standard than the subject specialists.

Out of all the responses (each student essay evaluated on 35 statements by three evaluators) in 6% of the evaluations only one of the evaluators had a positive view on the specific aspect of the essay. In all these cases, the positive view came from the student evaluation. In other words, in 6% of the evaluations both the subject and language specialists felt that the specific standard was not met – while the student felt that the standard was met. There was no instance where one of the specialists was the only one to have a positive view.

Delving deeper into the instances in which the students provided the only positive response – ignoring the one statement on the configuration (quality of the title) – resulted in more interesting observations. From the essays written by low performers, the only positive response came from the peers in 4.4% of the questions. This increases to 8.1% for the medium performers and 6.3% for the best performers. The students, therefore, tend to be more positive in their evaluation of medium and top performers compared to low performers; that is, they are able to distinguish between poor and better work. Concentrating on the different categories of text quality, the students were more positive in their evaluation than the academics regarding the use of sources (8.3% of the questions) and paragraphing (8.3%). Language quality and macrostructure came in at lower levels of 6.3% and 4.6%.

Data set B: Agreement on comprehension of elements of writing

In Figure 3 the cumulative count of cases where evaluators are in agreement are plotted for each topic of the eight different categories (Title [T], language [L], paragraphing [P], introduction [I], background [B], specific position [S], conclusion [C], and resources [R]). The data were arranged from top to bottom in the order of highest agreement to lowest agreement for all evaluators.

Hundred per cent (100%) (6/6) in agreement: On only one topic, all evaluators were in agreement: identifying the background of QE in the United States. The criteria asked if it was clear if the student used QE in the US. Therefore, because the subject terms were easy to identify, all the evaluators could correctly interpret the student's explanation. This, therefore, falls in the Dunning-Kruger category of known-knowns.

Eighty three per cent (83%) (5/6) in agreement: The three topics in which a good agreement is still observed for all evaluators, are all related to clearly identifiable contextual themes: did the writer explain his/her position, did they discuss the potential impact, and did they recap the main results in the conclusion. Thus, independent of subject knowledge, all evaluators can understand and identify main context elements.

Sixty seven per cent (67%) (4/6) in agreement: The moment the topics turned to more in-depth subject background questions, some evaluators started to disagree. Interestingly, regarding the overview of the introduction the students and the subject specialist were in full agreement. On the other hand, in respect of the subject-specific knowledge in explaining QE, the students, writing specialist and consultant agreed on five of the six documents with only the subject specialist needing more information.

Fifty per cent (50%) (3/6) in agreement: All the topics in this section relate to a misunderstanding of what was required in terms of subject content. Furthermore, the students, writing specialist and consultants were all in agreement about the quality of the background on the topic specified in the introduction. Only the subject specialist wanted a more in-depth discussion of the background.

Thirty three per cent (33%) (2/6) in agreement: The lack of expertise of each evaluator for the other's disciplines clearly starts to show. In the case of more in-depth subject knowledge in which students need to evaluate or integrate their knowledge, only the subject specialist along with the students could identify issues (specific policy, reason for specific policy, and contrast with alternatives). However, in the case of specific language elements (word choice and sentence length), the writing specialist and consultants require more sophisticated standards. Furthermore, in the case of good source use, the students tend to agree with the writing specialist and not the subject specialist who wanted better source use. This may be due to the subject specialist exhibiting a greater awareness of the quality of subject-specific sources available.

Seventeen per cent (17%) (1/6) in agreement: Four topics relating to referencing fall under this category: matching of sources between the in-text and the reference list, correctness of in-text referencing, the referencing style and the correctness of quoted sources. This result clearly highlights the issue of discrepancies in plagiarism prevention and the (mis-) understanding students illustrate regarding the stringent requirements of good referencing practice in academic writing. This could be labelled as a problematic unknown-unknown Dunning-Kruger effect.

Zero per cent (0%) (0/6) in agreement: The problem of identifying good grammar and spelling is a fundamental issue based on the standard of English writing proficiency of the students. Both the students and writing consultants overestimate the level of the students' language proficiency. Only more experienced writers (subject specialist and writing specialist) know that a higher standard is needed. However, the problem of identifying good paragraph elements is problematic because both the students and the subject specialist are in agreement, highlighting the gap in writing expertise between writing specialists and writers more focused on content.

In a typical educational situation like this, the assumption could be that the economics lecturer is the subject specialist or content expert, with some experience in writing. The student on the other hand, is a novice in writing, with limited subject knowledge, albeit more subject knowledge than the writing specialist or the writing consultant.

These assumptions are not born out in the data (Figure 4). There is a lack of agreement between the pair student and
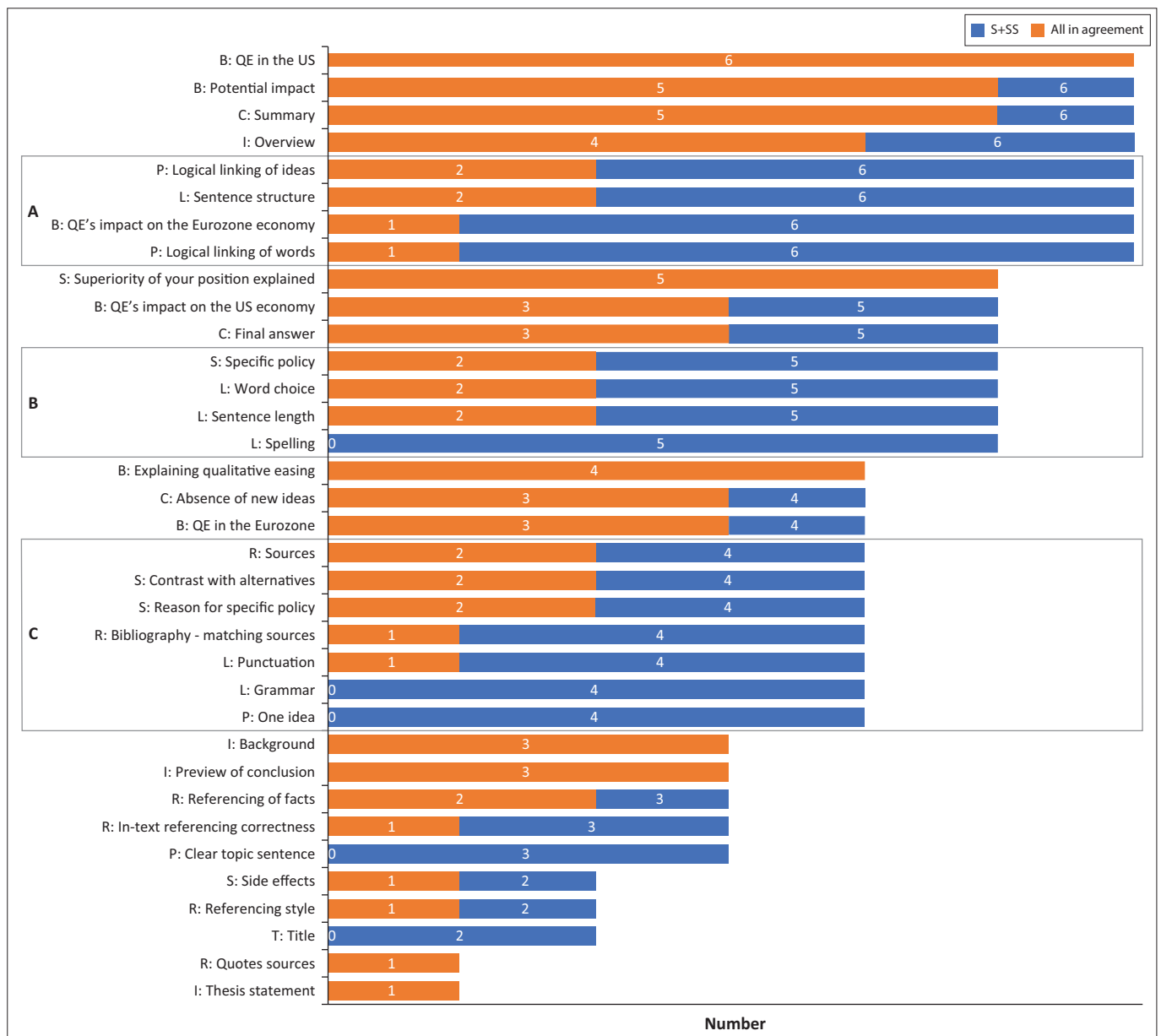
QE, quantitative easing.

FIGURE 4: The number of cases in which there is an agreement on each topic in all six student documents between the student (S) and the subject specialist (SS) (blue) where the number of all evaluators are in agreement (orange) plotted against the topics that were evaluated in the eight different categories.

TABLE 3: Types of writing per topic where all evaluators had the most disagreements.

Topic	Type of writing	Who disagreed†
I: Background	All	None
I: Preview of conclusion	All‡	None
I: Thesis statement	All‡	All
L: Spelling	All	WS+C
P: Clear topic sentence	All	WS+C
P: One idea	All	WS+C
S: Side effects	All	All
T: Title	All	All
C: Absence of new ideas	Bad	None
C: Final answer	Bad	WS+C
L: Grammar	Bad	WS+C
L: Punctuation	Bad	WS+C
L: Sentence length	Bad	WS+C
P: Logical linking of ideas	Bad	WS+C
R: In-text referencing correctness	Bad	WS+C
R: Referencing style	Bad	All
R: Sources	Bad	WS+C
S: Reason for specific policy	Bad	WS+C
B: Potential impact	Average	None
B: QE in the Eurozone	Average	None
B: QE's impact on the US economy	Average	None
I: Overview	Average	None
L: Word choice	Average	WS+C
R: Quotes sources	Average	All
R: Referencing of facts	Average	WS+C
S: Contrast with alternatives	Average	WS+C
S: Specific policy	Average	WS+C
B: Explaining qualitative easing	Good	None
B: QE's impact on the Eurozone economy	Good	WS+C
C: Summary	Good	None
L: Sentence structure	Good	WS+C
P: Logical linking of words	Good	WS+C
R: Bibliography - matching sources	Good	WS+C
S: Superiority of your position explained	Good	None
B: QE in the US	None	None

QE, quantitative easing.

†, This is based on the analysis shown in Figure 4. Therefore, this column indicates which group (the students and the subject specialist [S+SS], or the writing specialists [WS]), disagreed with the rest, or if all or none disagreed. These are rough approximations.

‡, Only student not in agreement with the rest.

subject specialist (S+SS), as well as the pair writing specialist and writing consultant (WS+WC) with regards to many of the variables. In Figure 4 we show the discrepancy between the agreement of the writing specialists and the subject specialists more clearly. The four blocks in Figure 4 denote areas of special interest. In these blocks, the subject specialists (economic student and economic lecturer) disagree with the writing experts (writing lecturer and writing consultant). Due to lack of space, all of these cannot be discussed in depth, but a few salient examples are mentioned.

In blocks B and C, it is clear that the subject specialists have the same understanding of issues of grammar, spelling and punctuation, but this understanding is not shared by the writing experts.

The next area is the concept of what entails good paragraphing. In paragraphing, a good paragraph has one main idea, and

this idea is presented with coherently linked sentences. While both the subject specialist and student had similar evaluations regarding their perceptions of these two paragraph elements, the writing specialists differed in opinion.

In Table 3 we analysed if the level of the students' writing ability influences the ease with which evaluators could identify the correctness of the writing topics. In other words, is it easier to spot problems in bad writing or good writing? The topics in the table are grouped according to the type of writing in which the evaluators disagreed the most. The type of writing is then arranged from top to bottom in the following order: all types, bad, average, good and none (no one disagreed on that/these topic[s]). The last column shows who disagreed the most about each topic as based on the analysis done in Figure 4. Therefore, on each topic we evaluated which type of writing was the most difficult to evaluate and who disagreed the most. Thus, if for instance all evaluators differed in opinion on the correctness of the background in any type of writing, it will be listed as all. If, on the other hand, they agreed, for example, about the correctness of the grammar in good and average writing but not in bad writing, the grammar will be listed as 'bad'.

It is, therefore, of interest to note which topics group together and why. In the case of the topics that are listed as 'bad' writing, we see almost all the disagreements are between the two 'camps': the writing specialists, and the students and the subject specialist. This trend indicates that the students and the subject specialist have the same idea or understanding of those topics and deem that in the 'bad' writing these topics are correct.

Furthermore, the topics listed for average writing illustrate the topics in which the evaluators have some knowledge but not sufficient to confidently analyse the correctness. It is interesting to note that in average to good writing all evaluators agreed on the topics in which basic level subject content could easily be identified. The disagreements in good writing came from the topics in which evaluators venture into the 'unknown-unknown' field, each determined by their expertise and perceived understanding of the others' discipline.

Implication of in-depth analyses data

Thus, it seems that when elements of language, like grammar, spelling and punctuation, are to be evaluated, external (language expert) assistance will be required by both students and subject specialists alike. This is not surprising, since most universities at present require their post-graduate texts to be proofread by professional language editors and many journals nowadays refuse to accept articles without proof of professional proofreading as well.

As far as effective communication (the graduate attribute mentioned earlier) is concerned, it is vital to obtain a logical flow and focus in written texts. The data illustrates that

assistance from writing pedagogues may be needed to assist students in economics to refine their writing in order to meet the graduate attributes expected of them, that is, formulate, present and communicate ideas and arguments effectively.

Conclusion: Implications and applications

The overall outcome of the study confirms that students' peer reviews of their honours essays tend to be much more positive, compared to that of the subject lecturer and writing specialists: in line with the Dunning-Kruger effect. On the positive side, however, despite the more generous evaluations, the students did display the ability to distinguish between varying skills levels, that is, good and bad writing. They ranked the elements of quality writing in the same order as the specialists and were more generous towards good writing compared to bad writing. Based on the evaluations from all the role players, paragraphing (a required characteristic of good writing) received the highest marks. It is thus comforting to know that this specific group of students, although in the first half of their honours academic year, are deemed to be able to build an argument and write coherently – the building blocks of economic rhetoric. On the negative side, they were not deemed to be eloquent in the use of sources and, being second-language English writers, certain issues regarding language use were highlighted by all the role players – including the students themselves.

Non-surprising findings were that students were able to evaluate some easier aspects of writing quite accurately, whereas the areas of writing skill illustrating greater finesse, were evaluated less accurately. Where the findings become troublesome as far as pedagogy is concerned, is when both the economics student and lecturer differed in their analyses from the analyses done by the analysers trained in writing and language, illustrating discrepancies in the experience or definition of what constitutes effective writing. A second area of contention is where only the subject lecturer disagreed in evaluation from the other three parties, illustrating either substandard subject knowledge and application from the students, or possible unclear expectations regarding academic requirements on the side of the lecturer.

Implications of this for pedagogy are quite simply that more conscious effort needs to be put into teaching economics students the importance and value of effective writing. Furthermore, the skills, requirements and qualities (definitions) of what is considered to be effective writing need to be spelt out and illustrated in more elaborate detail. Economics lecturers, and probably the syllabus, need to put more emphasis on the correct use of sources. This includes paraphrasing (not merely copying of other's ideas), as well as referencing (both in text and the compilation of a reference list). This, in combination with clear academic expectations regarding subject knowledge and academic application to content and context, may assist in raising the awareness of students about what it really takes to become an effective communicator on matters relating to their chosen profession.

An improvement on this study would entail a few more pre-exercise steps, such as a pre-test of students' theoretical knowledge of concepts and an investigation into how students acted upon the information they received about the mismatches between their self-evaluations and those of others. A self-evaluation on the same form as the peer evaluation would also provide additional quantifiable data.

References

Adrianatos, K., Fritz, R., Jordaan, A., Louw, H., Meihuizen, E., Motlhankane, K. et al., 2018, 'Academic literacy – Finding, processing and producing information within the university context', Centre for Academic and Professional Language Practice, North-West University, Potchefstroom.

Ames, D.R. & Kammrath, L.K., 2004, 'Mind-reading and metacognition: Narcissism, not actual competence, predicts self-estimated ability', *Journal of Nonverbal Behavior* 28, 187–209. https://doi.org/10.1023/B:JONB.0000039649.20015.0e

Brindley, C. & Scoffield, S., 1998, 'Peer assessment in undergraduate programmes', *Teaching in Higher Education* 3(1), 79–90. https://doi.org/10.1080/1356215980030106

Cheetham, G. & Chivers, G., 2005, *Professions, competence and informal learning*, Edward Elgar Publishing, Cheltenham, England.

Creswell, J.W. & Clark, V.L.P., 2017, *Designing and conducting mixed methods research*, Sage, Los Angeles, CA.

Creswell, J.W., Plano Clark, V.L., Gutmann, M.L. & Hanson, W.E., 2003, 'Advanced mixed methods research designs', in A. Tashakkori & C. Teddlie (eds.), *Handbook of mixed methods in social and behavioral research*, pp. 209–240, Sage, Thousand Oaks, CA.

Dodd, N.M. & Snelgar, R., 2013, 'Zulu youth's core self-evaluations and academic achievement in South Africa', *Journal of Psychology* 4(2), 101–107. https://doi.org/10.1080/09764224.2013.11885499

Dreyfus, S.E. & Dreyfus, H.L., 1980, *A five-stage model of the mental activities involved in directed skill acquisition*, Storming Media, Washington, DC.

Dunning, D., 2011, 'The Dunning-Kruger effect: On being ignorant of one's own ignorance', *Advances in Experimental Social Psychology* 44, 247–296. https://doi.org/10.1016/B978-0-12-385522-0.00005-6

Dunning, D., Johnson, K., Ehrlinger, J. & Kruger, J., 2003, 'Why people fail to recognize their own incompetence', *Current directions in Psychological Science* 12(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Edwards, L., 2000, 'An econometric evaluation of academic development programmes in economics', *The South African Journal of Economics* 68(3), 455–483. https://doi.org/10.1111/j.1813-6982.2000.tb01178.x

Horn, P. & Jansen, A., 2009, 'Tutorial classes – Why bother? An investigation into the impact of tutorials on the performance of economics students', *The South African Journal of Economics* 77(1), 179–189. https://doi.org/10.1111/j.1813-6982.2009.01194.x

Horn, P., Jansen, A. & Yu, D., 2011, 'Factors explaining the academic success of second-year economics students: An exploratory analysis', *The South African Journal of Economics* 79(2), 202–210. https://doi.org/10.1111/j.1813-6982.2011.01268.x

Hyland, K., 1990, 'Providing productive feedback', *ELT Journal* 44(4), 279–285. https://doi.org/10.1093/elt/44.4.279

James, C., 1997, *Errors in language learning and use – Exploring error analysis*, Longman, Edinburgh Gate.

Joubert, C.G. & Feldman, J.A., 2017, 'The effect of leadership behaviours on followers' experiences and expectations in a safety-critical industry', *South African Journal of Economic and Management Sciences* 20(1), a1510. https://doi.org/10.4102/sajems.v20i1.1510

Kruger, J.M. & Dunning, D., 1999, 'Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments', *Journal of Personality and Social Psychology* 77, 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Matoti, S. & Shumba, A., 2011, 'Assessing the writing efficacy of post-graduate students at a university of technology in South Africa', *Journal of Social Science* 29(2), 109–118. https://doi.org/10.1080/09718923.2011.11892961

McCloskey, D., 1985, 'Economical writing', *Economic Inquiry* 23(2), 187–222. https://doi.org/10.1111/j.1465-7295.1985.tb01761.x

NWU, 2015, '2015 Yearbook', Faculty of Economic and Management Sciences (postgraduate), North-West University, Potchefstroom.

O'Neill, K.S., 2018. 'Applying the Pareto principle to the analysis of students' errors in grammar, mechanics and style', *Research in Higher Education Journal* 34, 1–12.

Parker, K., 2006, 'The effect of student characteristics on achievement in introductory microeconomics in South Africa', *The South African Journal of Economics* 74(1), 137–149. https://doi.org/10.1111/j.1813-6982.2006.00054.x

Pretorius, A.M., Prinsloo, P. & Uys, M.D., 2009, 'Student performance in introductory microeconomics at an African open and distance learning institution', *Africa Education Review* 6(1), 140–158. https://doi.org/10.1080/18146620902857574

Smith, L.C., 2009, 'An analysis of the impact of pedagogic interventions in first year academic development and mainstream courses in microeconomics', *The South African Journal of Economics* 77(1), 162–178. https://doi.org/10.1111/j.1813-6982.2009.01195.x

Smith, L.C. & Edwards, L., 2007, 'A multivariate evaluation of mainstream and academic development courses in first-year microeconomics', *The South African Journal of Economics* 75(1), 99–117. https://doi.org/10.1111/j.1813-6982.2007.00102.x

Spencer, B., 1998, 'Responding to student writing: Strategies for a distance-teaching context', D. Litt thesis, University of South Africa.

Van der Merwe, A., 2006, 'Identifying some constraints in first year economics teaching and learning at a typical South African university of learning', *The South African Journal of Economics* 74(1), 150–159. https://doi.org/10.1111/j.1813-6982.2006.00055.x

Van der Walt, J.L., 2013, *Advanced skills in academic literacy*, Andcork, Potchefstroom.

Van Walbeek, C., 2004, 'Does lecture attendance matter? Some observations from a first year economics course at the University of Cape Town', *The South African Journal of Economics* 72(4), 861–883. https://doi.org/10.1111/j.1813-6982.2004.tb00137.x

Wentzel, A., 2017, *A guide to argumentative research writing and thinking: Overcoming challenges*, Routledge, London.

Williams, E.F., Dunning, D. and Kruger, J., 2013. 'The hobgoblin of consistency: Algorithmic judgment strategies underlie inflated self-assessments of performance,' *Journal of Personality and Social Psychology*, *104*(6), p.976

Yani, B., Harding, A. & Engelbrecht, J., 2019, 'Academic maturity of students in an extended programme in mathematics', *International Journal of Mathematical Education in Science & Technology* 50(7), 1037–1049. https://doi.org/10.1080/0020739X.2019.1650305