

Semantic tone of research ‘environment’ submissions in the UK’s Research Evaluation Framework 2014

Andy Thorpe, University of Portsmouth

Russell Craig, University of Portsmouth

Glenn Hadikin, University of Portsmouth

Sasa Batistic, Tilburg University

Abstract

This paper applies *DICTION* computer-assisted text analysis software to evaluate the tone of research ‘Environment’ submissions by Business and Management Studies schools in the UK’s 2014 Research Evaluation Framework. We find that submissions contain distinctive differences in semantic tone between high-ranked and low-ranked universities, particularly in terms of *DICTION*’s master variable, *ACTIVITY*. The language of high-ranked institutions has a tone of *low* *ACTIVITY*, whereas the language of low-ranked institutions has a tone of *high* *ACTIVITY*. More adjectives are used than expected: by high-ranked universities to bolster strong public reputations; and by low-ranked universities to atone for weaknesses. High-ranked universities are advantaged because they are more likely to be represented on assessing panels and be better-attuned to reader expectations. The results suggest that low-ranked universities could have achieved higher scores by reflecting on particular areas of word choice and the potential effects of those choices on assessors.

‘People use words to make impressions on other people. It has always been thus’
(Hart, Childers and Lind 2014: 3)

Introduction

This paper explores the semantic tone of formal submissions by universities to government administrators. Prior studies have analysed the leadership communication of university leaders with stakeholder groups (e.g. Fortunato, Gigliotti and Ruben 2017); the use of social media by universities (e.g. Naidoo and Dulek 2016); and the relationship between the

language of university mission statements and performance (Short and Palmer 2008). Myers (1991; 1993) explored how universities presented their case for government funding to establish a research centre. His analysis on how the writing of the submissions was shaped, and the effect of particular words, was based principally on a close reading of the proposals and related documents, and interviews with proposal writers. Authors (2017) adopted a similar approach in applying an ‘impression management lens’ and close reading analysis to scrutinise how university submissions to the UK Research Excellence Framework 2014 (REF2014) were crafted, and whether stylistic differences in submissions could be linked to published outcomes.

Our study extends the work of Author (2017) by using computer assisted text analysis (CATA) techniques to explicitly explore the semantic tone of research ‘Environment’ submissions made to REF2014. CATA techniques have many potential benefits. Hart, Childers, and Lund (2013: 13), for example, describe a wide range of practical uses of CATA techniques and contend that their general advantages reside ‘in the ability of computers to remember, detect continuities and discontinuities, track associations across semantic space, and note characteristic word choices from one person to another.’ Hart (2015: 155) makes a case that CATA techniques enable scholars to bring ‘granularity’ to their analysis, and to

... go beyond simple dichotomies (good/bad, happy/sad) when describing a text. The computer-using scholar can now examine multiple texts simultaneously, using the semantic web to shed light on their properties ... computer programs like *DICTION* attack a text from forty or fifty different angles simultaneously ... telling the user when a text conforms to, or deviates from, a set of norms.

DICTION is a particularly useful CATA technique because it assumes that words mean things, and that repetitive patterns of words over time mean even more than isolated instances.¹ *DICTION* assumes that word choices are an important factor in any kind of public discourse; that they aggregate into patterns of meaning for readers and listeners; and that the

¹ Other quantitative CATA programs that explore word counts or word co-occurrence, with or without custom dictionaries, include *VBPro*, *Yoshikoder*, *WordStat*, *General Inquirer*, *Profiler Plus*, *LIWC*, *PCAD*, *WORDLINK* and *CATPAC*.

patterns of words accumulate in readers' minds to form identifiable tonalities (such as, 'too slick,' 'rather bookish,' 'warm and genuine') (Lowry 2008: 485).²

The focus of our analysis is REF2014, one of the most costly³ research assessment exercises conducted in the UK. REF2014 covered the period January 2008 to December 2013 and was designed to help the four UK higher education funding bodies allocate about £2 billion of research funding per year among 154 universities. In making their submissions, universities could decide which research and researchers they wanted to submit, and which Units of Assessment (UoA) to submit to. This led to accusations of 'game-playing' as universities sought to maximise their funding allocation and/or their grade point average (GPA), and hence their ranking. The 1,911 written submissions across 36 UoAs were assessed by a panel of peers in terms of 'Outputs' submitted (65% of overall assessment score), 'Impact' of research undertaken (20% of score), and the 'Environment' in which the research was undertaken (15%). A score of '4' denoted an environment was 'world-leading', '3' was 'internationally excellent', '2' was 'recognised internationally', and '1' was deemed 'nationally recognised.'

In effect, the submissions were exercises in persuasion directed at assessors. We conceive the submissions as part of 'the extraordinary amounts of time, effort and money [universities spend] developing the "right" image, and shaping others' impressions and expectations...' (Alvesson and Kärreman 2017: 118). In addition to their 'gameplaying attempts by creating facts' (e.g., submitting some researchers but not others), the submissions also engaged in 'gameplaying by wording.'⁴ Our use of *DICTION* software to analyse the semantic tone of 'Environment' submissions contributes to development of a deeper

² For a compact summary of the key features of *DICTION* see Murphy (2013: 60-61). For more detailed explanations of *DICTION* see Hart and Carroll (2014), Hart et al. (2014), and Amernic et al. (2010). Fuller details regarding *DICTION* can be found too at www.dictionsoftware.com. This website also contains an exhaustive list of scholarly publications that have used *DICTION* across a wide range of disciplines.

³ Stern (2016: 6) estimated the total cost of REF2014 at approximately £246 million, with £212 million borne by the submitting institutions. This was 133% higher than for the preceding research assessment exercise in 2008.

⁴ We are indebted to an anonymous reviewer for highlighting this point.

understanding of research evaluation methods and outcomes, including the possibility of better appreciating the incidence and features of ‘gameplaying by wording.’

There has been a relatively low level of scholarly scrutiny of the ‘Environment’ component of REF2014 (exceptions are Pidd and Broadbent [2015]; Mellors-Bourne, Metcalfe and Gill [2017]; and Author [2017]). This is surprising for two reasons. First, because of the very large sums of public monies that were allocated to universities solely as a consequence of their ‘Environment’ score; and second, because ‘Environment’ submissions are conducive to scholarly enquiry by virtue of their conformity to a pre-specified stylistic narrative format and a pre-determined permissible maximum page length.

We are motivated by belief in a strong likelihood that the semantic tone of language used in research ‘Environment’ submissions was associated with the scores awarded by assessors. This was prompted, in part, by an *ad hoc* close reading conducted by one of the present authors, of a sample of four REF2014 ‘Environment’ submissions to UoA 19, Business and Management Studies (by Cardiff, Portsmouth, Plymouth and Westminster). His tentative conclusion was that rankings would have been more favourable if submissions made greater use of ‘action verbs’, preferred active voice to passive voice, and preferred present tense to past tense. That is, the rankings would have been higher if the ‘tone’ of the submission was one of *activity*, and reflected *present concerns*.

Our decision to explore semantic tone was strengthened too by findings of a study involving a qualitative close reading of REF2014 submissions in UoA 4 for Psychology, Psychiatry and Neuroscience (Blank et al. 2015, Part II). That study drew a strong link between the ambient tone created by reading a submission, and the ranking obtained; and it distinguished strongly between submissions with the tone of an ‘academic narrative’, and those with the tone of a ‘managerial bulletin.’ The authors of that study rated academic narratives to be ‘good’ because they ‘tell a story in which research groups grow organically’ and give ‘the impression of a clear, reasonable, no-nonsense, viable, healthy and sustainable approach.’ In contrast, submissions they deemed to be ‘managerial bulletins’ were said to be ‘poor’, and to ‘hardly tell a coherent story at all but essentially list managerial structures and procedures and [to be] full of mundane/ boring/ irrelevant detail and jargon.’

The present study therefore explores the ability of one type of CATA software program (*DICTION*) to identify linguistic features of the type referred to above in REF2014 ‘Environment’ submissions. In particular, the research question explored is:

Were there distinctive differences in semantic tone between high-ranked and low-ranked universities in the REF2014 Business and Management (UoA19) environment submissions?

In the following section we review literature pertaining to the semantic tone of written text, aspects of the ‘reader-writer relationship’, and how the software programme *DICTION* has been used in prior studies to explore semantic tone and nuances of written communication. We follow this by outlining details of the research method employed, including an elaboration on the features of *DICTION*, before presenting and discussing results and entering conclusions.

Semantic Tone

Semantic tone affects how people perceive others (Hart et al. 2013: 6). In investigating semantic tone, we are mindful that ‘there are few words in the English language more mysterious than *tone* ... [and that it] is an omnipresent if ill-defined concept (Hart et al., 2013: 5). Nonetheless, as with Hart et al. (2013: 9), we view tone as a device that (sometimes unwittingly) ‘create[s] distinct social impressions via word choice.’ For the present study, the social impression of a university’s research ‘Environment’ submission is deemed to be reflected in the REF2014 score awarded for that submission.

In the present context of universities becoming increasingly corporatized, market-driven and highly concerned with such issues as ‘branding’ (Huzzard, Benner, and Kärreman 2017), it is also pertinent to consider the literature on impression management. Tedeschi (2013) provides a useful review of studies over the last 50 years, from a social psychology perspective, by discussing how social actors interact with society to protect their self-image. This is, in effect, what REF2014 ‘Environment’ submissions were engaged in: the use of

written texts to manage the image of an academic institution so as to maximise the score awarded to the institution for that submission.

The authors are aware of only one (recent) example of work where semantic tone has been used to assess written text in an academic evaluation setting. Mellors-Bourne et al. (2017) explored CATA techniques to explore how equality and diversity issues were addressed in REF2014 environment statements. They found a positive relationship between REF2014 scores and references to key equality and diversity terms within submissions. There is, however, some related work in business settings. Henry (2008), for example, discusses how language choices affect earnings press releases and highlights the important effect of the communications genre in which a written text is presented. She frames her discussion in terms of the overall reason a text is written and the different objectives of authors. Henry highlights the care that should be taken to position positive items of information and repeat key points. She discusses the claim that passive voice is more likely in management reports of firms in financial decline (as suggested by Thomas, 1997). In the context of scholarly writing (and presumably in the REF2014 submissions analysed here *about* scholarly research), Sigel (2009: 479) argues that passive voice ‘bogs down the narrative’ and ‘makes for imprecise arguments.’

An important concept that can help in understanding tone is the *reader-writer relationship*. Jameson (2004: 227) describes this in terms of how a writer can visualise the reader and adopt the reader’s viewpoint to maximise effect by ‘artfully interweaving multiple rhetorical and linguistic elements.’ One way of exploiting the ‘reader-writer relationship’ for positive effect is by adopting a *you-attitude*: that is, by writers ‘intentionally subordinat[ing] their priorities to those of readers ...’ (Jameson 2004: 228). Modifiers, intensifiers, ‘vivid expressions’ and adjectival structures are also important when analysing tone (Jameson 2004: 243). In regard to adjectival structures, Edo Marzá’s (2011) study of the use of adjectives on hotel websites shows that general forms of appraisal (such as *special*, *perfect* and *flexible*) comprise the largest category of evaluative adjectives.

Sydserrff and Weetman (2002), Short and Palmer (2008), Amernic, Craig, and Tourish (2010), and Craig and Amernic (2015), among others, have analysed semantic tone in written communication using *DICTION* text analysis software in business settings. The use of *DICTION* is consistent too with lexical priming (Hoey 2005) and lexical selection (Hadikin

2015) in the sense that passages of text are being used in a context where the writer expects them to bring a social benefit. In this vein, Sydserff and Weetman (2002), for example, relate the language of the chairman's report and manager's report in each of 26 investment trusts to findings of studies in impression management (e.g. Steinbart 1989). They partitioned trusts into 'good performers' and 'bad performers' and reported that words leading to high scores for the *DICTION* master variable *ACTIVITY*⁵ were associated with poor performers. Sydserff and Weetman (2002) suggest that poor performing trusts sought to present themselves as stronger performing trusts by making linguistic choices that highlighted forward-looking and progressive activities. The corollary to this finding (which we discuss later) is that low scores for *ACTIVITY* are associated with good performers.

Short and Palmer (2008) combined *DICTION* text analysis with human-scored content analysis to investigate mission statements of university business schools or faculties accredited by the Association to Advance Collegiate Schools of Business (AACSB). They also compared the mission statements of public colleges versus private colleges, and explored whether there was any association between the content of mission statements and performance. They concluded that better performing private colleges were associated with three factors: positive entailments or endorsements of specific people and groups; a higher use of adjectives in structures such as 'highest quality' and 'brightest business students'; and simpler language.

Amernic et al. (2010) highlighted the power of the annual letters to shareholders of a charismatic leader, Jack Welch (when CEO of General Electric, 1980 to 2000). Welch's linguistic choices resulted in high scores for four of *DICTION*'s five master variables (*CERTAINTY*, *OPTIMISM*, *ACTIVITY*, and *REALISM*).⁶ The implicit implication is that in promotional business texts, strong performers choose language that stands out, is conducive to enhancing corporate relations with stakeholders, and will 'separate them from the crowd.' We contend that the findings reported in the literature reviewed above are highly relevant in the academic context, given the increasing corporatization of the modern

⁵ *ACTIVITY* is defined to be 'Language featuring movement, change, the implementation of ideas and the avoidance of inertia' (Hart and Carroll 2014).

⁶ These terms are defined in the immediately following Research Methods section.

university – a matter highlighted strongly in an edited volume of essays by Huzzard et al. (2017).

This present study of semantic tone in REF2014 ‘Environment’ submissions is therefore important because of the insights provided to the culture of universities and their attitudes to research. Those insights have the potential to reveal a tone of ‘self-assured smugness and confidence’, [a] lack [of] humility’, or a matter of fact[ness] – as was found in the context of the language use by CEOs of major companies (Amernic et al., 2010: ix). A case can be made too that the ambient tone of submissions affected the reader-writer relationship, and ensuing assessments. We enter no judgement on this, but contend that REF2014 ‘Environment’ submissions had the ability to portray accurate (or, alternatively, misleading) institutional attitudes and values to the broader community, including to REF2014 assessors.

Research Method

REF2014 reviewed the research outputs of 52,061 academics across 36 UoAs in UK universities. Here we analyse the publicly available corpus of 98 research ‘Environment’ submissions for UoA 19: Business and Management Studies,⁷ where 3,602 staff were put forward for assessorial scrutiny, together with 12,204 outputs and 432 impact case studies. The submissions were downloaded from <http://results.ref.ac.uk/Results>. They comprised 565,553 words, with an average length of 5,770 words.

The permissible length of each ‘Environment’ submission was determined by the number of full time equivalent (FTE) staff submitted. This ranged from seven pages (< 15 FTE submitted) to fifteen pages for the Universities of Manchester and Lancaster (each submitted 122 FTE). The submissions comprised an ‘Overview’ and four other sections addressing ‘Research Strategy’, ‘People’, ‘Income Infrastructure and Facilities’, and ‘Collaboration and Contribution to the Discipline.’ There was no limit on the length of any sub-section, provided the aggregate length limit (based on FTE submitted) was complied with.

⁷ ‘Environment’ sub-profiles were not publicly available for Buckinghamshire New University, University of Cumbria, and University of South Wales. They all submitted 3 or fewer staff and public release was deemed likely to enable inferences to be drawn about the quality of outputs of individual submitted staff.

Published assessments did not disclose individual component scores, but reported, for example, that for ‘Institution X’, 50 per cent of the submission was graded at 4*, 37.5 per cent at 3*, and 12.5 per cent at 1*. The results were summarised widely in a standardised ‘Grade Point Average’ (GPA). In the example just mentioned, the GPA is 3.25 ($[4 \times 0.50] + [3 \times 0.375] + [1 \times 0.125]$). As with Author (2017), we ranked the submissions according to their GPA score for ‘Environment’ (from best to worst) and then allocated them to quartile groups, designated as Q1, Q2, Q3 and Q4 (see Table 1). To resolve deadlocks arising from an equal GPA, the higher(est) rank was assigned to institutions submitting the larger(est) number of FTE staff.

TABLE 1 ABOUT HERE

Five higher education institutions [HEI] (Lancaster, LSE, Cardiff, Strathclyde and Cambridge) scored a maximum of 4. Only two small HEI (Sunderland and York St. John) scored less than one. We used *DICTION* (version 7.0) to explore the semantic tone in the 98 submissions. Below, we describe the main features of *DICTION*. We also outline some of the scholarly inspiration underlying its master variables; and highlight its assumptions, benefits and limitations.

Roderick P. Hart is the deviser of *DICTION*. He describes *DICTION* as :

... us[ing] some ten thousand search words apportioned across thirty-three word lists or dictionaries. It includes several calculated variables as well. None of the search terms is duplicated in these lists, giving the user an unusually rich understanding of a text. The program also produces five master variables by combining (after standardization) the subaltern variables. These master variables include *CERTAINTY* (indicating the resoluteness of a text), *OPTIMISM* (the endorsement of some person, group, or experience), *ACTIVITY* (movement, change, or the implementation of ideas), *REALISM* (words describing tangible, everyday matters), and *COMMONALITY* (language highlighting a group’s values and commitments). In essence, *DICTION* uses lexical layering to account for tone, something that becomes more identifiable when word families are comingled. (Hart 2015: 157)

Hart (2001: 45-46) describes the output arising from applying *DICTION* as comprising:

... frequency scores for each variable [rated] as being within, above or below a normal range. This range is calculated on a text type which the researcher chooses as comparable to the one under analysis. There are six broad classes of text types: Business, Daily Life, Entertainment, Journalism, Literature, Politics and Scholarship. These classes are further subdivided into thirty-six individual text types, representing both speech and writing.

The scholarly inspiration for each of *DICTION*'s five master tonal variables can be paraphrased from Hart et al. (2013: 14-15) as:

- CERTAINTY 'derives from the work of general semanticists, particularly Wendell Johnson (1946)';
- OPTIMISM was 'inspired by James David Barber's *Presidential Character* (1992)';
- ACTIVITY is 'indebted to the work of Osgood, Suci and Tannenbaum (1957)';
- REALISM 'taps into the pragmatism John Dewey (1954) found endemic to Western experience'; and
- COMMONALITY 'draws on the social theorizing of Amitai Entzoni (1993) and Robert Bellah (Bellah et al. 1991).'

DICTION's analysis of these five major types of tone has been found to be effective and reliable in discerning tonal qualities, positive psychological motifs, and positive residual feelings (Lowry 2008: 485).

In this study we are particularly interested in differences between the semantic tone of submissions of universities ranked in the upper quartile (1st to 24th) and those ranked in the lower quartile (76th to 98th). Our initial focus was on the five 'master variables' that indicate the general semantic tone of a text. We selected the *DICTION 7.0* processing option that analyzes a block of text averaged into sections of 500 words each. As our normative word-list referent, we chose the Corporate Public Relations dictionary from the 36 custom dictionaries available. This dictionary seemed to offer the closest fit to the impression management agenda of REF2014 submissions that was elaborated by Author (2017). Almost all of the 36 available specific comparator dictionary norms available for selection are clearly inappropriate for use here.⁸ The choice of the Corporate Public Relations dictionary was

⁸ These included those titled: Legal Documents, TV Advertising, Computer Chat Lines, Religious Sermons, Telephone Conversations, Celebrity News, Entertainment Reviews, Music Lyrics, Sports News, TV Comedies,

influenced by the recent conflation of the corporate and academic management of universities and the considerable resources university marketing and media departments devote to promoting the corporate/university brand⁹ and/or protecting institutional reputation (Mount and Belanger, 2004; Huzzard et al., 2017). We acknowledge that the Corporate Public Relations dictionary is not a perfect match to the context being discussed, but we contend that it provides a very good starting point and a baseline for future research.

We also acknowledge the general disadvantage of CATA techniques in treating texts as linear, complete, and independent of a reception community (Hart 2015: 154). We are mindful too of the imperfections of *DICTION*. These are reported results should be read with regard for *DICTION*'s basic assumptions of transformativity (that is, it is sensible to quantify language); additivity (more uses of a term have more effect than fewer uses); semantic independence (words have meaning independently of the context in which they are used); evanescence of context (texts and their interpretations change over time); and that the nuances of messages are ignored (see Hart 2001: 52-54; Hart et al. 2014: 13-18).¹⁰

Results and Discussion

Results from processing the 98 submissions using *DICTION* are displayed in Table 2. The content of each of the three panels in Table 2 is discussed, in turn.

TABLE 2 ABOUT HERE

Panel A

Panel A reveals that the mean scores for the master variables OPTIMISM, CERTAINTY, REALISM and COMMONALITY are within normal range (of ± 1 standard

TV Dramas, Letters-to-the-Editor, Poetry & Verse, Theatre Scripts, Social Movement Speeches, Philosophical Essays, Science Writing, and Student Essays.

⁹ An extensive and growing literature (e.g. Furey, Springer and Parsons 2014) has examined the motives and consequences of corporate-style branding in the HEI sector.

¹⁰ For a response to these criticisms see Hart et al. (2013: 16-17).

deviation from the expected mean) in all quartiles, with one exception (CERTAINTY in Q2, non-significant). The mean scores for ACTIVITY are all beyond the normal range (low) for each quartile. However, the only significant out-of-range mean score for any master variable is for (low) ACTIVITY in Q1 (5% level, two-tailed).

The fact that language indicative of ACTIVITY was strongly under-represented in Q1 is surprising. This is notwithstanding the corollary (described earlier) to the findings of Sydsæff and Weetman (2002): that low ACTIVITY is associated with good performers (including here, ostensibly, in terms of a good quality institutional research ‘Environment’). Nonetheless, intuitively, one would expect language reflecting ‘low activity’ would not indicate ‘vitality’ (one of the two major criterion for assessors) and would not be indicative of a top-rated university research environment. One explanation for this is that the low ACTIVITY scores in Q1 submissions arise because they came predominantly from long-established and ‘settled’ institutions – ones that knew their strategic direction and were comfortable with their identity, research agenda, and staffing profile. (The mean age of Q1 universities was 168 years). Generally, such institutions are not seeking to make major change. Most have an established niche in the higher education environment and are not in the process of making disruptive major transformations of their research priorities, agendas, and staffing profiles. This ‘steady state’ stands in contrast to the situation in ‘newer’ universities where it is common to read of recent or impending implementations of new research plans, institutional re-structurings, and high levels of staff turnover. (The mean age of Q4 universities was 26 years). Thus, it is understandable that the language of the longer-established universities reflected stability and a tone of ‘low activity’.

Consistent with such a view is the fact that by far the *lowest* score for ACTIVITY (of 26.35) was recorded by the university ranked in Table 1 as having the best research ‘Environment.’ This was the University of Lancaster (established by Royal Charter in 1964) – a university 50 years old at the time of REF2014. The three *highest* scores for ACTIVITY were grouped tightly from 50.95 to 51.05. These were for Staffordshire, Greenwich, and the University of the West of England — all of whom were granted university status in 1992.

We explored some of the differences in ACTIVITY of the first two sections (‘Overview’, and ‘Research Strategy’) in the ‘Environment’ submissions of Lancaster and Greenwich. These two sections comprised 4303 words for Lancaster and 3421 words for

Greenwich. Lancaster used many more ‘neutral’ verbs, such as ‘to be’ and ‘to have.’ Lancaster’s submission was not as ‘energetic’ as that of Greenwich: in parts, it simply became a tedious (often) ‘verbless’ managerial bulletin of staff achievements. Greenwich’s higher score for ACTIVITY seems to have arisen from much more frequent use of verbs (and their derivatives) such as ‘encourage’ (Greenwich 9 uses, Lancaster 1); ‘increase’ (Greenwich 12 uses, Lancaster 2); and ‘provide’ (Greenwich 13 uses, Lancaster 6). Additionally, the Microsoft Word grammar check detected 34 instances of passive voice in the Lancaster submission (1 per 127 words) but only 20 in the Greenwich submission (1 per 171 words). Consistent with Sigel (2009), this suggests that the Lancaster narrative was ‘more bogged down’ and less ‘active’ than Greenwich’s.

In the Q1 database of 175,925 words, the word ‘is’ occurs 1506 times (8.56 times per 1000 words). In the Q4 dataset the corpus of 110,226 words, the word ‘is’ occurs 974 times (8.83 times per 1000). Superficially, the results are similar. However, nuanced differences can be hidden by these raw data. For example, the expression ‘*culture is*’ helps to clearly identify ambient research culture. This expression appears in Q1 seven times in sentences such as: 1) Our research *culture is* inclusive, interdisciplinary and international; and 2) Our *research culture is* one that encourages and supports inter-disciplinary work. However, the expression ‘*culture is*’ is not used at all in the Q4 dataset.

Panel B

In Panel B of Table 2 we seek to further understand what influenced the low ACTIVITY scores reported in Q1 (and to a lesser extent in Q4) by exploring the seven sub-component variables comprising the ACTIVITY master variable. Of the 28 reported scores (seven sub-component variables across four quartiles), 11 lie outside the normal range. However, only two are significant extremes. Both are for the calculated variable, EMBELLISHMENT. These occur for Q1 institutions (0.1% level) and Q4 institutions (1% level).

EMBELLISHMENT is defined as ‘A selective ratio of adjectives to verbs ... that ... slows down a verbal passage by de-emphasizing human and material action’ (Hart and Carroll 2014: 11). This concept of embellishment is based on the work of Boder (1940). He argued that ‘precision’ (presumably a positive characteristic of a good research environment, and consistent with the ‘academic narrative’ stereotype described by Blank et al. 2015) is

accomplished ‘to a large extent through [an increased frequency of] the adjective’ (p. 328). Therefore, it is not surprising that better-ranked submissions make ‘heavy use of adjectival constructions’ (Hart et al. 2014: 153) to evoke an air of precision, whilst taking the opportunity to ‘embroider claims’ (p. 146), at the margin.

In elaborating on their achievements, submission writers in Q1 and Q4 embellished their claims with a ratio of adjectives to verbs that was significantly higher than expected according to *DICTION* norms. In the case of Q1 institutions, the adjectives used had a distinctive semantic tone evoking a ‘safe’, ‘solid’ and ‘established’ research environment. The epitome of this was Lancaster University Management School [LUMS]. Its submission commenced by highlighting that ‘after achieving outstanding ratings in the previous 4 RAEs, the research activity of LUMS has continued to grow and develop in a sustainable manner.’ In terms of the reader-writer relationship, this suggests the submission writer is visualising panel members will respond positively to such elaboration.

We explored the LUMS submission more closely, as it also had the highest EMBELLISHMENT score (11.63, very highly significant). We sought indicative examples of adjectival ‘embroidery’ by conducting a close reading of the principal narrative or ‘non-managerial bulletin’ sections of its submission (Sections 1 and 2, dealing with ‘Overview’ and ‘Research Strategy’: 4300 words), In particular, we searched for adjectival uses of the words ‘leading’, ‘world-leading’, ‘(very) significant’, and ‘high*/high-profile.’¹¹ Results are reported in Table 3.

TABLE 3 ABOUT HERE

LUMS has a reputation for being a high quality business and management school. It has many achievements of which to be proud. Three of the 37 adjectival uses reported in Table 3 should be recognized as ‘aspirational’ rather than as embroidered claims of current achievements (e.g., aspiring to be ‘world leading’ in research activity, aspiring to employ ‘world leading’ staff). The remaining 34 all bolstered performance claims, with almost all stated as matters of fact. Yet, in respect of many of these claims, no substantiating evidential

¹¹ The asterisk denotes words with this ‘stem.’ Thus, ‘high*’ refers also to ‘higher’ and ‘highest.’

source or support is proffered. For example, a staff member is said to ‘be listed as one of the top 2% of economists worldwide’, but the list in which this occurs is not cited. Readers seem to be expected to accept such claims unquestioningly. Perhaps this arises because the writer can ‘visualize’ the predisposition of assessors from similarly high-ranked institutions to concur with these statements, without the need for evidentiary support. Or perhaps it arises from an implicit expectation that evidentiary support for such claims is contained in the other two major components of REF submissions (‘Research Outputs’ and ‘Impact’).

Several of the adjectival descriptors used in making claims of excellence seem to be hyperbole. For example, LUMS asserts it has the ‘world leading research group in Operation Research’ and that ‘the OR group’s theoretical research makes world-leading contributions to the mathematical foundations of optimisation, forecasting and data mining, simulation methods and stochastic processes.’ This sits oddly with the view of the *QS World University Rankings by Subject 2015 – Statistics & Operational Research*¹² which ranks Lancaster as ‘between 51 and 100.’ Just how high (or low?) does one have to appear in rankings tables to claim the mantle of ‘world leading’? While the research of this group might well be outstanding in many respects, the claim to be ‘world-leading’ resonates of ‘embroidery.’

The embroidery of claims in Q4 submissions has a slightly different pattern. Leeds Beckett University (GPA = 2.0, FTE = 17, Q4, EMBELLISHMENT = 2.51, significant at 0.1%) uses ‘leading’/ (very) significant/ high* and high profile’/ and ‘world-leading’ on only 13 occasions. Several of these uses are aspirational and several others deflate performance claims (e.g., ‘significant decline’, ‘significant turnover’). Nonetheless, in the space of 245 words, Leeds Beckett uses the adjective ‘strong’ four times, variously to describe its research development, orientation of applied research practice, relationships with business and government, and capability to generate external research income. Such embroidery is more subdued than with LUMS. Nonetheless, it seems to have failed to convince assessors that these embellishments met the desired characteristics of being ‘feasible, well-considered and convincing’ (Pidd and Broadbent 2015: 8).

¹² [http://www.topuniversities.com/university-rankings/university-subject-rankings/2015/statistics-operational-research#sorting=rank+region="+country="+faculty="+stars=false+search](http://www.topuniversities.com/university-rankings/university-subject-rankings/2015/statistics-operational-research#sorting=rank+region=). A reviewer has drawn attention to rankings such as this not being based on ‘fact.’

Detailed analysis of the adjectives used highlights the importance of checking individual words. The full corpus for Q1 has 239 cases of *leading*, *world-leading*, *Europe-leading* and *leading edge* in expressions such as *leading business schools*, *world's leading companies* and *leading journals* (that is, 1.36 uses per 1000 words). There are only 76 comparable cases in the Q4 dataset (0.69 uses per 1000 words).

Panel C

DICTION defines EMBELLISHMENT in terms of four component variables: BLAME, PRAISE, PRESENT CONCERN and PAST CONCERN. To explore the reasons for this disparity between Q1 and Q4 institutions in terms of the likely effect of their adjectival embellishment on assessors, we analysed scores for these four component variables of EMBELLISHMENT. Of the 16 reported scores (four sub-component variables across four quartiles), only four (all for PRESENT CONCERN) fell outside the normal range.

We found low and significant (5% level, two-tailed) out-of-normal-range scores for PRESENT CONCERN in Q1, Q3 and Q4 submissions. PRESENT CONCERN is: 'A selective list of present-tense verbs point[ing] to general physical activity (cough, taste, sing, take), social operations (canvass, touch, govern, meet), and task-performance (make, cook, print, paint)' (Hart and Carroll 2014: 9). The submission of Keele University (GPA = 1.625; FTE = 18; Q3; and with the highest EMBELLISHMENT score in Q3 of 3.99) resonates with high activity verbs, for example, describing the individual research work by staff that variously 'takes', 'considers', 'analyses', 'explores', 'applies', 'hosts', 'studies', 'brings', 'spans', 'builds' and 'extends.' In contrast, York St John (GPA = 0.625, FTE = 7; Q4; and with the lowest EMBELLISHMENT score of 0.007) uses past tense verbs to describe their research 'Environment' ('hosted', 'established', 'developed', 'achieved', 'monitored', 'operationalized').

As shown in previous sections, it is useful to delve into detail to understand the effect of particular word choices in context. Consider the word *takes*. The frequency in the full Q1 and Q4 datasets are comparable - 14 cases (0.08 per 1000 words) and 12 (0.11 per 1000 words) respectively. A striking difference, however, is that nine out of the twelve occurrences in Q4 are part of the expression 'takes place', in examples such as 'collaboration takes place' and 'induction takes place annually.' By contrast, there are only three cases of 'takes place' in the

Q1 data set. Examples of different structures from Q1 include: ‘the school takes an active role in hosting conferences,’ ‘the school takes an interdisciplinary view of research’, and ‘the committee takes a holistic approach to developing and monitoring research activity.’ At a glance, such instances of the use of ‘takes’ may appear similar. However, the Q4 institutions make heavy use of the expression ‘takes place’ to simply say that certain events happen and how frequently. The Q1 institutions use the same verb but do so to draw the reader’s attention to their approach to research activity.

Summary

The submissions of high-ranked universities used language with a tone of low ACTIVITY, whereas lower-ranked universities used language with a tone of high ACTIVITY. Both high-ranked and low-ranked universities embellished their submissions with a high number of adjectives – but did so for different purposes. High-ranked universities sought to bolster performance claims, and engaged in hyperbole at the margin to do so. With low-ranked universities, adjective use was more likely to deflate claims, be set in an aspirational context, or atone for shortcomings. The embellishment of text was facilitated in high-ranked and low-ranked universities by the use of present tense rather than past tense.

A pure *DICTION*-based approach can, however, potentially disguise nuanced differences if not complemented by detailed reading and analysis of words in context. We recommend the use of both techniques to obtain a fuller understanding of the effects of word choice.

Conclusions

Previous analyses of government-initiated research assessment exercises across the globe have, for example, compared institutional costs and consequences of UK and Italian research assessment (Geuna and Piolatto 2016; Rebora and Turri 2013); explored the micro-politics of resistance by Czech researchers to their country’s assessment exercise (Linkova 2014); proposed methods to normalise assessment results across academic disciplines (Kenna and Berche 2011); debated whether the UKs RAE 2008 was implicated in the decline of some disciplines (Saunders, Wong and Saunders 2011); and teased out general lessons on how

performance-based university research funding can enhance understanding of research policy (Hicks 2012).

The ability of narrative to distort and/or embroider claims in research assessment exercises was highlighted by Wilsdon *et al.* (2015: 129) in a Higher Education Funding Council for England (HEFCE) report, as follows: ‘the narrative elements were hard to assess, with difficulties in separating quality in research environment from quality in writing about it.’ However, the narrative aspect of research assessment exercises has received minimal scrutiny to date in the academic literature.¹³ Nevertheless, concerns about the effect of writing quality caused the Stern Review (2016) and the subsequent HEFCE consultation exercise to propose the introduction of a ‘more structured template’ that ‘decreases the narrative elements’ (HEFCE 2016, note 112) – and instead relies on a greater use of quantitative metrics to evaluate research environments.

The present analysis using *DICTION* software provides tangible evidence that low-ranked UoAs use different language to high-ranked UoAs. This is revealed in some distinctively different semantic tone characteristics between high-ranked submissions and low-ranked submissions. In particular, and perhaps counter-intuitively, there is a much lower level of ‘ACTIVITY’ in long-established and ‘settled’ Q1 institutions. They know their strategic direction, are comfortable with their public profile, and are not generally engaged in major institutional change. Q1 institutions have a distinctive semantic tone that evokes a ‘safe’, ‘staid’, ‘orthodox’, ‘conservative’, and ‘settled’ environment that is not disturbed (unduly at least) by reform, disruption or major staff turnover. This reinforces the point of Author (2017) that top business schools engage in heavy use of self-referencing terms to benefit from their (generally highly-regarded) institutional ‘brand name.’ Low ACTIVITY corresponds to both solidity and stolidity, and reaps its reward in terms of higher GPA scores.

The findings do not imply that a Q4 institution could suddenly change to the tone of a Q1 institution and expect better REF results as a consequence. A more subtle approach is needed, based on what the writer thinks the reader already knows about their business school. Detailed analyses of the variations of tone and meaning between individual words and

¹³ Exceptions are Taylor (2011), Pidd and Broadbent (2015), and Authors (2017).

expressions may be something for Q4 institutions to reflect upon when contemplating the composition of their future submissions.

The following moot point arises as a possible focus for future research. Would REF 'Environment' submission writers have achieved maximum positive effect by visualising the panel, and crafting their submission accordingly? A plausible case can be made that the reader-writer relationship exerts some influence. In the present study, universities ranked in the upper quartile (Q1) of universities were predominantly Russell Group members¹⁴ and were much more strongly represented on the assessing panel, as Authors (2017) have reported. Thus, it would be beneficial to explore whether submission writers of Russell Group universities were advantaged in terms of the 'reader-writer' relationship. Did such an advantage arise because they were better-placed than universities which were not represented on the panel, to 'visualise the panel' and chose language that would appeal to panel members.

In the context of REF2014, determining reader priorities was not easy, despite the clear instruction to each panel assessor (reader) to assess the 'vitality' and the 'sustainability' of the research environment. How assessors might do so was largely left unsaid in the guidance criteria. Thus, unless writers knew panel members personally, or were 'kindred spirits' with them, they were in the invidious position of having to second-guess what they should prioritise in their submission to impress the assessors (Author 2017).

Our analysis reveals a general disposition of submission writers to dwell, more than usual, on past achievements, and to not focus on present, current physical activity. This is understandable given the writing task at hand. In elaborating on their achievements, most submission writers embraced 'embellishment', and used a higher than expected ratio of adjectives to verbs. This was nuanced in the case of Q1 and Q4 institutions, yet it was manifest in quite distinctive ways. In the case of Q1 institutions, the adjectives seemed to evoke an air of precision and to bolster the universities' high public reputation. In contrast, Q4 universities seemed to use adjectives to disguise weaknesses and atone for inadequacies.

The narrative text of research assessment exercises can have critical outcomes. In REF2014, for example, Lancaster's top ranking delivered a dividend (for research

¹⁴ The Russell Group, comprising 24 universities claims to represent 'research intensive world-class universities.' Membership details can be found at: <http://www.russell-group.ac.uk/>

‘Environment’ in the Business and Management UoA alone) estimated at approximately £620,000 per annum. This is a handsome return for just 15 pages of narrative. But is there a better non-narrative alternative? While aspects of the research environment of a UoA can be quantified by metrics such as levels of external research income generated or the number of postgraduate research (PGR) students, who is to say what is ideal? Is a low ratio of PGR students to staff members associated with an arid research environment? Does a high PGR to staff member ratio limit the extent of staff-student interchanges?

Quantitative performance metrics are unable to capture the richness and complexity of interactions between physical infrastructure, institutional codes and regulations, and the academics (and non-academics) employed therein. We concur with the Chair and Deputy Chair of UoA19, Pidd and Broadbent (2015: 579, 575) that quantitative information can only function as ‘a crude indicator or overall activity;’ and that it was better to make judgements based on: ‘Does this sound like a great place to work, in which senior and junior researchers should thrive.’ If some form of narrative is to be retained in future research assessment exercises (as seems likely in REF2021 in the UK) then there appears to be a compelling case for enlisting business communications experts to devise the assessment framework. Such experts could be commissioned to help minimise (or perversely, maximise) linguistic game-playing, and to advise how the subtleties of narrative might be explored more effectively in evaluating whether a UoA had the desired characteristic of being ‘a great place to work.’

References

- Alvesson, M., and Kärreman, D. (2017) ‘Uncreative Destruction: Competition and Positional Games in Higher Education.’ In Huzzard, T., Benner, M., and Kärreman, D. (2017) *The Corporatization of the Business School: Minerva Meets the Market* (pp. 111-127). London: Routledge.
- Amernic, J., Craig, R., and Tourish, D. (2010) *Measuring and Assessing Tone at the Top Using Annual Report CEO Letters*, Institute of Chartered Accountants in Scotland, Edinburgh.
- Authors. (2017)

- Barber, J.D. (1992) *The Presidential Character: Predicting Performance in the White House*. Englewood Cliffs, NJ: Prentice Hall.
- Bellah, R.N., Madsen, R., Sullivan, W.M., Swidler, A., and Tipton, S.M. (1991) *The Good Society*. New York: Kopf.
- Blank, H., Hope, L., López, B., Ost, J., Reddy, V. and Stafford, L. (2015) 'Environment analysis REF 2014 for UoA 4 (Psychology)', Report to the Faculty of Science, University of Portsmouth. Available on request to the first author.
- Boder, D.P. (1940) 'The Adjective-verb Quotient: A Contribution to the Psychology of Language', *Psychological Record*, 3: 310-343.
- Craig, R., and Amernic, J. (2016) 'Are There Language Markers of Hubris in CEO Letters to Shareholders?' *Journal of Business Ethics*, doi:10.1007/s10551-016-3100-3
- Derrick, G., and Samuel, G.N. (2016) 'The Evaluation Scale: Exploring Decisions About Societal Impact in Peer Review Panels', *Minerva*, 54: 75-97.
- Dewey, J. (1954) *The Public and its Problems*. Chicago: Swallow Press.
- Edo Marzá, N. (2011) 'A Comprehensive Corpus-based Study of the Use of Evaluative Adjectives in Promotional Hotel Websites', *Odisea: revista de estudios ingleses*, 12: 97-123.
- Etzioni, A. (1993) *The Spirit of Community: Rights, Responsibilities, and the Communitarian Agenda*. New York: Crown Publishers.
- Fortunato, J.A., Gigliotti, R.A., and Ruben B.D. (2017) 'Racial Incidents at the University of Missouri: The Value of Leadership Communication and Stakeholder Relationships', *International Journal of Business Communicators*. Published online January 19.
- Furey, S. Springer, P., and Parsons, C. (2014) 'Positioning University as a Brand: Distinctions Between the Brand Promise of Russell Group, 1994 Group, University Alliance, and Million+ universities', *Journal of Marketing for Higher Education* 24(1): <http://dx.doi.org/10.1080/08841241.2014.919980>.
- Geuna, P., and Piolatto, M. (2016) 'Research Assessment in the UK and Italy: Costly and Difficult, but Probably Worth it (At least for a While)', *Research Policy*, 45(1):260-71.
- Hadikin, G. (2015). 'Lexical Selection and the Evolution of Language Units', *Open Linguistics* 1: 458-466.

- Hart, R.P. (2001) 'Redeveloping DICTION: Theoretical considerations'. In M. West (Ed.), *Theory, Method, and Practice of Computer Content Analysis* (pp. 43–60). New York, NY: Ablex.
- Hart, R.P. (2015) 'Genre and Automated Text Analysis: A Demonstration.' In J Ridolfo and W. Hart-Davidson (Eds), *Rhetoric and the Digital Humanities* (pp.152-168). Chicago, IL: University of Chicago Press.
- Hart, R.P., and Carroll, C.E. (2014) *Help Manual: DICTION 7.0*, Austin, Texas, USA: Digitex.
- Hart, R.P., Childers, J.P., and Lind, C.J. (2013) *Political Tone: How Leaders Talk and Why*. University of Chicago Press, Chicago, IL.
- HEFCE, (2016) *Consultation on the Second Research Excellence Framework*. London: HEFCE.
- Henry, E. (2008) 'Are Investors Influenced by the Way Earnings Press Releases are Written?' *Journal of Business Communication*, 4 (45): 363-407.
- Hicks, D. (2012) 'Performance-based University Research Funding Systems', *Research Policy*, 41(2): 251-61.
- Hoey, M. (2005) *Lexical Priming: A New Theory of Words and Language*. London: Routledge.
- Huzzard, T., Benner, M., and Kärreman, D. (2017) *The Corporatization of the Business School: Minerva Meets the Market*, London: Routledge.
- Jameson, D.A. (2004) 'Conceptualizing the Reader-writer Relationship in Business Prose', *Journal of Business Communication*, 41 (3), 227-264.
- Johnson, W. (1946) *People in Quandaries: The Semantics of Personal Adjustment*. New York, NY: Harper.
- Khazragui, H., and Hudson, J. (2015) 'Measuring the Benefits of University Research: Impact and the REF in the UK', *Research Evaluation*, 24: 51-62.
- Kenna, R., and Berche, B. (2012) 'Normalization of Peer-evaluation Measures of Group Research Quality Across Academic Disciplines', *Research Evaluation*, 20(2):107-16.
- Linkova, M. (2014) 'Unable to Resist: Researcher's Responses to Research Assessment in the Czech Republic', *Human Affairs*, 24(1):78-88.

- Lowry, D.T. (2008) 'Network TV news framing of good vs. bad economic news under democrat and republican presidents: A lexical analysis of political bias', *Journalism & Mass Communication Quarterly*, 85(3): 483-498.
- Mellors-Bourne, R., Metcalfe, J., Gill, A. (2017) *Exploring Equality and Diversity using REF2014 Environment Statements*, CRAC LTD Report to HEFCE. <http://www.hefce.ac.uk/pubs/rereports/year/2017/edinref/>
- Mount, J., and Bélanger, C.H. (2004) 'Entrepreneurship and Image Management in Higher Education: Pillars of Massification,' *The Canadian Journal of Higher Education*, XXXIV(2): 125-140.
- Murphy, A.C. (2013) 'On "True" Portraits of Letters to Shareholders and the Importance of Phraseological Analysis', *International Journal of Corpus Linguistics*, 18: 57-81.
- Myers, G. (1991) 'Conflicting Perceptions of Plans for an Academic Centre', *Research Policy*, 20: 217-235.
- Myers, G. (1993) 'Centering: Proposals for an Interdisciplinary Research Center', *Science, Technology, & Human Values*, 18: 433-459.
- Naidoo, J., and Dulek, R. (2016) 'Leading by Tweeting: Are Deans Doing It? An exploratory analysis of tweets by SEC business school Deans', *International Journal of Business Communicators*. Published online October 30.
- Osgood, C.E., Suci, G.J., and Tannenbaum, P. (1957) *The Measurement of Meaning*. Urbana: University of Illinois Press.
- Pidd, M., and Broadbent, J. (2015) 'Business and Management Studies in the 2014 Research Excellence Framework', *British Journal of Management*, 26: 569-581.
- Rebora, G., and Turri, G. (2013) 'The UK and Italian Research Assessment Exercises Face to Face', *Research Policy*, 42:1657-66.
- REF2012 'Panel criteria and working methods'. http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12_2C.pdf
- Samuel, G.N., and Derrick, G.E. (2015) 'Societal Impact Evaluation: Exploring Evaluator Perceptions of the Characterisation of Impact under the REF2014', *Research Evaluation*, 24: 229-41.

- Saunders, J., Wong, V., and Saunders, C. (2011) 'The Research Evaluation and Globalization of Business Research', *British Journal of Management*, 22: 401-19.
- Short, J.C., and Palmer, T.B. (2008) 'The Application of *DICTION* to Content Analysis Research in Strategic Management', *Organizational Research Methods*, 11(4): 727-752.
- Sigel, T. 2009. 'How Passive Voice Weakens Your Scholarly Argument', *Journal of Management Development*, 28(5): 478-480.
- Steinbart, P.J. (1989) 'The Auditor's Responsibility for the Accuracy of Graphs in Annual Reports: Some Evidence of the Need for Additional Guidance', *Accounting Horizons*, 3(3): 60-70.
- Stern, N. (2016). *Building on Success and Learning from Experience: An Independent Review of the Research Excellence Framework led by Lord Stern*. Department for Business, Energy and Industrial Strategy. <https://www.gov.uk/government/publications/research-excellence-framework-review>.
- Sydserrf, R. and Weetman, P. (2002) 'Developments in Content Analysis: A Transitivity Index and *DICTION* Scores', *Accounting, Auditing & Accountability Journal*, 15(4): 523-545.
- Taylor, J. (2011) 'The Assessment of Research Quality in UK universities: Peer Review or Metrics?' *British Journal of Management*, 22: 202-17.
- Tedeschi, J.T. (Ed) (2013) *Impression Management Theory and Social Psychological Research*. London: Academic Press.
- Thomas, J. (1997) 'Disclosure in the Marketplace. The Making of Meaning in Annual Reports', *Journal of Business Communication*. 34: 47-66.
- Wilsdon, J., Allen, L., Belfiore, E., Campbell P., Curry S., Hill S., Jones R., Kain, R., Kerridge, S., Thelwall, M., Tinkler J., Viney, I., Wouters, P., Hill, J., and Johnson B. (2015) *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. London: HEFCE.