

1 *Article*

2 **An Exploration of the Sub-Register of Chemical Engineering Research**
3 **Papers Published in English**

4
5 Myrto-Panagiota Zacharof*^a and Anna Charalambidou^b

6 ^a Centre for Cytochrome P450 Biodiversity, Institute of Life Science, Swansea University Medical
7 School, Swansea, Wales SA2 8PP, UK

8 ^b Media Department, Faculty of Arts and Creative Industries Middlesex University London, The
9 Burroughs, Hendon, London, NW4 4BT, UK

10

11 *Corresponding author : ^a Centre for Cytochrome P450 Biodiversity, Institute of Life Science, Swansea
12 University Medical School, Swansea, Wales SA2 8PP, UK myrtozacharof1981@yahoo.com,
13 M.Zacharof@swansea.ac.uk Tel: +44(0) 7455858883

14 ^bMedia Department, Faculty of Arts and Creative Industries, Middlesex University London, The Burroughs,
15 Hendon, London, NW4 4BT, UK , a.charalambidou@mdx.ac.uk Tel: +44(0) 7904365654

16

17

18 **Abstract:** The combination of increased pressures for high-volume, high-impact
19 publications in English language with the high rejection rates of submitted manuscripts for
20 publications presents an often unsurpassable obstacle for (early career) researchers. At the
21 same, the register requirements of peer-reviewed journals -that can contribute to whether a
22 paper is accepted for publication- has received little attention. This paper redresses this gap,
23 by investigating the linguistic choices in 60 published manuscripts in four journals, with
24 impact factor (IF) above 2; all 4 journals, publish original research papers in the field of
25 chemical engineering science and specifically focus on wastewater treatment. Our survey
26 shows that chemical engineering research publications tend to comply to a set of unwritten
27 requirements: multidisciplinary, brevity, co-authorship, focus on the description of
28 practical results (rather than methods), and awareness of non-specialised audiences. It is
29 found that less discipline-specific vocabulary was used in higher IF journals and this is
30 interpreted within the current context of manuscript publication and consumption. Also, a
31 complex relationship between the advertised scope of each journal and the actual published
32 papers exists, indicating that guide for authors and aims and objective published by the
33 journal's editorial office should be critically evaluated.

34 **Keywords:** chemical engineering; journal publications; lexical choices; collocations; impact factor; training

35

36 **1. Introduction**

37 Chemical engineering science is a versatile, multifaceted scientific field integrating physics,
38 mathematics, biology and chemistry. Chemical engineers employed both in the academic
39 world and in industry are called to act upon a wide variety of subjects, from pharmaceutical
40 and cosmetics fabrications, to hydrocarbons, food production and processing and
41 environmental pollution. In academia, in particular, the research activity occurring has
42 factual outputs, such as communications of various character and nature, that are
43 quantifiable; for instance patents, presentations databases, protocols and publications [1].
44 Researchers, often non native speakers, are expected to gather information, process and
45 evaluate them, take practical steps and make comments and finally communicate these
46 findings in a concise form [2,3]. The prevailing form of communication of research -and
47 therefore its certification- is scientific journal publication, while publishing in co-authorising
48 teams is now the dominant *modus operandi* [1,4]. Researchers and scientists are under
49 constant pressure to publish their results [5], as this would enhance their employment
50 prospects and career development, their funding and consultancy prospects and, on the
51 whole, their professional reputation [6,7]. Chemical engineering researchers are further
52 challenged by the multifaceted nature of their discipline, since they are called to
53 communicate their findings to a wider audience of fellow scientists, both during the
54 manuscripts' writing process and its peer review. Moreover, high rejection rates of submitted
55 manuscripts for publications have been observed, with 62% of published paper having been

56 rejected at least once [8]. Numerous reasons influence rejection, including technicalities,
57 such as limits in pages of publications per year (printed pages per issue, volume), limited
58 time between submission and publication [9], but mainly lack of clear, succinct explanation
59 of the findings and their significance to their scientific field [8] which is often attributed to
60 the use of English language.

61 Although the acceptance of a manuscript for publication is an achievement, only high-quality
62 publications in high-ranking scientific journals are widely accepted by the scientific
63 community, authors' affiliations, employment and funding bodies [39,40]. For instance, the
64 European Commission has formally recognised the importance of bibliometric indicators for
65 policy purposes and is deeply engaged in and strongly encourages scientometric analysis [7].
66 The great number of predatory publishers [10,11], the increasing rate of generation of
67 scientific findings, the globalisation of scientific communication through electronic media,
68 the different sets of regulations regarding manuscript length, peer review and evaluation have
69 contributed to the widening importance of assessing the value of a publication by (a) the
70 quality of the journal described by the journal's impact factor and (b) the individual citations
71 the publication receives [6]. A journal's impact factor, despite being continuously and
72 increasingly scrutinised [12], is the most popular numerical measure for the evaluation of a
73 scientific publication.

74 The impact factor has been originally designed as an aid to librarians all over the world, to
75 select journals that were most relevant to the public the library addresses or aims to address
76 [13, 14]. It is a ratio calculated by the total number of citations a journal receives over the
77 preceding two years divided by the total number of citations of articles published during that

78 time [15,16,38]. Nowadays, impact factors have been converted to a vital part in decision
79 making regarding scientific impact [14] influencing decisions regarding career prospects,
80 recruitment and appointments [1,16, 17]. Therefore, mastering the art of scientific writing is
81 of utmost importance for every researcher [5] since, research scientists are requested to
82 produce publications of exceptionally high standards, not only related to the novelty and
83 validity of the results presented, but also in a style that would make the manuscript a good
84 read, hence enhancing the potential of publication (by reducing editor's time) and increasing
85 its citability potential [18].

86 Despite its importance and even though some writing-related training across the curriculum
87 at student level exists [4,19,20], training scientists in publishing their research findings is not
88 an elemental part of chemical engineering education [18]. Post-doctoral researchers are
89 expected to have already obtained the skills required for formulating high quality
90 publications, presentations or talks during their postgraduate education or to have learnt by
91 osmosis, ergo reading published manuscripts from fellow researchers, a tactic that might be
92 highly ineffective, time consuming and lead to failed attempts to publish [21]. Several
93 books have been published offering guidelines for writing papers [22] in science, chemistry
94 and engineering however these give general advice on the structure the papers need to have
95 related to the analysis of experimental data of quantitative and qualitative nature without
96 focusing on the use of language [23]. In addition, there is concrete evidence of lexical
97 variation of texts within the same academic discipline, depending on the type of publication
98 (i.e. journal article, research proposal, scientific poster, textbook, popular science article)
99 and, consequently, on its intended audience (expert, scientific, student, general public)

100 [20,24-26]. However, the issue of content and register variation among articles published at
101 different types of peer-reviewed periodicals has received little attention and is a much needed
102 addition to chemical engineering education at university level and researcher development,
103 in general.

104 Hence, this paper aims to identify and investigate the linguistic choices in 60 published
105 manuscripts in four different journals of impact factor above 2. All four journals publish
106 original research papers in the field of chemical engineering science, and specifically in one
107 of its most prominent and complex subject areas, environment conservation and
108 sustainability, focusing on wastewater treatment (Fig. 1). This study explores possible links
109 (or lack thereof) between the impact factor and scope of each journal on the one hand and
110 register of the manuscripts (with a focus on lexical choices and discourse moves) , on the
111 other.. To the authors' best knowledge, register variation between different types of
112 *published, professional original research* articles has not been researched. This paper, thus,
113 aims to investigate how lexical choices and content of scientific manuscripts relate to the
114 advertised scope and impact factor of the journal, in which they are published. This can
115 contribute in helping chemical engineering researchers better adapt their papers to suit the
116 specific register of their chosen journal, so as to positively influence their publication record,
117 career prospects and attract citations and possible collaborations.

Chemical Engineering Science

Environmental Chemical Engineering: Energy, Water, Environment, Sustainability

Selection of 60 original research articles published in four high impact factor journals (IF 2-7), of wide and narrow scope, in 2012 (15 papers per journal)

Qualitative/Quantitative Analysis:

1. Quantitative analysis of the format and length;
2. Qualitative analysis of the scientific concepts of each paper and addressed audience;
3. Analysis of lexical choices (aided by ManyEyes software): (a) word frequencies of the entire corpus and (b) collocations of selected lemmas

Selection of lemmas for collocation analysis:

1. General; related to environmental chemical engineering:
'Water' and 'Waste'

2. Specific; descriptive of waste:
'Sludge' and 'Effluent'

3. Specific; related methods of treatments/results/effectiveness:
'Treatment' and 'Removal'

118

119 Fig. 1: Schematic representation of the methodology developed and followed in this case
120 study.

121 2. Materials and Methods

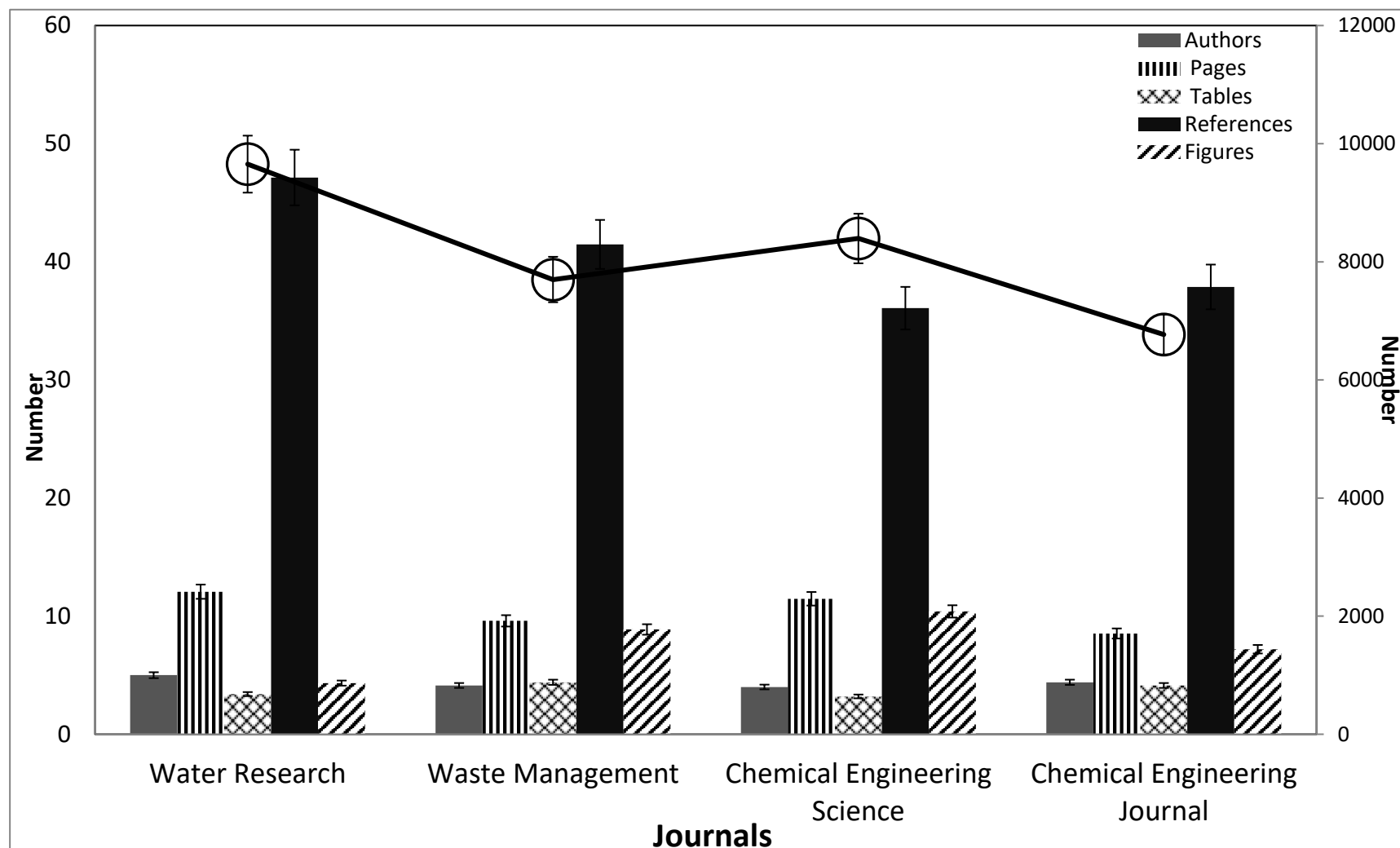
122 2.1. Materials

123 Four journals related to chemical engineering, with impact factor above two have been
124 selected, namely *Water Research (WR)*, *Waste Management (WM)*, *Chemical Engineering*

125 *Science (CES)* and *Chemical Engineering Journal (CEJ)*. The selection of the four journals
126 was based on the following criteria:

- 127 1. The topics the journal addresses, as advertised in the website of each journal, needed to
128 include environmental chemical engineering wastewater treatment and management;
- 129 2. The intended scientific audience, as advertised in the journal's website, needed to include
130 chemical engineering professionals;
- 131 3. The journal needed to have at least 15 original research articles published in 2012 focusing
132 primarily on various aspects of wastewater treatment and management, for example
133 industrial and agricultural wastewater, separation science etc.;
- 134 4. Journal's impact factor above 2, considered 2- and 3star, the quality is recognised
135 internationally in terms of originality, significance and rigour [42]

136 **CES** and **CEJ** were considered journals of wider scope; due to the great variety of scientific
137 categories within chemical engineering from which they accommodate publications (Table
138 1), while **WR** and **WM** were regarded as specialised scope due to their more concentrated
139 focus on areas relevant to environmental chemical engineering. Each journal publishes
140 various types of papers related to environmental chemical engineering and its major areas of
141 energy, water, environmental impact and sustainability (Table 1).



142

143 Fig. 2: Volumetric characteristics of the analysed published papers.

144

145 15 original papers, i.e. research-related scientific manuscripts describing, analysing and
146 discussing experimental trials and case studies were selected, of every journal totalling in 60
147 papers, all of them published in 2012. We chose to focus on volumes published 2012, as that
148 would give us a period of five years to track accumulate citations. Restraints in the type of
149 papers selected were placed to ensure a homogenous, consistent sample, in order to extract
150 meaningful results and draw useful conclusions, since the vast majority of published papers
151 in sciences, including chemical engineering, correspond to the type of factual research
152 related manuscripts.

153 **2.2. Methods**

154 A multi-layered analysis of the collected papers was devised, employing a mixture of
155 qualitative and quantitative methods as well as lexical analysis methods (Fig. 1). Quantitative
156 analysis related to the length of the papers (word counts, number of authors, references,
157 pages, tables, and figures) and was conducted in order to identify similarities and common
158 trends, using Portable Document Format (.pdf) to MS Office Word 2007 converter software
159 by freepdf solutions (www.freepdfsolutions.com). Further analysis was done using MS
160 Office Excel 2007, using linear regression analysis to obtain the average data and estimate
161 standard error and standard deviation (below < 5%).

162 Qualitative analysis of the corpus followed previously published methodologies [27,28]
163 focusing on the main scientific concepts each published manuscript was addressing. Each
164 paper was broken into clusters according to the classic practical sciences report writing style,

165 which is introduction, materials and methods, results, discussion and conclusion. Each
166 paragraph contained in the clusters was then conceptually analysed aiming at a literal
167 description, analysis and understanding of the stated research including its methodology,
168 findings, conclusions and addressed audience. Two independent examinations were carried
169 out by each of the authors to minimise each reader's subjectivity and bias regarding the
170 manuscript content. Any disagreements that arose were resolved after thorough discussion
171 among the reviewers, until a unanimous consensus was reached. Lexical variation was
172 examined through computational analysis of word association and frequencies, facilitated by
173 ManyEyes software (www-958.ibm.com). This software allows for the creation of visualisation
174 from large datasets. The following three visualisations were chosen, as they were the most
175 pertinent to the type of data (text) and research objectives (see supplementary material):

- 176 - Tag clouds: visualizations of word frequencies, which enable the researcher to see how
177 frequently a given word appears in the corpus.
- 178 - Phrase nets: This visualisation shows patterns of frequent pairs of words. Words are
179 connected when they are separated by 'and'; 'of the'; 'is', space, 'at', 'a', 'is', and 'the' in
180 the source text.
- 181 - Word trees: This visualisation enables the analyst to pick a word or phrase and shows all
182 the different contexts (i.e. immediately prior or upcoming text) in which the word or phrase
183 appears.

184 Many Eyes software can account for large amounts of text and provide accurate and fast
185 calculations, reducing researcher's bias. It can highlight the contrast between our intuitions

186 about word use and actual patterns in authentic language. An additional benefit is that it has
187 the potential of finding exceptional cases. For the analysis and interpretation of word
188 frequencies and collocations in the various journals analytical tools from corpus linguistics
189 were employed [29,30].

190 **3. Results**

191 In order to better contextualize the findings of the fine-grained analysis of the lexical choices
192 in the different journals, an overview of the format and length of the papers is provided,
193 followed by qualitative analysis of their targeted audience.

194 **3.1. Format of the collected papers**

195 In practical sciences, including chemistry, physics or engineering, manuscripts are generally
196 considered shorter in length compared to liberal sciences and arts [31]. Commonly within a
197 breadth of 6 to 12 printed two-column pages, including tables, figures and references [32] the
198 authors are expected to satisfactorily demonstrate and explain their reason for research and
199 findings. Reduction in volume and size of research papers have been implemented
200 unanimously to physical sciences journals due to the constantly increasing rate of
201 submission, leading to the need to accommodate a higher number of published papers within
202 journals printed issues or volumes [33]. Shorter length of such papers is also supported by the
203 ability of the authors to visualize their findings into meaningful figures that need little or no
204 explanation as well as reducing the amount of words and development of long, articulate
205 arguments by tabulating their core finding [34].

206 These findings are also supported in this case study. The papers' length was between 8 to 13
207 printed pages, including figures and tables, with a word count between 6800 to 9700 words

208 including references, highlights, abstracts and tables and figures legends (Fig. 2). As regards
209 to the length of the papers, similarities were found between **CES** and **WR** (average 11 pages,
210 8.300 words) and **CEJ** and **WM** (average 8 pages, 6.700 words). Cited literature serves in
211 supporting the findings and explaining the reasoning behind the trials, but also saving space,
212 as the authors are not forced to refer extensively to previously developed knowledge.
213 References in all papers ranged between 36 to 48, with similar trends found among the wider
214 scope journals **CES** and **CEJ** (on average 37 references) and the specialized scope **WM** and
215 **WR** (on average 41 references) (Fig. 2).

216 Figures and tables are the core part of the published manuscripts, varying in numbers, 4 to 10
217 figures and 3 to 5 tables, proving essential for the understanding and scientific evaluation of
218 the papers. Within that context, the text serves for analyzing, explaining and discussing these
219 visual aids to the audience. Papers in **CEJ** and **WM** were small in size, quite densely written,
220 and comprising mainly graphs and figures without analytically describing numerical results.
221 **CES** and **WR** publish longer papers with numerous figures and analytical numerical data,
222 encouraging elaboration and explanation of findings while **WR** has a balance between
223 figures, tables and discursive sections.

224 The quantitative analysis suggests that the selected papers from each of the four journals
225 share similar quantitative characteristics, thus rendering the four datasets comparable.

226 **3.2. Multidisciplinary nature of the analyzed papers**

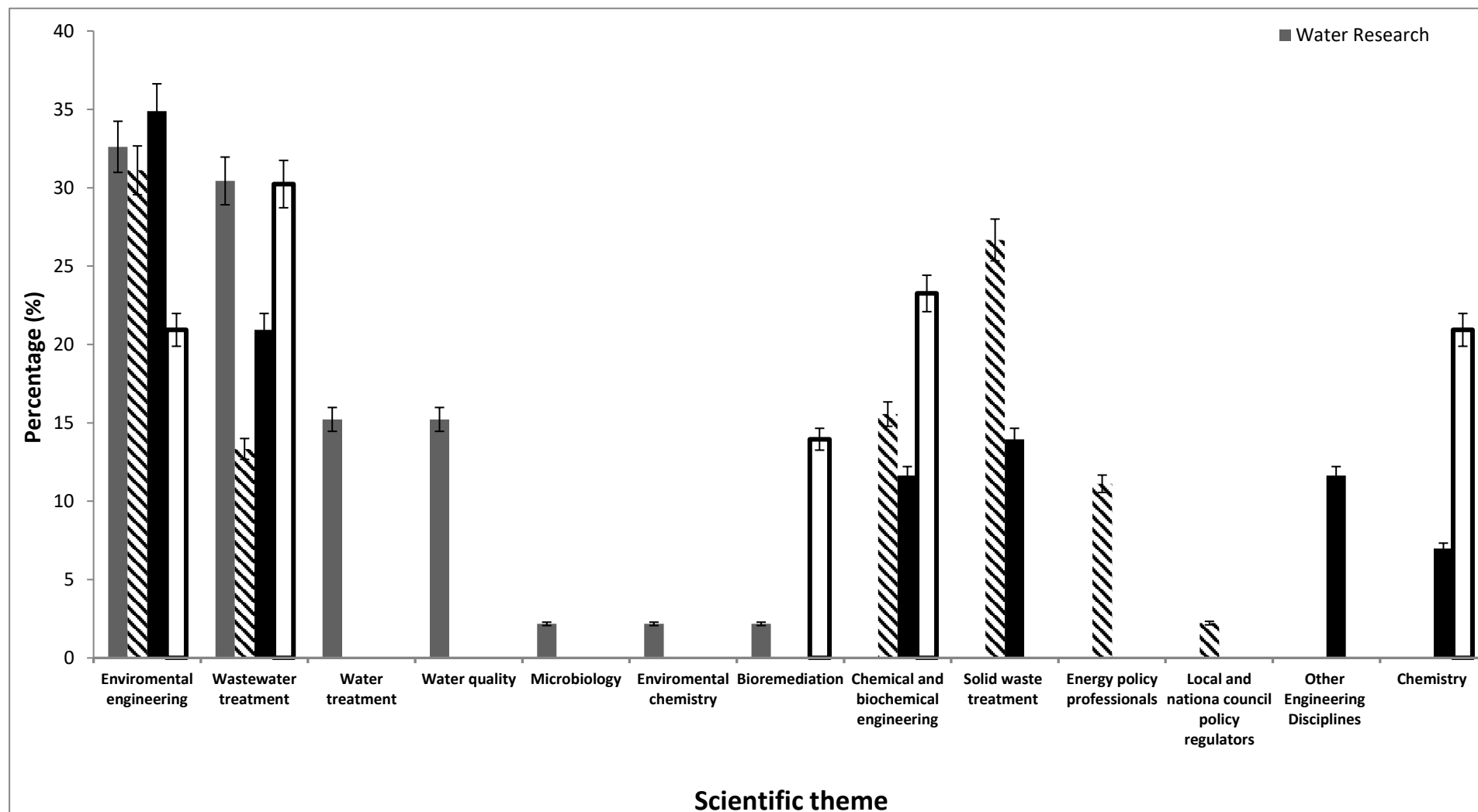
227 Despite their moderate size, all published manuscripts were the outcome of collaborative
228 efforts, with the mean number of authors being four. The multidisciplinary nature of

229 chemical engineering calls for extensive cooperation, since specialists from many disciplines
230 are required to perform the integral experimental trials to prove the scientific concept and
231 reasoning developed in the manuscripts (Ware and Mabe, 2009; White, 2006). The
232 multifaceted nature of the published papers in chemical engineering was clearly reflected in
233 this study, by the subject category (Fig. 3) and audience distribution (Fig. 4).

234 Out of the 60 papers investigated, the array of subjects of interests relevant to environmental
235 chemical engineering and specifically to waste treatment and management is wide (Fig. 3),
236 covering numerous scientific areas from biochemical engineering to environmental
237 chemistry, to other engineering disciplines such as mechanical, electrical or civil
238 engineering. The two most often-encountered areas were environmental engineering (up to
239 35%) and wastewater treatment (up to 30%) making these two (Fig. 3), while a more general
240 approach to biochemical and chemical engineering related paper was the next prevalent
241 subject area (up to 23%). Solid waste treatment (up to 26.7%) and chemistry (up to 20.93%)
242 are also covered in the journals. When compared to the advertised scientific subject of
243 interest for publication of each journal, a differentiation is found since the advertised subject
244 areas are broader to the categories that emerged from this research. The fact that journal
245 guidelines are not foolproof representations of a journal's actual remit of publications is not a
246 novel finding. What our research shows is that lexical visualisation can provide a quick way
247 for researchers to assess the specific areas that are most likely to be published in the journal.

248

249



250

251

Fig.3. Scientific theme distribution of the analysed published manuscripts.

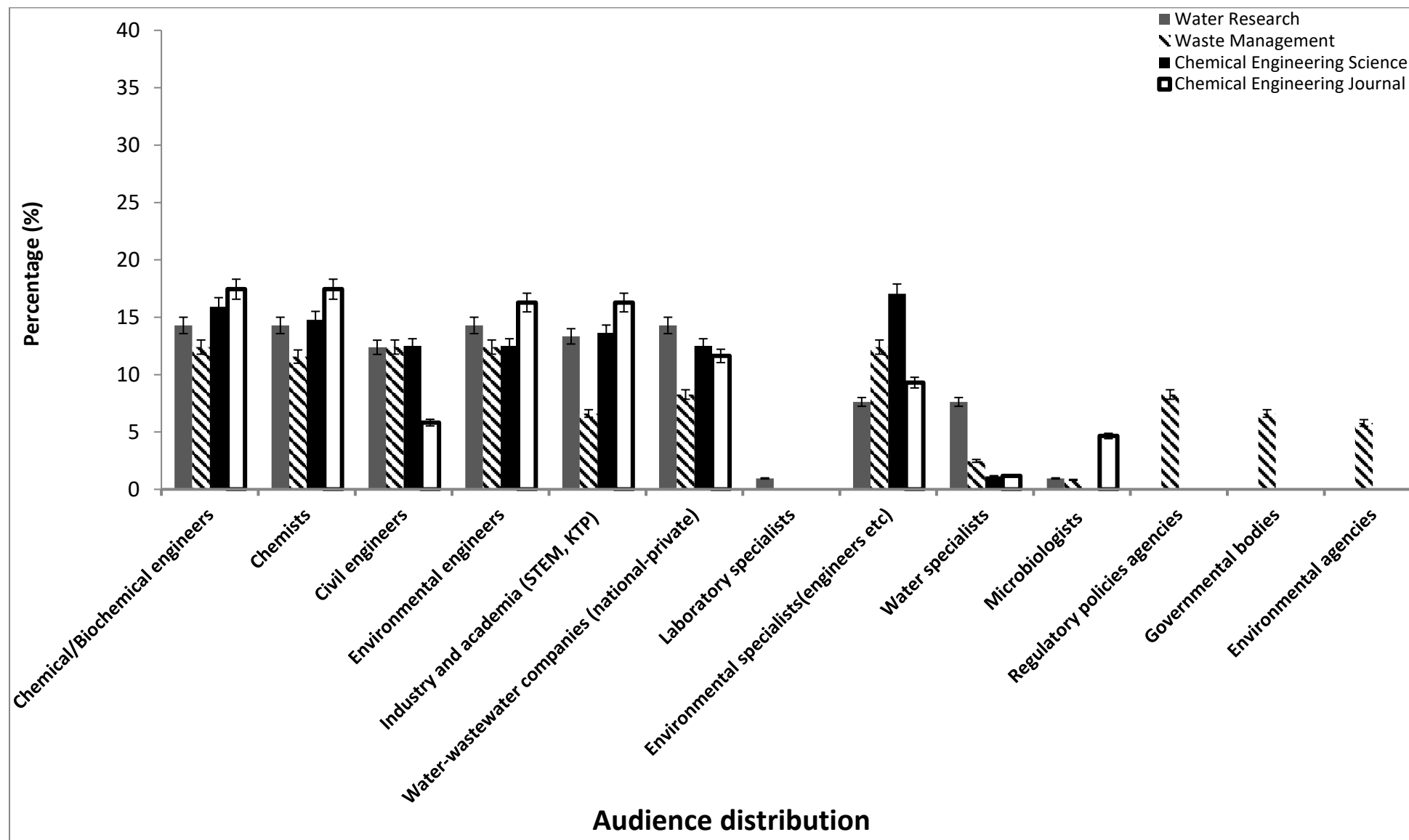


Fig. 4: Distribution of potential audience of the analysed published manuscripts

1

2 The topics covered in the papers in our sample was found to be of potential relevance to a
3 broad audience, not restricted to academia, but also to other bodies such as policy regulators,
4 small and medium size companies and enterprises, or environmental agencies. In fact, the
5 collected papers addressed an audience of 13 categories varying from water and environment
6 specialists to microbiologists and chemists, as well as governmental bodies, water and
7 wastewater companies (national, private) or regulatory policies agencies and law developing
8 and forming bodies (Fig. 4). In particular, 8 of these categories are represented in all the
9 selected journals into varying percentages (5.81% to 17.5%). This is a divergence from the
10 advertised audience in the website of each journal, where the focus is on specialist in
11 chemical engineering audience within the field.

12 The content analysis of the papers has shown that in **CEJ** and **CES** there is a stronger
13 tendency, compared to **WR** and **WM**, to appeal to the industry. That could be attributed to
14 the nature of studies, i.e. dealing with trials in pilot plant scale (large volumes of materials),
15 which are more attractive to the industry, since the authors have not only proven their
16 concept but have also implement it to a large scale. In contrast, **WR** and **WM** are primarily
17 addressing an academic audience, with **WM** publishing also on topics that are of interest to
18 the regulatory authorities of each country and globally, regarding waste; since a more holistic
19 approach is taken that accounts for financial and social parameters. Thus patterns have
20 emerged about the nuances of the addressed audience in the published manuscripts of each
21 journal, which are not clearly communicated in the journals' websites.

22 The wide range of potential audience of the published papers emphasises the need for clear,
23 concise and easily understood language, as readers coming from different academic
24 disciplines, even in close proximity, might fail to comprehend the concepts and rationale
25 expressed in the manuscripts. Figures and tables might, to a certain extent, describe the core
26 essence of the paper but the text, especially in the discussion and conclusions part are vital for
27 the overall understanding of the ideas. This is found also in this case study, where the words
28 “table” and “figures” are among the top ten words mostly used among all the journals (Table
29 2), implying that the text’s primary function, especially in the results sections, is to comment
30 upon the visual parts of the papers.

31 **3.3. Analysis of lexical choices**

32 As suggested in the introduction, linguistic and in particular lexical choices, are intimately
33 linked to the text type and intended audience. In order to investigate lexical variation among
34 different types of journal articles on wastewater treatment and management, the most
35 frequently used words in the corpus were analysed, as well as collocations of certain key
36 words, and correlations were explored between the results and the type of journal (wider or
37 specialized scope) and the journal’s IF. Six lemmas were chosen, to explore collocations and
38 consequently the context in which certain key terms are employed and variation in the
39 specific meaning that is ascribed to them (Fig. 1). These terms comprise *water* and *waste*,
40 which are generally used when referring to the environment and would be expectably
41 mentioned mostly in the introduction and discussion or conclusions parts of the papers, two
42 lemmas specifically related and descriptive of waste, *sludge* and *effluent*, that could be found

43 throughout the manuscript and especially in the results section and, finally, two lemmas
44 related to the experimental methodology used and the achieved results and relevant
45 conclusions, *treatment* and *removal* (see Fig. 5, for the frequency of occurrence of these six
46 lemmas in each journal). The collocations of the lemmas and consequently the specific
47 meaning they accrue because of their context of use (context is taken here as immediately
48 prior and upcoming text, see [41]) were analysed based on ‘word trees’ and ‘phrase net’
49 visualizations (see supplementary material). Below the key findings of the analysis of the
50 ManyEyes visualisations of the six lemmas are outlined.

51 In **CEJ** the lemma “water” was found 546 times in a total of 15 papers, and, as the analysis of
52 the visualization showed, was mainly conceptualised as a resource (ground water, surface
53 water, wastewater) either potable or as liquid waste. Focus was placed on reuse (removal of
54 harmful elements and use as washing water), recycling (water reclamation in the scope of
55 cost reduction, environmental load) and treatment (removal of toxic metals such as lead,
56 copper, harmful substances i.e. pesticides, hormones, pharmaceuticals) of water focusing on
57 wastewater treatment.

58 In the 15 **CES** collected papers, “water” occurred 176 times and was mainly understood as a
59 tool within the context of a chemical reaction, water as an aid in a chemical process for
60 example in the form of steam during sterilization, as solvent, as treatment method for other
61 elements or as a component to other substances.

62 As regards to the 15 **WR** and 15 **WM** papers, “water”, was found 792 and 244 times
63 respectively, and, as its collocations suggest, it was conceptualized as a matter worthy of
64 research, a resource, an object of analysis regarding quality, safety, treatment (potable water

65 treatment i.e. softening, salts and metals removal) wastewater (liquid waste of industrial,
66 municipal, domestic, agricultural, slaughterhouse, food, tanning industry origin), a resource
67 and water cycle (water as an environmental resource, ponds, rivers, lakes).

68 The word “waste”, found 413 times, in **CEJ** was used to indicate a problematic material that
69 has to be treated, managed and disposed. It was commonly found immediately preceding the
70 term ‘water’, forming the compound “wastewater” referring to the liquid or semi-liquid,
71 semi-solid nature of waste.

72 Similarly to the use of the lemma “water”, “waste”, occurring 131 times, in **CES**, was mainly
73 conceptualized as part or a tool of a chemical reaction, a part a chemical process, the
74 substrate or sample where the chemical process is applied on, as a component to other
75 substances.

76 In **WR** and **WM** “waste” was found 462 and 1150 times, respectively, and, similarly to the
77 word “water”, it was used in the context of a research subject deriving of numerous sources, a
78 subject of analysis regarding quality and treatment, but as well as a component or a resource
79 for the production of other materials.

80 Both lemmas “water” and “waste” were routinely found in the manuscripts of each journal,
81 and they were among the top 10 words most often-encountered words in the manuscripts, and
82 used in high frequency either combined, i.e. wastewater, or separately (Table 2). However, as
83 the analysis above indicates, in **CEJ** and **CES** the terms were recurrently employed in
84 different contexts than **WR** and **WM**. In **CEJ** and **CES** the words are used in a rather
85 specialized context compared to **WR** and **WM**, an interesting observation that did not
86 confirm the authors’ expectations, since both journals are of wider scope (Table 1), and it was

87 anticipated that a less restrictive use of the term would have been encountered. In **CEJ** and
88 **WR**, the words are found in analogous amounts; while in **CES** the amount of use is very
89 limited, suggesting the use of a scientific specialized vocabulary (e.g. the terms “liquid” or
90 “fluid” or “solvent”, were preferred over “water”). On the other hand, **WM** is standing out
91 since the lemma “waste” is used very frequently, suggesting a broader approach to the
92 subject (i.e., industrial, agricultural, slaughterhouse, domestic, municipal waste).

93 Further investigation of the observed trends, was achieved by examining the use of the words
94 “sludge”, “effluent”, “treatment” and “removal” (Fig. 1), as can be deduced from the
95 visualisations.

96 In **CEJ** the word “sludge” was found 165 times and was referred to as a problematic,
97 potentially harmful and hazardous material coming of waste. On the other hand, in **CES**
98 “sludge”, found 205 times, was used to describe a muddy, murky, highly viscous thick
99 material in the need of processing or treatment not necessarily harmful or indicative of a
100 problem. In **WR** and **WM** “sludge” occurred 129 and 85 times respectively and had a far
101 more complex meaning, as it was used in the context of harmful material coming out of
102 waste, physically looking as murky, muddy, soil based material, liquor or concentrated liquid
103 of a semisolid nature coming out of process treating sludge.

104 In **CEJ** and **CES** the word “effluent” is not found, implying the absence of mention of any
105 mechanical treatment process that would separate the solid from the liquid phase of sludge,
106 such as filtration, and the absence of any treatment involving large scale processes, a finding
107 that relates with the subject and audience distribution of the journal as defined by the
108 journal's author guidelines. In **WR** the word “effluent” was found 337 times, and was used to

109 explain any liquid coming out, discharged of a waste treatment or of waste producing
110 process, while in **WM** it occurred only 11 times, and was used when referring to any liquid
111 discharged of a leaching related process.

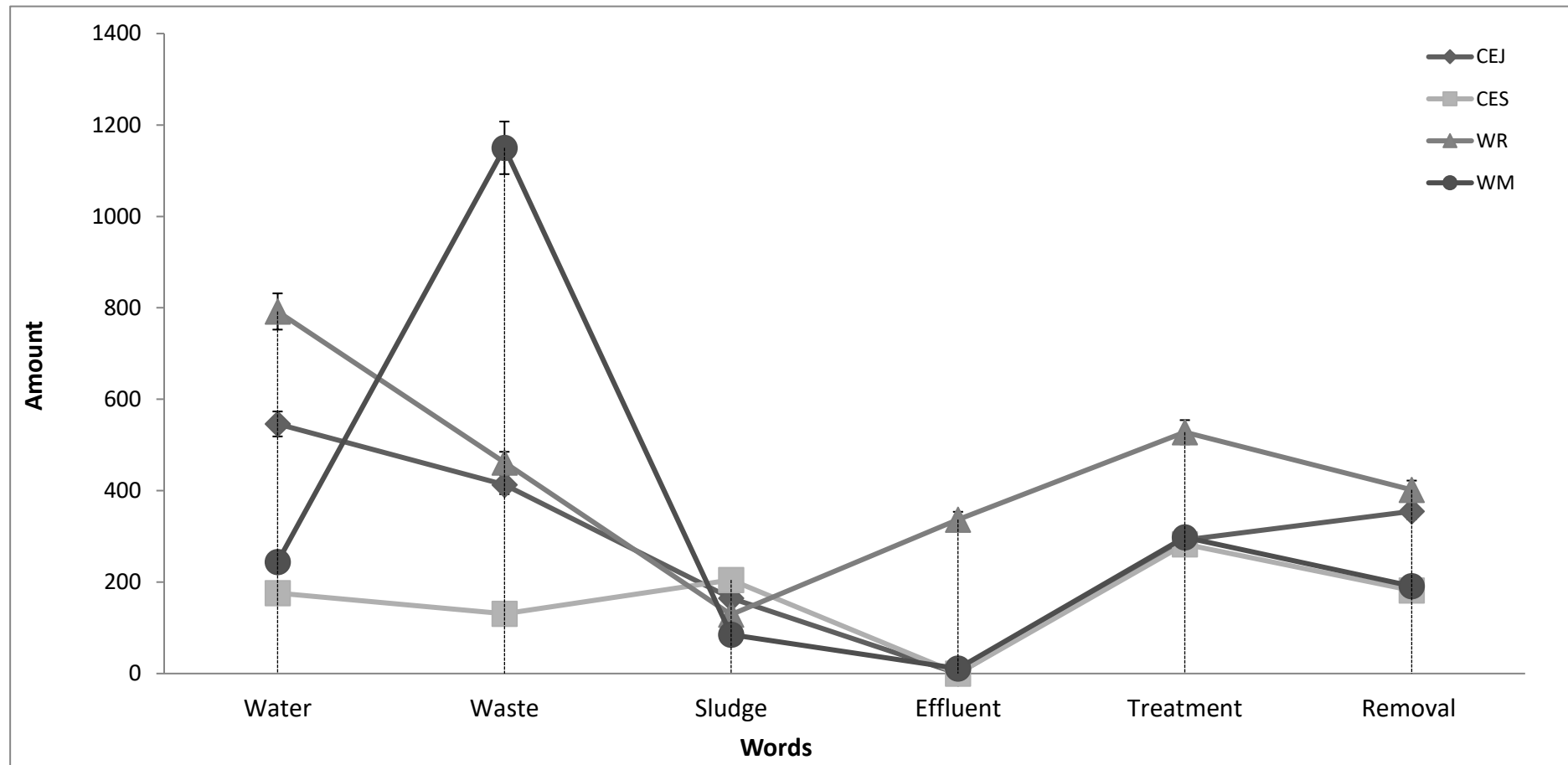
112 In **CES** the word “treatment”, occurring 283 times, referred to any method and/or process
113 used to uncouple sludge or wastewater of its harmful, dangerous, hazardous, toxic elements.
114 In **CEJ** and **WM** “treatment”, found 292 and 298 times correspondingly, was used to
115 describe any process used, developed or applied to water, wastewater and sludge, without
116 specifically explaining whether it is done to remove hazardous substances or simply for
117 treatment. In **WR** “treatment” occurred 528 times and had a more generic meaning, referring
118 to any process in which waste is involved, for example anaerobic digestion for combined heat
119 and electricity production, to technologies or systems used to remove the harmful
120 components.

121 Finally, the word “removal”, found 325 and 182 times in **CEJ** and **CES** respectively, was
122 employed to refer to any method and or process used to recover nutrients from the waste or
123 remove all the components that are harmful and /or toxic, and its effectiveness and efficiency.
124 In **CES** “removal” also represented the main scope of the project developed in the
125 manuscript. In **WR** and **WM**, “removal”, occurring 402 and 191 times, was used in the
126 context of referring to any process or method applied to the removal of harmful elements
127 from the discharged effluents, wastes, sludge or wastewater.

128

129

130



131

132

Fig. 5: Distribution of the selected keywords among the analysed manuscripts

133 4. Discussion

134 The analysis of visualisations and word frequencies (see supplementary material) has shown
135 emerging trends in lexical choices that also have implications about the specific subject-area
136 and approach preferred in each journal and which -interestingly- do not necessarily
137 correspond with the advertised scope of each journal or with the authors' expectations.
138 Among the four journals selected, **WR** and **WM** were considered of specialized scope (based
139 on their advertised scope), thus expected to accommodate a highly specialized and technical
140 lexis, whereas **CEJ** and **CES** were expected to use less discipline-specific lexis, due to their
141 wider range of scientific areas and potential audiences, as described on the journals' websites
142 (Table 1).

143 However, these expectations were not completely supported by the findings. **WM** and **CES**
144 were found to be the journals where a more specialized vocabulary is used, especially in **WM**.
145 The high occurrence of discipline-specific vocabulary is not only associated with the scale of
146 the experiments, but also with the methodology and experimental phase meaning the size, the
147 accommodating volume of the equipment and the size of volume eligible to be processed by
148 the proposed methodology, rather than the results and their impact and applicability. The
149 technical vocabulary was mainly associated with quantifiable data, experimental trial
150 chemical reaction and processing, for example "model", "fig.", "lysimeter", "system",
151 "reaction", as the lists of the most frequently used words in these journals indicates (Table 2).
152 This finding is also supported by the close reading of the published manuscripts, that has
153 revealed that the manuscripts in **WM** and **CES** refer to specialized and complex methods of

154 chemical engineering (Fig. 3 and Fig. 4). For example in **CES** instead of plainly using “water”
155 other related terms are used such as “concentrations”, “phase” which point towards to
156 chemical processing, whereas in **WM** terms related to water such as “leachate” are used to
157 point residuals of solid wastes.

158

159

Journal	Water Research	Waste Management	Chemical Engineering Science	Chemical Engineering Journal
Affiliations	International Water Association (IWA)	-	-	-
Website	www.journals.elsevier.com/water-research	http://www.journals.elsevier.com/waste-management/	http://www.journals.elsevier.com/chemical-engineering-science/	http://www.journals.elsevier.com/chemical-engineering-journal/
Publisher	<i>Elsevier B.V.</i>			
Audience	Chemists, biologists, microbiologists, immunologists, limnologists, civil engineers, sanitary engineers and chemical engineers.	Scientists, engineers and technical managers concerned with waste treatment and the engineering problems related to environmental protection laws. scientists, engineers, and managers, regardless of their discipline, who are involved in scientific, technical and other issues related to solid waste management.	Industrial and academic researchers in chemical and process engineering.	Chemical and process engineers, applied chemists and product engineers, biochemical engineers and biotechnologists
Impact Factor (IF) ¹	6.942	4.030	2.895	6.216
Publication Rate	20 issues per year (1 volume per year)	10 issues per year (1 volume per year)	12 volumes per year (1 issue per volume)	No issues, 39 volumes per year
Mean Number of publications per issue/volume	36	25	20	56
Types of papers published	Full papers, review papers, comments	Full papers, review papers, letters to the editor, columns	Original papers, review articles, short communications, letters to editors	Original papers, review articles, short communications, letters to editors
Scientific subjects published	No specific scientific sections, the journal interested in water quality and its management. It publishes original research on treatment processes for municipal, agricultural and industrial water and wastewaters, water quality standards and	Emphasis is placed on integrated approaches, major areas in which papers are solicited: generation and characterization, minimization, recycling and reuse, storage, collection, transport, and transfer, treatment (mechanical, biological, chemical, thermal, other), landfill disposal (including design, monitoring, remediation of old sites), environmental	Publication of papers on the fundamentals of chemical engineering, including. Industrial areas covered by the journal include biotechnology, chemicals, energy, food, materials, microelectronics, nanotechnology, specialty chemicals and pharmaceuticals. biomolecular and biological engineering, biochemical and bioprocess engineering, energy, water, environment, and sustainability materials engineering, particle technology; process	Three aspects of chemical engineering: chemical reaction engineering, environmental chemical engineering, and materials synthesis and processing.

¹ 2012 Journal citations report by Thomson Reuters <http://thomsonreuters.com/journal-citation-reports/> (last accessed 29 Mar. 18).

	analysis by chemical, physical and biological methods	considerations, financial and marketing aspects, policy and regulations, education and training, planning and implementation.	systems engineering reactions, separations science and technology	
Abstract	Concise and factual, descriptive (up to 250 words)			
Graphical abstract	Optional			
Highlights²	Mandatory			

160 **Table 1:** Summary description the prerequisites request by each journal for the submission of manuscripts based on of the full aims and scope and guide for
161 authors, publically available on the journals' websites.

² Highlights are a short collection of bullet points that convey the core findings and provide readers with a quick textual overview of the article. These three to five bullet points describe the essence of the research (e.g. results or conclusions) and highlight what is distinctive about it. There should be a maximum of 85 characters, including spaces, per highlight.

162

163 On the other hand, CEJ and WR use a less discipline-specific vocabulary, with salience of
164 terms “wastewater”, “effluents”, “samples” which are far less frequent in **CES** and **WM**
165 (Table 2). **WR** published papers are indicating a holistic approach to water-related research
166 focusing on the findings of the experimental trials and their applicability in the society,
167 addressing social, financial and legal aspects. This also corresponds with the frequent use of
168 the lemmas “removal” and “environmental”.

169

170

171

172

Journals								
Water Research		Waste Management		Chemical Engineering Journal		Chemical Engineering Science		
<u>Words</u>	<u>Amount</u>	<u>Words</u>	<u>Amount</u>	<u>Words</u>	<u>Amount</u>	<u>Words</u>	<u>Amount</u>	
water	700	waste(s)	1151	concentration(s)	602	water	478	
treatment	540	leachate	748	model	470	pH	405	
concentration(s)	520	landfill	671	fig	468	concentration(s)	351	
effluent(s)	520	lysimeter	503	gas	382	removal	326	
removal	410	fig	396	mm	327	fig	325	
samples	408	emissions	376	CO ₂	323	wastewater	296	
wastewater	350	system	332	rate	322	mg	283	
environmental	276	collection	298	absorption	283	treatment	277	
mg	307	treatment	297	reaction	267	min	244	
table	264	cod (chemical oxygen demand)	289	pH	phase	262	phosphate	239
Total number (15 papers/journal)	144798	115491		101519		125910		

173

174 **Table 2:** The top ten words occurring in each journal and their total number of occurrence.

175

176 In the case of **CEJ** the findings, from the qualitative analysis of the papers, regarding the
177 multidisciplinary nature of the papers are mirrored in the results of lexical analysis. It further
178 confirms that the use of a less discipline-specific vocabulary enhances the readability of the
179 journal, which can reach a wider audience, including industrials and policy regulators. On the
180 other hand, while **WR** has been found in the qualitative analysis to target in the main an
181 academic audience, the use of simplified vocabulary boosts its readability among scientists
182 from a wide range of varying disciplines.

183 Such findings indicate a correlation between increased intelligibility (beyond the narrowly
184 conceived discipline of environmental chemical engineering) and citability of the journals,
185 since **WR** and **CEJ** have the highest impact factors of 4.655 and 3.473 respectively.
186 Technology has facilitated tremendously knowledge exchange shifting from only printed
187 media to a combination of available online, easily downloadable articles and printed media,
188 expanding significantly the availability of a paper, as the readers are not depended only on
189 the printed resources that exist in libraries and repositories across the world [17, 37].

190 Literature searches are not necessarily guided by advisors, supervisors or assisting librarians,
191 and are being partially replaced by specialized research engines such as Google Scholar or
192 Scopus and the relevant webpages of the main academic publishers such as Springer,
193 Elsevier, Sage or Wiley. This leads to reading of the majority of published papers, on an
194 individual unsupervised basis, from an audience that may not have an extensive knowledge
195 on the subject (postgraduate students, early career, professionals, researchers, academics and
196 fellows), and may be novices on the specific subject area of the article. Employing highly

197 complex, scientific lexis might not facilitate the understanding of the manuscripts by readers
198 and will possibly result in lower citability. This can explain the association that was found in
199 this study between more accessible, less specialised vocabulary and higher IF.

200 When comparing these findings to the advertised scope of each journal, certain differences
201 are found. Among all four journals, only **WR** published papers reflect the journal's very
202 broad approach, focusing on innovation without disregarding new approaches to current
203 techniques. **CEJ** and **CES** have a narrowed thematology, addressing highly specific subjects
204 contrary to the journals advertised spectrum. In the published manuscripts, emphasis is
205 placed on optimization of existing methods, mainly chemical treatments rather than
206 innovation, which cannot be as easily and quickly applicable. A similar tendency is found in
207 **WM**, where, in spite of the advertised wide array of publishing subjects, the published
208 manuscripts do not cover such a wide spectrum, and focus primarily on waste management
209 and relevant regulations, reflecting the anisomorphy between the advertised and the actual
210 scope of the journals.

211 **5. Conclusions**

212 This is a case study and results are not unproblematically generalizable across journals of
213 practical sciences, let alone all disciplines. However, due to the depth of the investigation this
214 snapshot of trends in published chemical engineering research has offered an insight on the
215 implications of publishing research findings that can be extended beyond the four journals.
216 Some tentative conclusions that could be deduced regarding the lexical and thematic choices
217 in original chemical engineering research articles and which could be incorporated in

218 learning and teaching material for chemical engineers, but also researchers from other
219 disciplines that seek to publish their research include the following.

220 • Highly discipline-specific vocabulary use, including extensive use of acronyms, should be
221 avoided where possible, to aid favorable consideration of manuscripts at higher IF journals
222 and to increase the citability potential of the article.

223 • There is a complex relationship between the thematology, the audience and the scope, as
224 they are advertised in the journal's website, and the actual published manuscripts.

225 • Guide for authors and journal aims and objectives, published by the journal's editorial office,
226 should be taken into account, to help authors make an initial decision regarding the journal
227 that is most suitable for the submission of their research, but should be critically viewed.

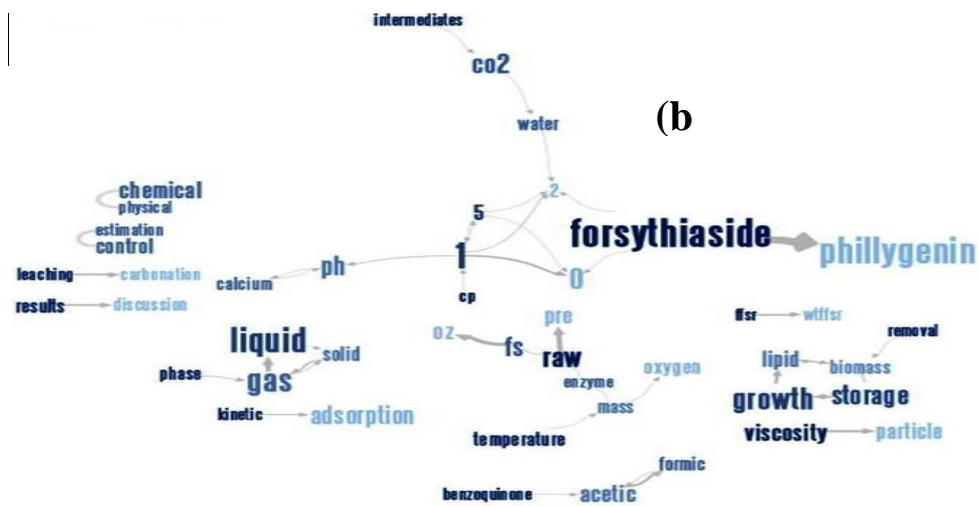
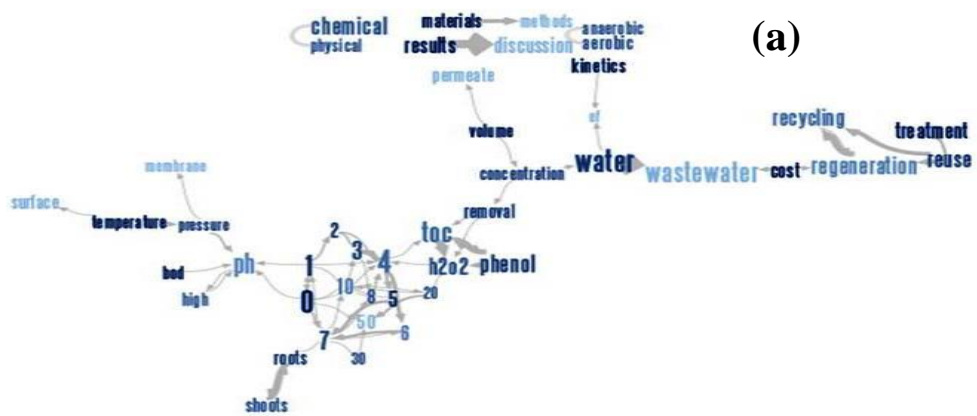
228 • It is recommended for prospective authors to collect a number of publications, of their
229 journal of choice, published within close proximity, to the potential submission date, in
230 order to get a better understanding of the journal's thematology, the approaches favored and
231 preferred discourse style.

232 • Visualisations of word choices and associations, which can be fairly easily and quickly
233 done with the aid of freely available software, is a very powerful tool in providing an
234 accurate overview of both the preferred content and approach of each journal, as well as its
235 preference as regards to lexical choices. They can be an indispensable tool for chemical
236 engineering students and novice researchers that wish to gain an emit understanding of the
237 actual scope of the plethora of journals within each discipline, without having to engage in
238 the labor-intensive close reading of a large corpus of published papers.

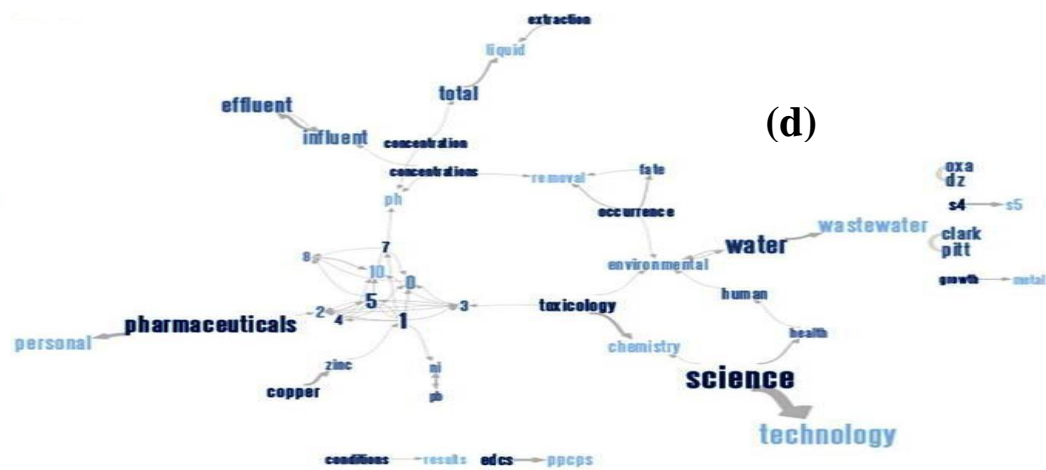
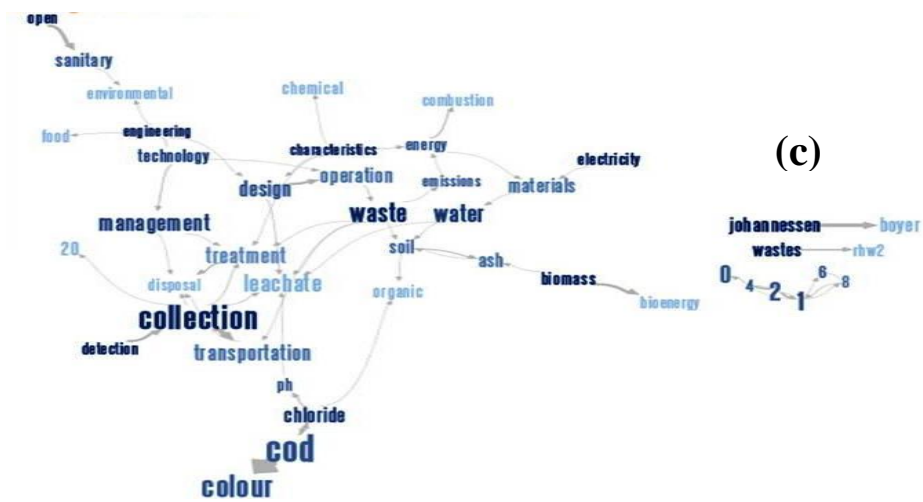
239 Extending this research to similar investigations of a larger size of text samples, representing
240 more fields of science would be desirable, so that the findings will then be more
241 representative of scientific writing in English. Further exploration of links between linguistic
242 choices and citability, impact factor, new media use and altmetrics (online traffic of journal's
243 published manuscripts) could lead to the development of a methodology that would help the
244 researchers to write in a style that best suits their target journal.

245

246 Supplementary Materials:



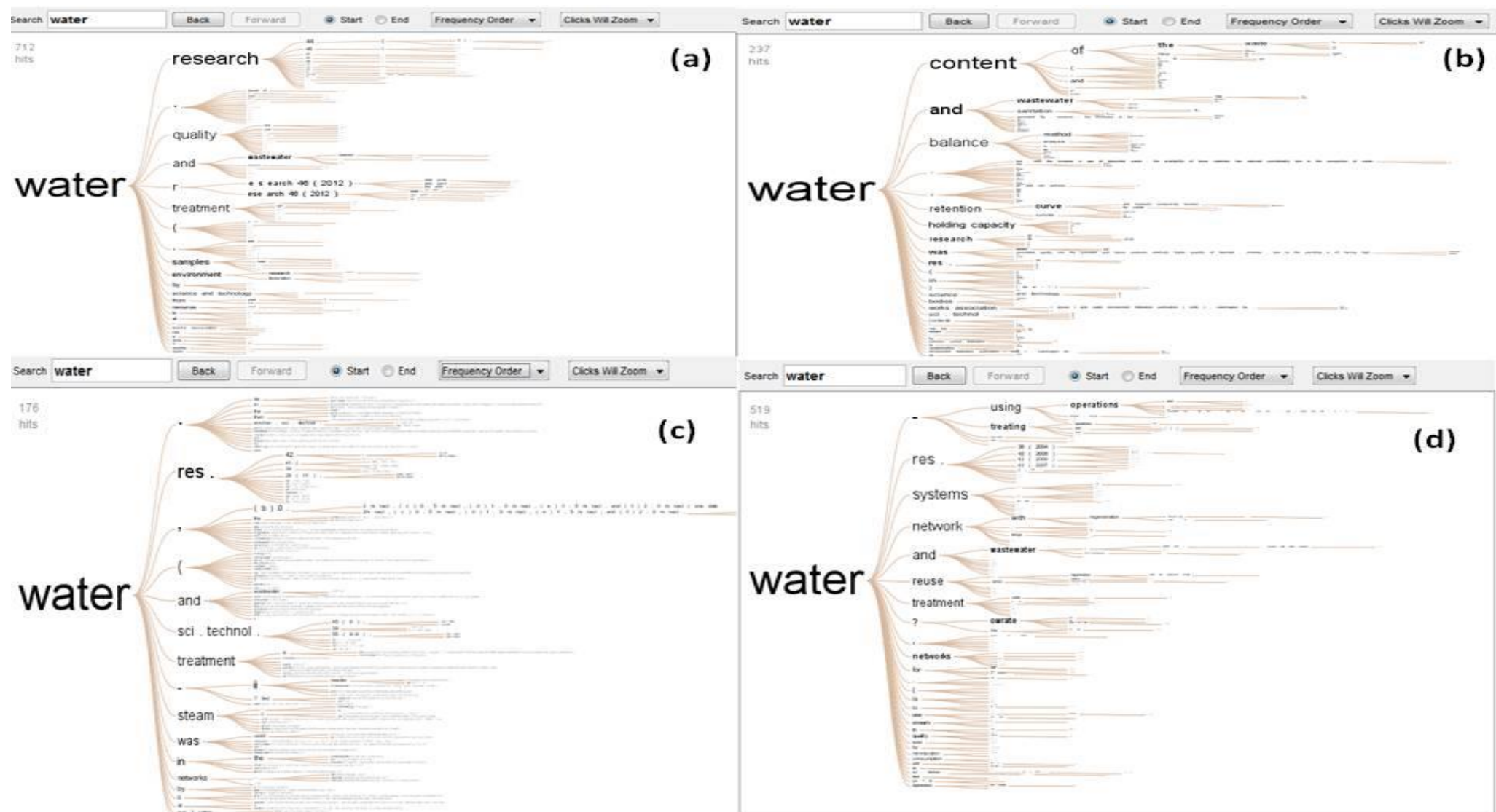
247



248

249

250 Fig.1: Phrase nets graphical images depicting the collocations between the selected words for analysis and the remaining words in the selected published manuscripts in (a)CEJ, (b)CES, (c)WR,
251 and (d)WM provided by the lexical visualisation software Many Eyes and used for the qualitative analysis of the published manuscripts in this case study.
252



253

254

255

Fig.2: Word trees depicting the word “water” in (a)CEJ, (b)CES, (c)WR, and (d)WM and its collocations (word associations) provided by the lexical visualisation software Many Eyes and used for the qualitative analysis of the contexts of use of selected lemmas in the published manuscripts in this case study.

261 **Acknowledgments:** The authors would like to thank fellow researchers in the College of
262 Engineering, Swansea University for their valuable advice in the research discussed here.
263

264 **References**

- 265 [1] Abramo, G., D' Angelo C.A., Rosati F., 2014. Career advancement and scientific performance in
266 universities. *Scientometrics*, 98, 891-907.
- 267 [2] Massoudi, M., 2003. Can scientific writing be creative? *J. of Sci. Edu. Tech.*, 12, 115-128.
- 268 [3] Vanclay, J., 2012. Impact factor: outdated artefact or stepping-stone to journal certification?
269 *Scientometrics*. 92, 211-238
- 270 [4] Beall, H., 1998. Expanding the scope of writing in chemical education. *J. Sci. Edu. Tech.*, 7,
271 259-270.
- 272 [5] Uskokovic, V. 2012. Ten Commandments for writing a meritable scientific paper. *J. Postdoc.*
273 *Affairs*, 2, 2-7.
- 274 [6] Bosquet, C., Combes, P.-P., 2013. Are academics who publish more also more cited? Individual
275 determinants of publication and citation records. *Scientometrics*. 97, 831-857
- 276 [7] Porter, A., Rafols, I., 2009. Is science becoming more interdisciplinary? Measuring and mapping
277 six research fields over time. *Scientometrics*. 81, 719-745.
- 278 [8] Woolley, K., Barron, P., 2009. Handling manuscript rejection insights from evidence and
279 experience. *Chest*. 135, 573-5
- 280 [9] Drummond, C. W. E., Reeves, D., 2005. Reduced time to publication and increased rejection rate.
281 *J. Antimicrob. Chemo.* 55, 815-816.
- 282 [10] Beall, J., 2012. Beall's list of predatory publishers 2013, <http://scholarlyoa.com/publishers> (last
283 accessed 29 Mar 2018).
- 284 [11] Moore, A. (2009). The garbage collectors - Could a particular sector of author-pays journals
285 become silently acknowledged collectors of scientific waste? *BioEssays*, 31, 821.
- 286 [12] Amin, M., Mabe, M., 2000. Impact factor: use and abuse. *Persp.s in Publ.*, 1, 1-6.

- 287 [13] Fassoulaki, A., Sarantopoulos, C., Papilas, K., Patris, K., Melemeni, A., 2001. Academic
288 anaesthesiologists' views on the importance of the impact factor of scientific journals: a North
289 American and European survey. *Gen. Anaesthesia*. 48, 953-957.
- 290 [14] Ovalle-Peradones, M.-A., Gorraiz, J., Wieland, M., Gumpenberger, C., Olmeda-Gomez, C.,
291 2013. The influence of European Framework Programmes on scientific collaboration in
292 nanotechnology. *Scientometrics* 97, 59-74.
- 293 [15] Garfield, E., 1999. Journal impact factor: a brief review. *Can. Med. Assoc. J.* 16, 1979-980.
- 294 [16] Misteli, T., 2013. Eliminating the impact of the impact factor. *J. Cell Biol.* 201, 651-652.
- 295 [17] Oliveira, E. A., Peicots-Filho, R., Martelli, D. R., Quirino, I. G., Oliveira, M. C. L., Duarte, M.
296 G., Martelli-Junior, H., 2013. Is there a correlation between journal impact factor and researchers'
297 performance? A study comprising the fields of clinical nephrology and neurosciences. *Scientometrics*.
298 97, 149-160.
- 299 [18] Finegold, L., 2002. Writing for science as scholarly communication. *J. Sci. Edu. Tech.* 11,
300 255-260.
- 301 [19] Alaimo, P.J., Bean, J.C., Nichols, L., 2009. Eliminating lab reports: a rhetorical approach for
302 teaching the scientific paper in sophomore organic chemistry. *WAC J.*, 20, 17-32.
- 303 [20] Robinson, M.S., Stoller, F. L., Costanza-Robinson, M.S., Jones, J.K., 2008. *Write like a chemist*.
304 New York: Oxford University Press.
- 305 [21] Rodriguez, A.C., 2012. Teaching peers to talk to peers. *BioEssays*. 34, 918-920.
- 306 [22] Lebrun, J.-L., 2013. *Scientific writing: a reader and writer's guide*. Singapore: World Scientific
307 Publishing Co Pte. Ltd.
- 308 [23] Derntl, M., 2014. Basics of research paper writing and publishing. *Int. J. of Tech. Enh. Learn* 6,
309 105-123
- 310 [24] Conrad, S., 1996. Investigating academic texts with corpus-based techniques: an example from
311 biology. *Ling Edu.* 8, 299-326.

- 312 [25] Gunawardena, C., 1989. The present perfect in the rhetorical divisions of biology and
313 biochemistry journal articles. *Engl. Spec. Purp.*, 8, 265-273.
- 314 [26] Stoller, F., Jones, J., Costanza-Robinson, M. S., Robinson, M. S., (2005). Demystifying
315 disciplinary writing: A case study in the writing of chemistry. *Across the Disciplines*, 2.
316 <http://wac.colostate.edu/atd/lds/stoller.cfm> (last accessed 15 April 2015).
- 317 [27] Burnard, P., Gill, P., Stewart, K., Treasure, E., Chadwick, B., 2008. Analysing and presenting
318 qualitative data. *Brit. Dent. J.* 204, 429-432.
- 319 [28] Gill, P., Stewart, K., Treasure, E., Chadwick, B., 2008. Methods of data collection in qualitative
320 research: interviews and focus groups. *Brit. Dent. J.* 204, 291-295.
- 321 [29] Biber, D., Conrad, S., Reppen, R., 1994. Corpus-based approaches to issues in applied
322 linguistics. *App. Ling.*, 15, 69-189.
- 323 [30] Biber, D., Conrad, S., 2009. Register, genre and style. Cambridge, UK; New York: Cambridge
324 University Press.
- 325 [31] Gustavii, B., 2008. How to write and illustrate a scientific paper. Cambridge, UK; New York,
326 Cambridge: Cambridge University Press.
- 327 [32] Blackwell, J., Martin, J., 2011. A scientific approach to scientific writing. New York: Springer.
- 328 [33] Okulicz-Kozaryn, A., 2013. Cluttered writing: adjectives and adverbs in academia.
329 *Scientometrics.* 96, 679-681.
- 330 [34] Hosgood, G., 2011. How to write and publish a scientific paper. *Austral. Vet. Pract.*, 41, 137-144.
- 331 [35] Ware, M., Mabe, M., 2009. *STM: An overview of scientific and scholarly journal publishing*.
332 Oxford: International Association of Scientific, Technical and Medical Publishers.
- 333 [36] White, T., 2006. Principles of good research and research proposal guide. In: Policy, performance
334 and quality assurance unit. (pp. 1-5). London: London Council, Borough of Richmond upon Thames.
- 335 [37] Maffioli, F., Augusti, G., 2003. Tuning engineering education into the European higher education
336 orchestra. *Eur. J. Eng. Edu.* 28, 251-273.

- 337 [38] Saha, S., Saint, S., Christakis, D., 2003. Impact factor: a valid measure of journal quality? *J. Med.*
338 *Librar. Assoc.* 91, 42-46
- 339 [39] Barrow, L.H., 2003. Searching for educational technology faculty. *J. Sci. Edu. Tech.*, 12, 143-147.
- 340 [40] Tschardtke, T., Hochberg, M.E., Rand, T.A., Resh, V.H., Krauss, J., (2007). Author sequence and
341 credit for contributions in multiauthored publications. *PLOS Biol.* 5, 13-14.
- 342 [41] Schegloff, E. A., 1992. In another context. In A. Duranti and C. Goodwin (Eds.), *Rethinking*
343 *context: language as an interactive phenomenon.* (pp. 193-227). Cambridge; New York: Cambridge
344 University Press.
- 345 [42] <http://www.ref.ac.uk/2014/panels/assessmentcriteriaandleveldefinitions/> (last accessed
346 29.Mar.18)