

## Examining the phenomenon of quarter-life crisis through artificial intelligence and the language of Twitter

Article (Accepted Version)

Agarwal, Shantenu, Guntuku, Sharath Chandra, Robinson, Oliver C, Dunn, Abigail and Ungar, Lyle H (2020) Examining the phenomenon of quarter-life crisis through artificial intelligence and the language of Twitter. *Frontiers in Psychology*, 11. a341 1-11. ISSN 1664-1078

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/94673/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

# Examining the Phenomenon of Quarter-life Crisis through Artificial Intelligence and the Language of Twitter

1 Shantenu Agarwal<sup>1</sup>, Sharath Chandra Guntuku\*<sup>1</sup>, Oliver Robinson<sup>2</sup>, Abigail Dunn<sup>3</sup>, and Lyle  
2 Ungar<sup>1</sup>

3 <sup>1</sup> University of Pennsylvania, USA

4 <sup>2</sup> University of Greenwich, UK

5 <sup>3</sup> University of Sussex, UK

6

7 \*Corresponding author information: Sharath Chandra Guntuku, 3400 Civic Center Blvd,  
8 Philadelphia, PA, 19104, (e-mail: sharathg@sas.upenn.edu).

9

## 10 **Abstract**

11 Quarter-life crisis (QLC) is a popular term for developmental crisis episodes that occur during  
12 early adulthood (18-30). Our aim was to explore what linguistic themes are associated with this  
13 phenomenon as discussed on social media. We analyzed 1.5 million tweets written by over 1,400  
14 users from the UK and USA that referred to QLC, comparing their posts to those used by a control  
15 set of users who were matched by age, gender and period of activity. Logistic regression was used  
16 to uncover significant associations between words, topics, and sentiments of users and QLC,  
17 controlling for demographics. Users who refer to a QLC were found to post more about feeling  
18 mixed emotions, feeling stuck, wanting change, career, illness, school and family. Their language  
19 tended to be focused on the future. Of 20 terms selected according to early adult crisis theory, 16  
20 were mentioned by the QLC group more than the control group. The insights from this study could  
21 be used by clinicians and coaches to better understand the developmental challenges faced by  
22 young adults and how these are portrayed naturalistically in the language of social media.

23

## 24 **Keywords**

25 Quarter-life Crisis; Machine Learning; Natural Language Processing; Social Media; Emerging  
26 Adulthood

27

## 28 **Data availability statement:**

29 The datasets for this manuscript are not publicly available to protect the privacy of the users.  
30 Requests to access the datasets should be directed to Sharath Guntuku  
31 [sharathg@sas.upenn.edu].

32

## 33 **Conflict of Interest statement**

34 The authors declare that the research was conducted in the absence of any commercial or  
35 financial relationships that could be construed as a potential conflict of interest.

36

37

38

39

40 **Author Contributions Statement**

41 SA and SG originated the study. SA, SG, OR, and LU developed methods, interpreted analysis,  
42 and contributed to the writing of the manuscript. AD assisted with critical developments to the  
43 manuscript, including theoretical content.

44

45 **Contribution to the Field statement**

46 This study is the first to study via a developmental framework how the term quarter-life crisis is  
47 deployed on social media. We explored how the language associated with quarter-life crisis relates to  
48 theoretical frameworks pertaining to the developmental challenges of being a young adult. Over the  
49 past ten years, the term quarter-life crisis has become common parlance amongst young adults and is  
50 present in the self-help literature and in the media. We analyzed tweets written by over 1,400 Twitter  
51 users that referred to QLC, then compared these posts to those used by a control set of users who  
52 were matched by age, gender and period of activity. Using *Differential Language Analysis* and  
53 *Linguistic Inquiry and Word Count*, we found a range of associations with quarter-life crisis on social  
54 media that indicate it to be closely linked to psychological models of early adult crisis and the theory  
55 of emerging adulthood. The study provides convincing evidence that the textual analysis of social  
56 media posts can productively increase our understanding of specific periods of adult development,  
57 via constructs that are popularly used to convey challenging experiences that are life-stage-specific.

## 58 Introduction

59 Quarter-life crisis (QLC) is a phenomenon that has become widely discussed in the media  
60 and in popular writings about the challenges of early adulthood (Robbins & Wilner, 2001;  
61 Rosen, 2019; Wiley, 2015). Academic discourse on the phenomenon has also developed in the  
62 form of empirical and theoretical work on developmental crisis episodes during the first decade  
63 of adult life (Robinson, 2016) and also in applied disciplines such as coaching (Stapleton, 2012).  
64 The present study explores how the phenomenon is linguistically rendered in the social media  
65 space of Twitter, with the aim of better understanding the popular conception of the phenomenon  
66 and how that may help to convey the nature of key developmental challenges pertaining to being  
67 a young adult.

68 The theoretical frameworks used to analyse, explore, and interpret QLC are the theory of  
69 emerging adulthood (Arnett, 2000; Arnett, 2007) and the model of early adult crisis by Robinson  
70 and colleagues (Robinson & Smith, 2010; Robinson, Wright & Smith, 2013; Robinson, 2019).  
71 The theory of emerging adulthood proposes five defining developmental features of the age  
72 range of 18-28. These are: (1) feeling ambiguous in terms of adult status - young people in this  
73 age range typically describe themselves as in some ways an adult, yet in some ways not, and as  
74 being caught in between; (2) a period of active exploration of self and world; (3) a time of  
75 instability in roles and relationships, stemming from a continued lack of long-term ties that  
76 permits changes in lifestyle, role and residence; (4) a time of adaptive self-focus as young people  
77 attempt to invest in their own future; and (5) a time of future-focus and optimism (Arnett, 2007;  
78 Arnett & Mitra, 2018; Munsey, 2006).

79 Emerging adulthood as a purported life stage can be defined demographically too.  
80 Epidemiological data shows that most young adults in Western countries now choose to wait for  
81 a decade or more after turning 18 before having children, or before starting a marriage or civil  
82 partnership (Perelli-Harris & Lyons-Amos, 2015). This demographic fact of postponing  
83 parenthood and marriage is integral to the theory of emerging adulthood, as it is during the years  
84 prior to settling down that young adults can continue the exploration of their identity, roles and  
85 relationships, via the accrual of new and diverse life experiences (Arnett, 2000). Questions have  
86 been raised over whether or not the theory of emerging adulthood applies to individuals or  
87 subcultures who still settled down at the age of 18 or soon after (Nelson & Barry, 2005). Yet the  
88 theory pertains to what is normative rather than universal, hence accepts that there will be  
89 minorities that do not fit the model. There is now a substantial body of research attests to its wide  
90 generalisability across socio-economic groups and cultures at a normative level (Arnett, 2016;  
91 Robinson, 2016). However, the theory is also clear that economic and cultural contexts frame the  
92 experience of emerging adulthood, and hence research on the topic should be explicit in its  
93 acknowledgement of cultural context (Konstam, 2007). The current study focuses on the UK and  
94 the USA, two countries in which rates of tertiary education are relatively high and young adults  
95 must pay for such education, typically leading to large debts. There is evidence to suggest that  
96 levels of stress are high in emerging adults in the UK and the USA compared with other age  
97 groups within each country (Forth, 2018; Stone, Schwartz, Broderick & Deaton, 2010).

98 Early adult crisis episodes typically occur towards the latter end of the life stage of  
99 emerging adulthood, and last approximately a year (Robinson, 2016). They are periods of

100 instability, transition and heightened emotion, and are typically triggered when a person makes  
101 substantive and active efforts to settle down into a steady set of life roles but then experiences  
102 difficulties that lead to feeling overwhelmed and unable to cope (Robinson, 2016). As such, early  
103 adult crisis typically revolves around a struggle with either feeling *locked out* of adult  
104 commitments (being unable to find work or love), or the feeling of being *locked in* to life roles  
105 that are then experienced as a poor fit for one's identity, or as generally stultifying (Robinson,  
106 2016; Robinson, 2018). During an early adult crisis, there are reports of strong negative emotions  
107 but also of enhanced curiosity and opportunities for growth and change (Robinson, Demetre, &  
108 Litman, 2017). There is also an intensive focus on the question of personal and social identity,  
109 with people who are in crisis during this age group engaging in a process of questioning 'who I  
110 am' in the context of the roles and relationships, and making active attempts to answer such  
111 questions through exploration and trial-and-error (Robinson, Wright & Smith, 2013). Following  
112 an early adult crisis, there may be post-crisis growth if substantive lifestyle changes are made, or  
113 feelings of depression and lowered self-esteem if attempts to cope fall short (Robinson &  
114 Wright, 2013). Early adult crises are widespread; in a UK sample, 70% of people in their 30s  
115 report that they had one in their 20s (Robinson & Wright, 2013).

116 Based on a review of popular definitions of QLC and the model of early adult crisis,  
117 Robinson (2016) concluded that they can be considered essentially synonymous. Counsellors and  
118 coaches frequently use the term QLC to frame transitional difficulties and concomitant emotional  
119 challenges that young adults experience (Hapke, 2017; Stapleton, 2012). It has also become a  
120 topic of discussion in popular self-help literature and the media (Rosen, 2019; Wiley, 2015; Jay,  
121 2012). Given this widespread general usage, it is unsurprising that the phrase has also become  
122 part of the vernacular of many young adults who attempt to make sense of their personal  
123 challenges through the lens of QLC. We assume that given the relatively loose meanings  
124 surrounding the construct in the popular domain, references to QLC in social media are likely to  
125 be shorthand for many issues that pertain to being a young adult more generally, and so will not  
126 only link to early adult crisis, but also to many of the issues that pertain to emerging adulthood,  
127 such as uncertainty, stress, self-focus, and feeling caught in-between (Black, 2010).

128 The portrayal of autobiographical events and experiences on social media is an important  
129 frontier for psychological research, showing promise as a tool for studying lifespan development  
130 and mental health (Toseeb & Inkster, 2015). Facebook postings, tweets, emails and text  
131 messages that contain information about ongoing personal experiences and life events can  
132 collectively be referred to as micro-narratives (Giles, 2017), or small stories (Georgakopoulou,  
133 2017). Each of these brief texts tells a story to an intended audience, while drawing on generic  
134 constructs that link each posting to broader cultural concerns or popular academic theories (van  
135 Dijck, 2007). Social media postings that relate to actual life events and experiences can be  
136 argued to serve a developmental function, which is to represent and reify the passing of time into  
137 a simplified and publicly documented life story that can help the individual create a meaningful  
138 ongoing narrative of how their life is changing (Rettberg, 2009).

139 In terms of analyzing social media data, developing vectoral representations for words  
140 (i.e. word clouds) using AI machine learning systems such as Word2Vec has recently been  
141 gaining popularity as a way of representing language usage (Mikolov, Yin, & Zweig, 2013).

142 These methods permit the capture of local context order rather than just "bag-of-words"  
143 relatedness, which in turn leads to the capture of syntactic information.

144 The aim of this study was to explore if QLC is represented in social media using  
145 linguistic features that can provide empirical illuminations about emerging adulthood and early  
146 adult crisis *within the context of the USA and UK*. This study is, to the best of our knowledge,  
147 the first attempt to study the language of QLC through the application of natural language  
148 processing on social media data. *Twitter was selected as the social media platform on which to*  
149 *investigate this phenomenon for two reasons. Firstly, tweets are public and searchable, unless the*  
150 *user opts out by making them private. In contrast, posts on other social media platforms are*  
151 *mainly restricted to a defined audience and not publicly searchable. Secondly, it has been shown*  
152 *to be a very conducive platform for self-disclosure related to a wide variety of phenomena such*  
153 *as personality, stress, and other mental health categories (Guntuku, Buffone, Jaidka, Eichstaedt*  
154 *& Ungar, 2019; Guntuku, Preotiuc-Pietro, Eichstaedt, & Ungar, 2019; Preotiuc-Pietro, Liu,*  
155 *Hopkins & Ungar, 2017).*

156 A skip-gram AI model with negative sampling (Mikolov, Chen, Corrado, & Dean, 2013)  
157 was originally used to learn word embeddings from a corpus of 400 million tweets (Lampos,  
158 Aletras, Preotius-Pietro, & Cohn, 2014). This same method has been shown to successfully  
159 predict the income (Lampos et al., 2014) and personality of Twitter users (Guntuku, Lin,  
160 Carpenter, Ng, & Ungar, 2018).

161 We hypothesized that the topics discussed by users who expressed having a QLC would  
162 significantly differ from a matched control group who have not used this term, in ways that  
163 support existing theory on emerging adulthood and early adult crisis. An open vocabulary  
164 analysis approach was conducted in which key QLC topics were ascertained by clustering co-  
165 occurring tweets. Given its open and exploratory focus, no word-specific or term-specific  
166 predictions were made for this analysis.

167 For the second analysis, run via the Linguistic Inquiry Word Count (LIWC) system, we  
168 predicted, based on emerging adulthood theory, that movement-based words (reflecting  
169 exploration), negative emotion words (reflecting instability) and personal pronouns (reflecting  
170 self focus) would be more prevalent in the QLC group compared to the control group.

171 A third analysis was conducted on twenty terms selected to represent early adult crisis  
172 theory, based on a conceptual and thematic review of qualitative studies on early adult crisis  
173 (*Stuck; Trying; Leave; Change; Unemployed; Lonely; Hopeless; Overwhelmed; Unfair; Fail;*  
174 *Coping; Failing; Debt; Meaning; Trapped; Try; New; Identity; Sacked; Money*). We predicted  
175 that these would all be more prevalent in the QLC group than the control group.

## 176 **Methods**

### 177 *Participants*

178 Data used for this study were derived from public messages posted on Twitter from 2011  
179 to 2015. Using Twitter Search API<sup>1</sup>, we obtained a set of 3,200 unique users *aged 18-30 from the*  
180 *UK and the USA* who mentioned having a QLC. Tweets were filtered to deselect any retweets,

---

<sup>1</sup> <https://twitter.com/search-home>

181 URLs, advertisements and spam. Tweets with reference to 'Happy Birthday' were also removed  
182 to help avoid ironic mentions of QLC. Users were also filtered out if they did not have over 40  
183 messages to ensure there would be enough history and activity to analyze. After further validation  
184 of the number of words to obtain reliable language-based estimates (>500 words), there were 1,390  
185 users. For each user, we obtained their entire timeline of Tweets (maximum of 3,200) from Twitter  
186 API resulting in over 1.5 million messages across all users. These users (the QLC group) were  
187 matched with a control sample of Twitter users, who never mentioned having a QLC, consisting  
188 of the same age and gender distribution and who had posts around the same time period as the  
189 QLC group. The mean age of the QLC and control groups was 23.95, and the standard deviation  
190 was 2.74. Each groups contained 1,195 females and 195 males. The high percentage of females in  
191 the groups is congruent with previous findings that (a) more females than males self-report early  
192 adult crises (Robinson & Wright, 2013), and (b) women discuss emotional matters on Twitter more  
193 than men do (Kivran-Swaine, Brody, Diakopoulos & Naaman, 2012).

194 Table 1 highlights the process of data collection (Table 1A) and the composition of the  
195 data (Table 1B).

196  
197

#### 198 *Linguistic Analysis (LA)*

199 We used three sets of language analysis: a) Open-vocabulary clustering b) Linguistic  
200 Inquiry Word Count (LIWC) analysis c) Theory-based analysis. These language features have  
201 been shown to be predictive of several health outcomes, such as depression, schizophrenia,  
202 attention deficit hyperactivity disorder (ADHD), personality, and general well-being (Guntuku,  
203 Ramsay, Merchant, & Ungar LH, 2017; Guntuku, Yaden, Kern, Ungar, & Eichstaedt, 2017;  
204 Schwartz, Eichstaedt, Kern, et al, 2013; Schwartz, Sap, Kern, et al, 2016 ).

#### 205 *LA-a) Open-vocabulary clustering approach*

206 An open-vocabulary statistical learning and modelling approach to find topics that the QLC  
207 group talk about more than the control group. This was conducted using an open source language  
208 analysis toolkit (DLATK) (Schwartz, Giorgi, Sap, Crutchley & Ungar, 2017). From each post,  
209 words were identified (using an emoticon-aware tokenizer which also looked for tokens such as  
210 ':)', ':D' etc.) and multi-word expressions were selected, keeping 2- and 3-grams (two or three  
211 consecutive words) with the highest pointwise mutual information (PMI) or association between  
212 their words. PMI is the ratio of the joint-probability to the independent probability of observing  
213 the phrase:

$$214 \quad pmi(phrase) = \log \frac{p(phrase)}{\prod_{w \in phrase} p(w)}$$

215  
216 In practice, we kept phrases with PMI values greater than  $2 * \text{length}$ , where length is the number  
217 of words contained in the phrase, to ensure retained phrases were informative parts of speech  
218 and not just accidental juxtapositions. All word and phrase counts were normalized by each  
219 subject's total word use ( $p(\text{word} | \text{subject})$ ), and we applied the Anscombe transformation, where

220 vocab(subject) returns a list of all words and phrases used by that subject. These Anscombe  
221 transformed “relative frequencies” of words or phrases ( $p_{ans}$ ) were then used as the independent  
222 variables in all our analyses.

223  
224

$$225 \quad p(\textit{phrase} \mid \textit{subject}) = \frac{\textit{freq}(\textit{phrase}, \textit{subject})}{\sum_{\textit{phrase}' \in \textit{vocab}(\textit{subject})} \textit{freq}(\textit{phrase}', \textit{subject})}$$

226

$$227 \quad p_{ans}(\textit{phrase} \mid \textit{subject}) = 2 \sqrt{p(\textit{phrase} \mid \textit{subject}) + 3/8}$$

228

229 Artificial neural networks have recently been gaining popularity because they result in low-  
230 ranking word embeddings leading to state-of-the-art results for a number of semantic tasks  
231 (Mikolov, Yih, and Zweig, 2013). This study used a hidden layer size of 50 with the Gensim  
232 implementation.<sup>2</sup> Then a spectral clustering on these embeddings was applied to obtain hard  
233 clusters of words. This resulted in 200 hard clusters, i.e. one word belongs to only one topic. The  
234 importance score associated with every word represents how central the word is in its cluster.  
235 Clusters are computed using spectral clustering over a word-word similarity matrix generated by  
236 Word2Vec. These clusters, termed as Topics in subsequent analysis, are available online.<sup>3</sup>

237

238

239 *LA-b) Linguistic Inquiry Word Count (LIWC) analysis:*

240 This analysis consists of words grouped into 73 categories (such as Functional words,  
241 Money, Family etc.) and shown to previously predict multiple user traits such as stress, health,  
242 personality, etc. (Pennebaker, Boyd, Jordan & Blackburn, 2015; Tausczik & Pennebaker, 2010).  
243 From each post on Twitter, we extracted the relative frequency of single words and phrases  
244 (consisting of two or three consecutive words). Then, all words used by less than 1% of users were  
245 removed from analysis so as to remove uncommonly used words (outliers). All messages used to  
246 identify the study group (i.e. tweets containing #quarterlifecrisis) were removed so that the logistic  
247 regression model captures other linguistic attributes associated with the study group above and  
248 beyond this selection criteria. The distribution of LIWC dictionary features were also extracted for  
249 each post. For each user, we measured the proportion of word tokens that fall into a given LIWC  
250 category. Then, we compared it against the word tokens from the control data using an empirical  
251 distribution of the proportion of language attributable to each LIWC category. This approach can  
252 be written out in the following way:

---

<sup>2</sup> <https://radimrehurek.com/gensim/>

<sup>3</sup> <https://web.sas.upenn.edu/danielpr/resources>



253 
$$p(\text{category} | \text{subject}) = \frac{\sum_{\text{word} \in \text{category}} \text{freq}(\text{word}, \text{subject})}{\sum_{\text{word} \in \text{vocab}(\text{subject})} \text{freq}(\text{word}, \text{subject})}$$

254  
 255 where  $\text{freq}(\text{word}, \text{subject})$  is the count where the message contains the *word* and the  $\text{vocab}(\text{subject})$   
 256 is the entire list of words mentioned by the subject, i.e. Twitter user.  
 257

258 *LA-c) Theory-based analysis*

259 Based on a conceptual and thematic review of qualitative studies on early adult crisis, 20  
 260 central concepts were identified as linguistic features expected to be mentioned in social media  
 261 reference to QLC (Robinson, Wright & Smith, 2013; Robinson, 2018). The terms associated with  
 262 the QLC based on the holistic model of early adult crisis are: *Stuck; Trying; Leave; Change;*  
 263 *Unemployed; Lonely; Hopeless; Overwhelmed; Unfair; Fail; Coping; Failing; Debt; Meaning;*  
 264 *Trapped; Try; New; Identity; Sacked; Money.* These terms were analyzed against the data in a  
 265 similar manner as LIWC (LA-b).

266 *Identifying differentially expressed language features during QLC*

267 To determine if linguistic attributes (dictionary-based and open-vocabulary) and theory-  
 268 based words were associated with QLC group, we individually tested them as a predictor in an in-  
 269 sample logistic regression model, and report its standardized regression coefficient ( $\beta$ ) with the  
 270 associated significance. We used Benjamini-Hochberg p-correction for multiple comparisons and  
 271 use  $p < 0.05$  as a heuristic for identifying potentially meaningful correlations; the effect size was  
 272 measured using Cohen's D. Demographic variables such as age and gender are included as  
 273 covariates to obtain a unique effect of the language variables. Since we explored several features  
 274 simultaneously, we consider coefficients significant if they are less than a Bonferroni-corrected  
 275 two-tailed  $p$  value of 0.05. This sets an extremely stringent level for significance. So for example,  
 276 when examining 20,000 features, in the case of words and phrases, a required  $p$ -value is less than  
 277 0.05 divided by 20,000 which is  $2.5 \times 10^{-6}$ , or when examining 200 topics the required  $p$ -value is  
 278 less than  $2.5 \times 10^{-4}$ , and when examining 73 LIWC categories  $p$ -value is less than  $6.8 \times 10^{-4}$ ).

279 **Results**

280 *Open-vocabulary clustering approach*

281 Based on the open-vocabulary AI-based analysis using vector-based graphic  
 282 representations of term clusters that correlated with QLC, Figure 1 shows the most prominent  
 283 words and phrases in the Twitter messages posted by the QLC group compared with the control  
 284 group. In the figure, word size represents the strength of the correlation to QLC and word color  
 285 indicates relative word frequency.

286 Words relating to time ("night"; "weekend"; "morning"; "early"; "day") and work  
 287 ("work"; "working") had the highest frequency and correlation strengths. Also, a pronounced use  
 288 of first person pronouns ("I"; "my"; "myself") was observed in users going through QLC.

289 References focused on reflection and a willingness to conduct activities ("i\_should", "i\_could",  
290 "i\_can") have been identified.

291 Figure 2 shows topics that are associated with QLC. Each individual box represents an  
292 output from the Word2Vec vector analysis. These were grouped into four thematic categories by  
293 the authors for the purposes of parsimonious presentation. Every topic in each category was  
294 significantly associated with QLC at  $p < .05$  after Bonferroni p-correction. Figure 2a highlights  
295 topics of everyday life including issues of timing, exercise, fitness, travelling, sports, domestic  
296 settling down, and alcohol. Figure 2b indicates the emotional dysfunction that accompanies the  
297 strife and confusion of QLC. The range of emotions extends from positive ("awesomest",  
298 "#ilovemylife") to negative ("sadness", "agitated") sentiments. In addition to expressing  
299 themselves through emotional words, the QLC group tends to emote through elongated words  
300 (Figure 2c). Elongation is common in social media and provides nuance to digital communications  
301 by mimicking intonation found in vocal exchanges (Doll, 2013). A final cluster of topics (Figure  
302 2d) includes lifestyle and health issues, including employment, mental health problems, physical  
303 illness, family, food and studying.

304

### 305 *Linguistic Inquiry Word Count (LIWC) analysis*

306 LIWC conducts all inferential analysis using Cohen's  $d$ . Effect size between QLC and  
307 LIWC linguistic categories are shown in Table 2, grouped into categories for readability.  
308 Categories that were significant at  $p < 0.05$  after Bonferroni correction are shown in the table.  
309 Positive effect size represents an association between the category and the presence of QLC.

310 For QLC individuals, Time ( $d = 1.317$ ) had the largest effect size LIWC category followed  
311 by Space ( $d = 1.080$ ) and Motion ( $d = .718$ ) under Relativity, potentially indicating the phases of  
312 transition with school, career, and locations in emerging adults. Replicating previous findings (Jay,  
313 2012), we found a high effect size with Present Focus ( $d = .928$ ) relating to the present bias of  
314 QLC individuals. The large effect sizes associated with Biological Processes, specifically  
315 Ingestion ( $d = .958$ ) and Health ( $d = .713$ ) aligns with some of the previously identified common  
316 pressure pain points associated with QLC (Panchal & Jackson, 2007). QLC individuals also tend  
317 to use language associated with Cognitive Processes, including Certainty ( $d = .592$ ), Discrepancy  
318 ( $d = .568$ ), Tentativeness ( $d = .501$ ), indicative of mixed reactions and emotions. Similar  
319 observations can be made about QLC's association with grammatical categories such as  
320 Comparisons ( $d = .747$ ) and Interrogatives ( $d = .583$ ). As a part of emerging adulthood, individuals  
321 consider where to live, how to live, and what to do, consequently producing language associated  
322 with Personal Concerns, specifically Home ( $d = .850$ ), Leisure ( $d = .475$ ), Work ( $d = .432$ ), and  
323 Feelings ( $d = .676$ ). Language of QLC individuals has a high association with Anxiety ( $d = .631$ )  
324 and Sadness ( $d = .501$ ), and Risk ( $d = .453$ ). Other effect size with Nonfluencies ( $d = .534$ ) and  
325 Function words (e.g. 1st person pronouns,  $d = .774$ ) are in line with the results from the open  
326 vocabulary approach.

327

### 328 *Theory-based analysis*

329 In order to test the hypothesis that the 20 selected theoretical terms would be more prevalent in  
330 the QLC group than the control group, a language analysis was conducted using the same

331 method as LIWC. The results in Table 3 show which theoretically derived terms were  
332 significantly higher for the QLC group than the control group. Sixteen of the twenty terms were  
333 significantly different across QLC and control group, providing robust evidence that discussions  
334 of crisis on Twitter show a linguistic fit with what is known about crisis in young adults from a  
335 theoretical and empirical standpoint.

## 336 Discussion

337 This study leveraged social media data to uncover the online word usage of Twitter users  
338 from the UK and the USA who report experiencing a QLC. Using an open vocabulary AI-based  
339 clustering method, we hypothesized that we would find differences between those reporting QLC  
340 and matched non-QLC control group. This was supported, and the results shown in Figures 1 and  
341 2 show clear links with the theory of emerging adulthood and the holistic model of early adult  
342 crisis. The most common words in Figure 1 were *work, time, night, weekend* and *my*. *Work* was  
343 the word most strongly associated with QLC. This fits with the fact that accounts of QLC mainly  
344 revolve around problems with finding, or adapting to, work. In a large quantitative survey of crisis  
345 features, the most prevalent features in early adult crisis amongst men were “Feeling trapped in a  
346 job you didn’t want to be in any more”, and “Experiencing a high level of stress and pressure in  
347 your job” (Robinson & Wright, 2013). These features were also very prevalent amongst women.  
348 The association of the word *time* with QLC is illuminated by the top right cluster in Figure 2,  
349 which shows a future focus, with words like *tomorrow, preparing* and *anticipating*. This fits with  
350 one of the five core features of emerging adulthood, which is an optimistic preoccupation with the  
351 future. With regards to the use of first person pronouns in QLC tweets relative to the control (*my,*  
352 *myself, I*) as shown in Figure 1, this fits with previous findings showing a heightened use of  
353 personal pronouns on social media by users with mental health issues relative to a control (De  
354 Choudhury, Gamon, Counts, & Horvitz, 2013). It also fits with the theoretical postulates of both  
355 emerging adulthood and the model of early adult crisis. The former suggests that young people  
356 who are passing through emerging adulthood have a relatively strong self-focus compared with  
357 adults of other age groups (Arnett, 2000). If QLC is partly used in social media as a proxy for the  
358 challenges of emerging adulthood, as we have hypothesized, one would expect a high level of  
359 usage of personal pronouns in the social media language of QLC. The use of personal pronouns  
360 fits with early adult crisis theory insofar as the latter purports that crises involve an extended  
361 questioning of identity, in terms of ‘who I am’ in the context of society, roles and relationships  
362 (Robinson, Wright & Smith, 2013).

363 The general topics represented in Figure 2a - exercise, travel, alcohol, sport, time - reflect  
364 many of the topics that emerging adults engage in to both cope with stress and find meaning  
365 (Arnett, 2014). The time-related words in the top right box of Figure 2a include mainly future-  
366 focused terms, such as *tomorrow, preparations* and *anticipating*, as previously discussed. Exercise  
367 and fitness are effective ways of managing stress, so may be linked to QLC as coping strategies  
368 (Cairney, Kwan, Veldhuizen, & Faulkner, 2014). On the flipside, alcohol usage is also linked to  
369 QLC. Epidemiological research shows that alcohol consumption peaks in emerging adulthood  
370 (McManus, Bebbington, Jenkins & Brugha), and it has been theorized that this may be a form of  
371 self-medication for stress and anxiety (Cooper, Russell, Skinner, Frone, & Mudar, 1992). The

372 cluster of terms on tourism and travelling relates to the phase of exploration in the holistic model  
373 of QLC, which often involves taking time-out from long-term commitments to go abroad, with the  
374 aim of getting perspective on one's current life circumstances and priorities (Robinson et al.,  
375 2013). The cluster of terms on domestic furnishing fits with the fact that early adult crisis tends to  
376 occur just as a young adult is making active attempts to settle down and develop a stable lifestyle  
377 (Robinson, 2016). Many of these topics that associate with QLC are likely to be framed by cultural  
378 factors of the UK and the USA, for example the relative affluence and high employment rates of  
379 these countries, the relatively high tertiary education participation, the high stress levels in young  
380 adults (Forth, 2018; Stone et al., 2010). The potential generalization of the current study that we  
381 tentatively claim is to young adults within these two countries. While QLC is a phenomenon that  
382 is discussed in other cultures such as India and China (e.g. Mehta, 2008, Mei, 2017), the linguistic  
383 associations of the phenomenon in these other countries may differ substantively.

384 Social media captures users' emotions in an ecological and relatively immediate setting  
385 (Suler, 2014). Our finding that expressing *mixed* emotions (Figure 2b) associates with QLC  
386 supports the affective strand of the holistic theory of early adult crisis, which represents emotions  
387 as during a crisis episode as a combination of emotional conflict and negativity but also times of  
388 excitement, hope, and fun (Robinson, Wright, & Smith, 2013). The same mix of positivity and  
389 negativity has been found in other previous studies of QLC (Black, 2010; Panchal et al., 2007;  
390 Robinson, 2018). Furthermore, employment and family (Figure 2d) have been identified as  
391 pressure points during transitions in emerging adulthood (Panchal et al., 2007). The cluster of  
392 mental health terms in Figure 2d fits with the acknowledged overlap between QLC and mental  
393 illness, such that periods of crisis are times of heightened vulnerability for mental illness,  
394 particularly if the individual does not enact changes that permit resolution of the crisis, so the  
395 difficulties and instabilities associated with the crises become chronic (Robinson, 2016).

396 QLC episodes are periods of high intensity experience, in terms of major decisions to take,  
397 challenging problems to surmount, and strong emotions to manage. A novel finding from the study  
398 is the association between QLC and word elongations such as *meeeeee*, *pleaseeeee*, *yeaahhhhh*,  
399 *reeaallly*, *sooooo* and *yaaaaaaay*. We propose that this form of spelling idiosyncrasy may be used  
400 to convey intensity of experience - by expanding the word in size, the strength of its meaning is  
401 enhanced to the reader. They could even be considered a marker for possible QLC - an apparently  
402 mundane linguistic quirk that may have developmental meaning. Further research could explore  
403 this in relation to other high intensity life transitions.

404 With regards to LIWC findings, it was found that QLC is related to words referring to time,  
405 change, and movement. This reflects how QLC is often a time of transitional change and active  
406 movement. It was also found that QLC is associated with being focused on both the present  
407 moment and the future, reflecting how immediate concerns to cope with and struggle against  
408 pressing challenges draw attention to the present moment, and also to the question of where life is  
409 going. This finding reflects work by Meg Jay (Jay, 2012) on the 'present bias' of young adults  
410 more generally - such a bias may be amplified in times of QLC. The association with words about  
411 eating and health is likely to refer to concerns about well-being, stress and health that have been  
412 regularly found to precipitate a developmental crisis (Robinson, 2016) or create a sense of  
413 existential concern in young adults (Panchal et al., 2007). The other key categories; insight, feeling,

414 home, and anxiety also reflect various facets of what is known about QLC. Anxiety and feeling  
415 words reflect the strong affective content of crisis episodes; insight reflects the heightened  
416 curiosity and questioning that has been found to be present in crisis episodes (Robinson et al.,  
417 2017); home reflects the central issue in QLC of ‘where I fit in’ to the world and where one will  
418 end up living as an adult (Robinson, 2018).

419 The theory-led term analysis shown in Table 3 strongly supports the proposition that QLC  
420 links to the theoretical model of early adult crisis (Robinson, 2016). Sixteen of the twenty terms  
421 predicted to associate with QLC from this theory (*stuck, trying, leave, change, unemployed, lonely,*  
422 *hopeless, overwhelmed, unfair, fail, coping, failing, debt, meaning, trapped* and *try*) did indeed  
423 link to Twitter postings about it. The four words that did not were *new, identify, sacked* and *money*.  
424 Our interpretations of the absence of these four are speculative. It may be being *sacked* is more  
425 likely to be a feature of crisis in midlife or that *fired* is a more common term of use among young  
426 adults, and while debt is a clear feature of crisis, *money* was not associated perhaps because of its  
427 value neutrality. *Identify* may be too abstract a verb for Twitter postings, and *new* may be used in  
428 too wide a range of ways to make it a crisis differentiator.

429 Our use of data from Twitter has introduced a new lens on the analysis of the QLC and  
430 how the challenges of being a young adult are discussed on social media. There are however  
431 various limitations to this study. The individuals who post on Twitter may not represent a full  
432 range of socioeconomic status groups or may be systematically different from the general  
433 population in other unspecified ways. Hence, those who post about QLC on Twitter may not  
434 represent those who would report having QLC in the general population. However, given the  
435 naturalistic conditions of data collection and the relatively large sample of this study, it is arguably  
436 more likely to generalize than most surveys. We do not assume that QLC is a phrase used in other  
437 languages, so we do not generalize these findings to non-English-speaking countries. In addition,  
438 the study filtered for English-only tweets, irrespective of cultural origination, which could have  
439 introduced cultural confounds in language.

440 Finally, while we have framed our interpretation of the results through the theory of  
441 emerging adulthood and the model of early adult crisis, we do not have the scope here to  
442 systematically compare these with other potential theories as interpretative frameworks. We claim  
443 a good fit between data and theory in a host of ways, but other theories may also provide helpful  
444 abductive schemes.

445 Notwithstanding these limitations, this is the first academic study known to use artificial  
446 intelligence and social media to study the discourse surrounding QLC at scale. *Theory on emerging  
447 adulthood and early adult crisis has not previously made use of Big Data to test theoretical  
448 postulates and explore new areas. We argue that the study provides a new empirical lens on the  
449 developmental challenges that young adults experience and the language used to frame  
450 experiences on social media.* Important next steps for further research include exploring how the  
451 link between QLC and language is moderated by gender, age, geographical location and ethnicity.  
452 Another option for future studies includes using the same methodology to analyze midlife crisis  
453 and later life crisis. Finally, a longitudinal analysis would be revealing of how a mention of QLC  
454 predicts social media postings over time. It would be possible to take a group of individuals who  
455 mention QLC on Twitter for the first time and then explore the contents of their postings at several

456 time points (e.g. six months later and a year later) relative to a matched control group. As well as  
457 a stimulus for more research, the study has a number of possible practical applications.

458 Following further studies to ensure replicability and a fine-grained understanding of gender  
459 and culture, we intend to develop a guide for clinicians, coaches and university lecturers **in the UK**  
460 **and USA** to help understand how young adults verbally discuss their personal challenges with each  
461 other in the space of social media, and what kinds of issues and words used in tweets may be  
462 indicative of a personal crisis and hence in need ~~for~~of targeted support and help.

463

#### 464 **Figure Legend**

465 Figure 1: The most frequently used words for those who describe themselves as going through a quarter-  
466 life crisis.

467 Word size indicates the strength of the correlation to QLC and word colour indicates relative word  
468 frequency ( $p < 0.05$ , Bonferroni p-corrected)

469

470 Figure 2: Highly correlated topics (grouped into categories) significantly associated with quarter-life crisis  
471 at  $p < 0.05$  after Bonferroni p-correction

472 2a: General Topics (Time, Exercise, Travel, Entertainment)

473 2b: Mixed Emotions

474 2c: Elongation

475 2d: Career, Health, School, Family.

476

477

478 **References**

- 479 Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through  
480 the twenties. *American Psychologist*, 55(5), 469-480.  
481 <https://www.apa.org/pubs/journals/amp/>
- 482 Arnett, J. J. (2007). Emerging adulthood: What is it, and what is it good for? *Child Development*  
483 *Perspectives*, 1(2), 68–73. doi: 10.1111/j.1750-8606.2007.00016.x
- 484 Arnett, J. J. (2016). Does emerging adulthood theory apply across social classes? National data  
485 on a persistent question. *Emerging Adulthood*, 4(4), 227–235.  
486 <https://doi.org/10.1177/2167696815613000>
- 487 Arnett, J.J. (2014). *Emerging Adulthood: The winding road from the late teens through the*  
488 *twenties*. New York: Oxford University Press.
- 489 Arnett, J. J., & Mitra, D. (2018). Are the features of emerging adulthood developmentally  
490 distinctive? A Comparison of Ages 18–60 in the United States. *Emerging Adulthood*.  
491 <https://doi.org/10.1177/2167696818810073>
- 492 Black, A. S. (2010). “*Halfway between somewhere and nothing*”: A exploration of the quarter-  
493 *life crisis and life satisfaction among graduate students*. Charleston: Proquest.
- 494 Cairney, J., Kwan, M. Y., Veldhuizen, S., & Faulkner, G. E. (2014). Who uses exercise as a  
495 coping strategy for stress? Results from a national survey of Canadians. *Journal of*  
496 *Physical Activity and Health*, 11(5), 908-916. doi:10.1123/jpah.2012-0107
- 497 Cooper, M. L., Russell, M., Skinner, J. B., Frone, M. R., & Mudar, P. (1992). Stress and alcohol  
498 use: moderating effects of gender, coping, and alcohol expectancies. *Journal of Abnormal*  
499 *Psychology*, 101(1), 139. <https://www.apa.org/pubs/journals/abn/>
- 500 Coppersmith, G., Dredze, M., & Harman, C.M. (2014). Quantifying Mental Health Signals in  
501 Twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical*  
502 *Psychology: From Linguistic Signal to Clinical Reality*, (pp. 51-56). Association for  
503 Computational Linguistics.
- 504 De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via  
505 social media. *ICWSM*, 13, 1-10. Web
- 506 Doll, J. (2013). Why Drag It Out?. *The Atlantic*, 228. <https://www.theatlantic.com/world/>
- 507 Forth (2018). *Great Britain and Stress - How bad is it and why is it happening?* Retrieved from  
508 <https://www.forthwithlife.co.uk/blog/great-britain-and-stress/>
- 509 Georgakopoulou, A. (2017). Sharing the moment as small stories: The interplay between  
510 practices & affordances in the social media-curation of lives. *Narrative Inquiry*, 27 (2)  
511 311-333. doi:10.1075/ni.27.2.06geo
- 512 Giles, D.C. (2017). How do fan and celebrity identities become established on Twitter? A study  
513 of ‘social media natives’ and their followers. *Celebrity Studies*, 8(3), 445-460.  
514 doi:10.1080/19392397.2017.1305911
- 515 Guntuku, S. C., Lin, W., Carpenter, J., Ng, W. K., Ungar, L. H., & Preoțiu-Pietro, D. (2017).  
516 Studying personality through the content of posted and liked images on Twitter. In  
517 *Proceedings of the 2017 ACM on web science conference* (pp. 223-227). ACM.
- 518 [Guntuku, S. C., Buffone, A., Jaidka, K., Eichstaedt, J. C., & Ungar, L. H. \(2019, July\).](#)

- 519            *Understanding and measuring psychological stress using social media*. In Proceedings of  
520 the International AAAI Conference on Web and Social Media (Vol. 13, No. 01, pp. 214-  
521 225).
- 522 Guntuku, S. C., Preotiuc-Pietro, D., Eichstaedt, J. C., & Ungar, L. H. (2019, July). *What twitter*  
523 *profile and posted images reveal about depression and anxiety*. In Proceedings of the  
524 International AAAI Conference on Web and Social Media (Vol. 13, No. 01, pp. 236-  
525 246).
- 526 Guntuku, S. C., Lin, W., Carpenter, J., Ng, W. K., Ungar, L. H., & Preotiuc-Pietro, D. (2017,  
527 June). *Studying personality through the content of posted and liked images on Twitter*. In  
528 Proceedings of the 2017 ACM on web science conference (pp. 223-227).
- 529 Guntuku, S.C., Ramsay, J.R., Merchant, R.M., Ungar, L.H. (2017). Language of ADHD in  
530 Adults on Social Media. *Journal of Attention Disorders*.  
531 doi:10.1177/1087054717738083.
- 532 Guntuku, S.C., Yaden, D.B., Kern, M.L., Ungar, L.H., Eichstaedt, J.C. (2017). Detecting  
533 depression and mental illness on social media: an integrative review. *Current Opinion in*  
534 *Behavioral Sciences*, 18, 43-49. doi:10.1016/j.cobeha.2017.07.005.
- 535 Jay, M. (2012). *The defining decade: why your twenties matter--and how to make the most of*  
536 *them now*. New York: Twelve.
- 537 Hapke, C. (2017). *How To Apply Emotionally Focused Couple Therapy To Couples*  
538 *Experiencing a Quarter-Life Crisis*. (Doctoral dissertation, Alliant International  
539 University). Retrieved from ProQuest Dissertations Publishing.
- 540 Lampos, V., Aletras, N., Preotiuc-Pietro, D., & Cohn, T. (2014). Predicting and characterising  
541 user impact on Twitter. In *Proceedings of the 14th Conference of the European Chapter*  
542 *of the Association for Computational Linguistics* (pp. 405-413). Association for  
543 Computational Linguistics.
- 544 Kivran-Swaine, F., Brody, S., Diakopoulos, N., & Naaman, M. (2012). *Of joy and gender:*  
545 *Emotional expression in online social networks*. Proceedings of the ACM Conference on  
546 Computer Supported Cooperative Work, CSCW. 139-142. 10.1145/2141512.2141562.
- 547 Konstam, V. (2007). *Emerging and Young Adulthood: Multiple Perspectives, Diverse*  
548 *Narratives*. Springer, Boston, MA.
- 549 McManus S., Bebbington P., Jenkins R., Brugha T. (Eds.) (2016) *Mental health and wellbeing in*  
550 *England: Adult Psychiatric Morbidity Survey 2014*. Leeds: NHS Digital.
- 551 Mehta, N.T. (2008). *'Quarter-life crisis' grips Young India*. Retrieved from:  
552 [https://www.hindustantimes.com/entertainment/quarter-life-crisis-grips-young-](https://www.hindustantimes.com/entertainment/quarter-life-crisis-grips-young-india/story-WAeN4Ro3AMWzPo6vPARTCP.html)  
553 [india/story-WAeN4Ro3AMWzPo6vPARTCP.html](https://www.hindustantimes.com/entertainment/quarter-life-crisis-grips-young-india/story-WAeN4Ro3AMWzPo6vPARTCP.html)
- 554 Mei, Y. (2017). *China's millennials are dealing with a quarter-life crisis*. Retrieved from  
555 <https://thriveglobal.com/stories/china-s-millennials-are-dealing-with-a-quarter-life-crisis/>
- 556 Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word  
557 representations in vector space. Poster session presented at the International Conference  
558 on Learning Representations.
- 559 Mikolov, T., Yih, W. T., & Zweig, G. (2013). Linguistic regularities in continuous space word



560 representations. *In Proceedings of the 2013 Conference of the North American Chapter*  
561 *of the Association for Computational Linguistics: Human Language Technologies* (pp.  
562 746-751). Association for Computational Linguistics: Human Language Technologies.

563 Munsey, C. (2006). Emerging adults: The in-between age. *American Psychological Association*  
564 *Monitor*, 37 (7), 68. Retrieved 8th May 2019 from  
565 <https://www.apa.org/monitor/jun06/emerging>

566 Nelson, L. J., & Barry, C. M. (2005). Distinguishing features of emerging adulthood: The Role  
567 of self-classification as an adult. *Journal of Adolescent Research*, 20(2), 242–262.  
568 <https://doi.org/10.1177/0743558404273074>

569 Panchal, S., & Jackson, E. (2007). Turning 30 transitions: Generation Y hits quarter-life. *The*  
570 *Coaching Psychologist*, 3(2), 46-51. [https://shop.bps.org.uk/publications/publication-by-](https://shop.bps.org.uk/publications/publication-by-series/the-coaching-psychologist.html)  
571 [series/the-coaching-psychologist.html](https://shop.bps.org.uk/publications/publication-by-series/the-coaching-psychologist.html)

572 Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The Development and  
573 Psychometric Properties of LIWC2015. *University of Texas at Austin*. doi:  
574 10.15781/T25P41

575 Perelli-Harris, B., & Lyons-Amos, M. (2015). Changes in partnership patterns across the life  
576 course: An examination of 14 countries in Europe and the United States. *Demographic*  
577 *Research*, 33, 145-178. DOI: 10.4054/DemRes.2015.33.6

578 Rettberg, J. W. (2009). ‘Freshly Generated for You, and Barack Obama’: How Social Media  
579 Represent Your Life. *European Journal of Communication*, 24(4), 451–466.  
580 doi:10.1177/0267323109345715

581 Robbins, A., & Wilner, A. (2001). *Quarterlife crisis: The unique challenges of life in your*  
582 *twenties*. London: Penguin.

583 Robinson, O.C. (2016). Emerging adulthood, early adulthood and quarter-life crisis: Updating  
584 Erikson for the twenty-first century. In R. Žukauskiene (Ed.) *Emerging adulthood in a*  
585 *European context* (pp.17-30). New York: Routledge.

586 Robinson, O. C. (2019). A longitudinal mixed-methods case study of quarter-life crisis  
587 during the post-university transition: Locked-Out and Locked-In forms in combination.  
588 *Emerging Adulthood*, 7(3), 167–179. doi:10.1177/2167696818764144

589 Robinson, O. C., Demetre, J. D., & Litman, J. A. (2017). Adult life stage and crisis as predictors  
590 of curiosity and authenticity: Testing inferences from Erikson’s lifespan theory.  
591 *International Journal of Behavioral Development*, 41(3), 426–431.  
592 doi:10.1177/0165025416645201

593 Robinson, O. C., & Smith, J. A. (2010). Investigating the form and dynamics of crisis episodes in  
594 early adulthood: The application of a composite qualitative method. *Qualitative Research*  
595 *in Psychology*, 7(2), 170-191. doi: 10.1080/14780880802699084

596 Robinson, O. C., Wright, G. R., & Smith, J. A. (2013). *The holistic phase model of early adult*  
597 *crisis*. *Journal of Adult Development*, 20(1), 27-37. doi:10.1007/s10804-013-9153-y

598 Robinson, O. C., & Wright, G. R. T. (2013). The prevalence, types and perceived outcomes of  
599 crisis episodes in early adulthood and midlife: A structured retrospective-  
600 autobiographical study. *International Journal of Behavioral Development*, 37(5), 407–  
601 416. doi:10.1177/0165025413492464

602 Rosen, E. (2019). *The Radical Sabbatical: The millennial handbook to the quarter life crisis*.  
603 London: John Catt.

604 Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M.,  
605 ... Ungar, L. H. (2013). Personality, gender, and age in the language of social media: the  
606 open-vocabulary approach. *PloS one*, 8(9), e73791. doi:10.1371/journal.pone.0073791

607 Schwartz, H. A., Sap, M., Kern, M. L., Eichstaedt, J. C., Kapelner, A., Agrawal, M., ... &  
608 Kosinski, M. (2016). Predicting individual well-being through the language of social  
609 media. In *Biocomputing 2016: Proceedings of the Pacific Symposium* (pp. 516-527).

610 Schwartz, H. A., Giorgi, S., Sap, M., Crutchley, P., Ungar, L., & Eichstaedt, J. (2017). DLATK:  
611 Differential language analysis ToolKit. In *Proceedings of the 2017 Conference on*  
612 *Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 55-  
613 60). Association for Computational Linguistics.

614 Stapleton, A. (2012). Coaching clients through the quarter-life crisis: What works. *International*  
615 *Journal of Evidence Based Coaching and Mentoring, Special Issue, 6*, 130-145.  
616 <https://radar.brookes.ac.uk/radar/items/b6bb9783-f20a-44f6-9e07-f9bdf4437eb1/1/>

617 Stone, A. A., Schwartz, J. E., Broderick, J. E. and Deaton, A. (2010). A snapshot of the age  
618 distribution of psychological well-being in the United States. *Proceedings of the National*  
619 *Academy of Sciences of the United States of America, 107*, 9985–90.  
620 <https://doi.org/10.1073/pnas.1003744107>

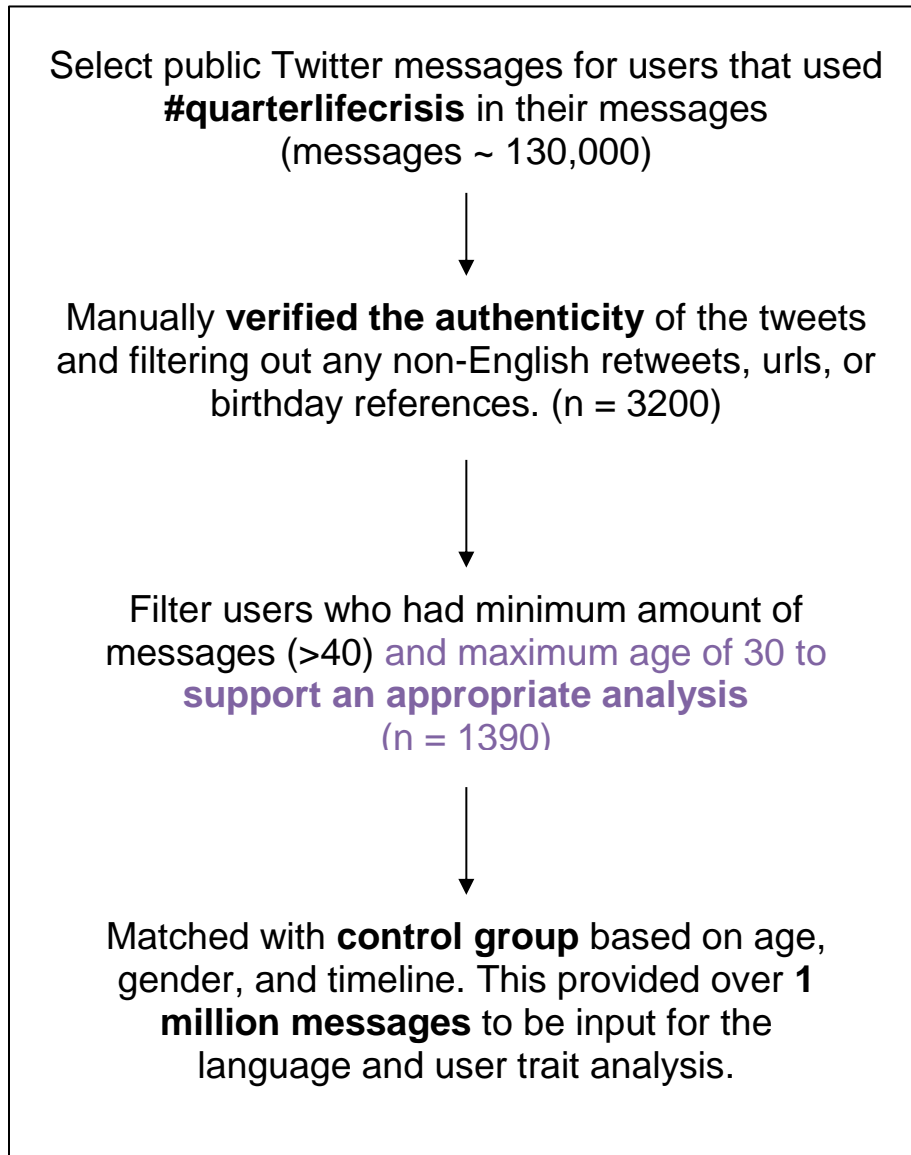
621 Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior, 7*(3), 321-326.  
622 doi:10.1089/1094931041291295

623 Toseeb, U., & Inkster, B. (2015). Online social networking sites and mental health research.  
624 *Frontiers in Psychiatry, 6*(36). doi: 10.3389/fpsy.2015.00036

625 van Dijck, J. (2007) *Mediated Memories in the Digital Age*. Stanford, CA: Stanford University  
626 Press.

627 Wiley, F. (2015). Technology, Social Networks, and the Quarter-Life Crisis. *Not Your Average*  
628 *Millennial*. Retrieved from [https://averagemillennial.com/2015/11/10/technology-social-](https://averagemillennial.com/2015/11/10/technology-social-networks-and-the-quarter-life-crisis)  
629 [networks-and-the-quarter-life-crisis](https://averagemillennial.com/2015/11/10/technology-social-networks-and-the-quarter-life-crisis)

630 Table 1A: Process on participant recruitment and selection



631

632

633

634

635 **Table 1B: Data Set of Control and QLC Groups – Gender and Mean Age**

<b>Group</b>	<b>Count</b>	<b>Male</b>	<b>Female</b>	<b>Mean Age</b>	<b>Standard Deviation</b>
Control	1390	195	1195	23.95	2.74
QLC Group	1390	195	1195	23.95	2.74
Total	2780	390 (14%)	2390 (86%)	23.95	2.74

636

637

638 **Table 2: Results from LIWC Categories significant after Bonferroni p-correction**

LIWC Category Name	Sample Words	Cohen d
<i>Relativity</i>		
Time	When, Now, New	1.317
Space	In, On, Out	1.080
Motion	Go, Going, Come	0.718
<i>Time Orientations</i>		
Present Focus	Is, Have, Are	0.928
Future Focus	Will, Then, Gonna	0.789
Past Focus	Was, Got, Been	0.779
<i>Biological Process</i>		
Ingestion	Eat, Sweet, Drunk	0.958
Health	Sick, Tired, Living	0.713
<i>Cognitive Processes</i>		
Insight	Know, Think, Find	0.842
Differentiation	If, Or, Can't	0.671
Certainty	All, Never, Ever	0.592
Discrepancy	Want, Need, Would	0.568
Tentative	Same, Hope, Any	0.501

Causation	How, Why, Make	0.473
<hr/>		
Other Grammar		
<hr/>		
Comparisons	Like, As, Than	0.747
Quantifiers	More, Some, Much	0.794
Common Adjectives	As, More, New	0.792
Interrogatives	What, How, Who	0.583
<hr/>		
Perceptual Processes		
<hr/>		
Feel	Feel, Hard, Feeling	0.67
<hr/>		
Personal Concerns		
<hr/>		
Home	Home, Bed, House	0.850
Leisure	Fun, Twitter, Play	0.475
Work	Work, School, Class	0.432
<hr/>		
Psychological Processes		
<hr/>		
<i>Negative Emotion</i>		
<hr/>		
Anxiety	Awkward, Worry, Scared	0.631
Sadness	Miss, Sorry, Sad	0.501
Risk	Bad, Stop, Wrong	0.453
<hr/>		
Informal Language		
<hr/>		
Nonfluencies	Oh, Well, Ugh	0.534
<hr/>		

## Linguistic Dimensions

### *Function Words*

Common Adverbs	So, Just, When	1.164
Prepositions	To, Of, In	1.145
Articles	The, A, An	1.005
Auxiliary Verbs	Is, Be, I'm	0.974
Conjunctions	And, So, But	0.827
Impersonal Pronouns	It, This, That	0.792
1st Pers Singular	I, My, Me	0.774
Negations	Not, Don't, No	0.559

639

640

641 **Table 3: Common Words associated with quarter-life crisis in the Twitter dataset.**

Rank	Label	d	p	*	Confidence interval		Frequency
					Lower	Upper	
1	Stuck	0.376	<0.001	*	0.149	0.22	5220
2	Trying	0.339	<0.001	*	0.131	0.203	22990
3	Leave	0.330	<0.001	*	0.127	0.198	13316
4	Change	0.173	<0.001	*	0.049	0.122	12531
5	Unemployed	0.158	<0.001	*	0.043	0.116	256

6	Lonely	0.158	<0.001	*	0.042	0.115	1678
7	Hopeless	0.144	<0.001	*	0.036	0.109	317
8	Overwhelmed	0.126	<0.001	*	0.026	0.1	530
9	Unfair	0.106	0.01	*	0.021	0.094	415
10	Fail	0.112	0.01	*	0.02	0.093	2642
11	Coping	0.102	0.01	*	0.015	0.088	105
12	Failing	0.100	0.01	*	0.013	0.087	635
13	Debt	0.096	0.01	*	0.012	0.085	552
14	Meaning	0.096	0.01	*	0.011	0.085	1399
15	Trapped	0.086	0.03	*	0.006	0.08	584
16	Try	-0.104	0.01	*	-0.088	-0.015	18692
17	New	-0.070	0.08		-0.072	0.001	75993
18	Identify	0.068	0.09		-0.003	0.07	564
19	Sacked	0.064	0.1		-0.005	0.069	84
20	Money	0.772	0.36		-0.055	0.018	13649

---

642 \* significance  $p < 0.05$