



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The Effect of User Psychology on the Content of Social Media Posts

Citation for published version:

Chen, LL, Magdy, W & Wolters, M 2020, 'The Effect of User Psychology on the Content of Social Media Posts: Originality and Transitions Matter', *Frontiers in Psychology*, vol. 11, 526.
<https://doi.org/10.3389/fpsyg.2020.00526>

Digital Object Identifier (DOI):

[10.3389/fpsyg.2020.00526](https://doi.org/10.3389/fpsyg.2020.00526)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Frontiers in Psychology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



The Effect of User Psychology on the Content of Social Media Posts: Originality and Transitions Matter

Lucia Lushi Chen^{1,*}, Walid Magdy¹ and Maria Wolters¹

¹*School of Informatics, University of Edinburgh, Edinburgh, United Kingdom*

Correspondence*:

Lucia Lushi Chen

lushi.chen@ed.ac.uk

2 ABSTRACT

3 Multiple studies suggest that frequencies of affective words in social media text are associated
4 with the user's personality and mental health. In this study, we re-examine these associations
5 by looking at the transition patterns of affect. We analyzed the content originality and affect
6 polarity of 4086 posts from 70 adult Facebook users contributed over two months. We studied
7 posting behaviour including silence periods when the user does not post any content. Our results
8 show that more extrovert participants tend to post positive content continuously, and that more
9 agreeable participants tend to avoid posting negative content. We also observe that participants
10 with stronger depression symptoms posted more non-original content. We recommend that
11 transitions of affect pattern derived from social media text and content originality should be
12 considered in further studies on mental health, personality, and social media.

13 **Keywords:** affect, social media, emotion, Facebook, personality traits, depression, mental health

1 INTRODUCTION

14 Many people express rich moods and emotions in their social media posts. Psychologists use the word
15 "affect" to describe these experiences of feelings and emotions. Affect plays an important role in cognition
16 (Gross et al., 1998) and wellbeing (Silvera et al., 2008). Therefore, affective expressions on social media
17 text have emerged as a key variable for making inferences about users' personality traits (Bachrach et al.,
18 2012; Golbeck et al., 2011; Farnadi et al., 2013) or mental health (De Choudhury et al., 2013; Coppersmith
19 et al., 2014; De Choudhury and De, 2014; Bazarova et al., 2015).

20 Existing studies formulate the associations between affect and wellbeing based on the frequencies of
21 affective words used in social media text (Schwartz et al., 2013; Yarkoni, 2010; Golbeck et al., 2011; Park
22 et al., 2015; Chen et al., 2020). However, patterns of affect are an important class of symptoms of affective
23 disorders (Rottenberg, 2005; Frijda, 1993; Bylsma et al., 2011; Sheppes et al., 2015; Thompson et al.,
24 2012; Houben et al., 2015; Carlo et al., 2012). Personality may also predispose individuals to specific
25 moods (Rusting and Larsen, 1995; Rusting, 1998). With this in mind, we examined how patterns of affect
26 expressed in social media text is related to with users' mental health and personality.

27 While non-original content has been extensively studied in opinion mining (Agarwal et al., 2011; Balahur
28 et al., 2009), it has been comparatively neglected in the study of psychological interpretations of social
29 media data. However, social media users often use lyrics or quotes to communicate their emotions. Such

30 content comes from other media, such as literature, videos, films, or music, which can evoke strong
31 emotional experiences (Juslin and Laukka, 2004; Scherer et al., 2001; Scherer, 2004). Since the affect
32 of the non-original content may be different from the social media users' affect when they are post this
33 content, we differentiated between original and non-original content in our analysis.

34 This pilot study was designed to examine the following research questions:

- 35 1. **Changes in Affect:** To what extent do changes in the affect of social media posts correlate with users'
36 personality traits and mental wellbeing?
- 37 2. **Originality:** To what extent does the use of non-original material in their posts correlate with users'
38 personality traits and mental wellbeing?

39 Following best practice in sentiment analysis and opinion mining, we distinguish between positive, negative,
40 neutral, and mixed (both positive and negative) affect (Moilanen and Pulman, 2007; Rosenthal et al., 2015;
41 Agarwal et al., 2011).

42 We used a well known dataset, myPersonality (Bachrach et al., 2012; Youyou et al., 2015), that enriches
43 Facebook posts with many validated psychological measures. In MyPersonality, positive mental wellbeing
44 is measured using the Satisfaction with Life Scale (Diener et al., 1985, 1999), while the presence of
45 depressive symptoms is assessed using the Centre for Epidemiologic Studies Depression scale (CES-D)
46 (Radloff, 1977). Personality traits are established following the OCEAN model (McCrae and John., 1992),
47 which consists of the five traits Openness to Experience, Conscientiousness, Extroversion, Agreeableness,
48 and Neuroticism.

49 We included all 70 adult users who provided sufficient, regular Facebook data for two months before
50 completion of the CES-D questionnaire, and corrected for multiple comparisons in our statistical analysis.
51 We find that the transitions from one affective state to another expressed in social media posts give us a
52 highly nuanced view of personality traits. While the amount of non-original posts in ones' social media
53 status updates is closely linked to depression symptoms, this link is mediated by neuroticism.

2 BACKGROUND

54 Affect refers to both mood and emotion. Moods are slow-moving states that can be influenced by people,
55 objects or situations, whereas emotions are quick reactions to stimuli (Rottenberg and Gross, 2003; Watson,
56 2000), and highly situation- or object-specific (Bylsma et al., 2008). Mood influences the probability of
57 having emotions of the same valence—negative mood facilitates negative emotions, and positive mood
58 makes positive emotions more likely (Rottenberg, 2005; Fredrickson, 1998). Affect is an important
59 predictor of mental wellbeing, including a person's overall satisfaction with life (Headey et al., 1993;
60 Singh and Jha, 2008; Chen et al., 2017), and the level of symptoms of depression (Tsugawa et al., 2015;
61 Coppersmith et al., 2015; Resnik et al., 2015).

62 Personality also predisposes people to certain affective states (Rothbart et al., 2000). While neuroticism
63 is associated with negative affect (Pishva et al., 2011), positive affect is strongly linked to extroversion
64 (Watson and Clark, 1997; Fujita et al., 1991). Extroverts experience more positive affect because they
65 engage in more social situations (Diener and Emmons, 1984; Ryan and Deci, 2001). Individuals who
66 score high on agreeableness have a greater ability to regulate negative affect (Meier et al., 2006; Haas
67 et al., 2007). This relationship between affect and personality is also reflected in social media studies (Lin
68 et al., 2017; Golbeck et al., 2011; Schwartz et al., 2013; Pennebaker and King, 1999). For example, people

69 who use negative affective words in their social media posts tend to have lower conscientiousness, lower
70 agreeableness (Golbeck et al., 2011), and higher neuroticism (Pennebaker and King, 1999).

71 In psychology, quantitative representations of affect are typically multidimensional (Russell, 1980). In
72 this study, we focus on valence, which is represented in many classic affect models. Traditional measures,
73 such as the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988), report the strength of
74 positive and negative valence. Mixed valence can occur when people experience ‘dialectic’ emotion, which
75 is a mix of positive and negative emotions (Russell, 2003; Schimmack et al., 2002).

76 The personality trait measurements in myPersonality are based on Costa and McCrae’s well-validated
77 OCEAN model (McCrae and John., 1992). The model consists of five dimensions: extroversion,
78 agreeableness, conscientiousness, neuroticism, and openness to experience. Neuroticism refers to the
79 degree of emotional stability. Openness reflects the degree of creativity and curiosity. Conscientious
80 individuals tend to be careful and diligent. Extroversion refers to a tendency to be energetic and friendly.
81 Agreeableness reflects the tendency to be compassionate and to cooperate with others (Digman, 1990).
82 The five-factor structure has proved to be robust in both self and peer ratings (McCrae and John., 1992),
83 children and adult (Mervielde et al., 1995), across different cultures (McCrae and Allik, 2002), and stable
84 over time (McCrae and John., 1992).

3 DATA AND METHODOLOGY

85 The myPersonality data set (Bachrach et al., 2012; Youyou et al., 2015) contains more than 180,000
86 Facebook users, enriched with a variety of additional validated scales (Bachrach et al., 2012). The collection
87 of myPersonality data complied with the terms of Facebook service, informed consent for research use
88 was obtained from all users, and researchers had to seek permission to use the dataset. Permission for the
89 use of this database was obtained before it closed for new studies in 2018. The study was granted Ethical
90 Approval by the Ethics Committee of the School of Informatics, University of Edinburgh.

91 3.1 Choice of Scales

92 From the extensive data collected within myPersonality, we chose two scales for quantifying mental
93 wellbeing, the *Center for Epidemiologic Studies Depression Scale (CES-D)* and the *Satisfaction with*
94 *Life Scale (SWL)*. The CES-D scale measures a key aspect of mental health, the presence of depression
95 symptoms (Radloff, 1977). The scale has high internal consistency, test-retest reliability (Radloff, 1977;
96 Orme et al., 1986; Roberts, 1980), and validity (Orme et al., 1986). Following previous social media studies
97 (De Choudhury et al., 2013; Park et al., 2012), we adopt a score of 22 or higher as a cut-off value for
98 likely depressive disorder (maximum score: 60). The 5-item SWL scale has been tested across different
99 cultures and age groups (Pavot and Diener, 2009) and has been found to have high internal consistency and
100 temporal reliability (Diener et al., 1985). Personality traits were measured using a 100 item scale using
101 items from the open source International Personality Item Pool (Goldberg et al., 2006) that were validated
102 against the the NEO-PI-R (Schwartz et al., 2013) instrument.

103 3.2 Selection of Participants

104 The data set was originally designed for a study of the effect of mental wellbeing and values on social
105 media disclosure. We therefore selected only those participants who had completed the CES-D scale, the
106 SWL scale, and the Schwartz Value survey (Schwartz, 1992) in addition to the full personality questionnaire.
107 301 participants in myPersonality provided full data for all four scales.

108 To ensure we had enough posts to assess the frequency of affect transitions, we only included users in
109 our sample that regularly updated their public Facebook feed (*regular users*). We defined regular users
110 as individuals who posted on average twice a week or more. We estimated posting frequency using the
111 average post count per day during the sampling frame. If an individual had a post count per day of 0.3, this
112 individual made around 110 posts in 365 days, which was roughly equivalent to an average of 2 posts per
113 week. Of the original 301 participants, 122 (40.5%) were regular users.

114 Since the CES-D asks about symptoms in the past week, we excluded a further 31 users who had not
115 posted any content in the week before completing the CES-D scale. We then focused on a 60-day span
116 (two months) before CES-D completion, to ensure we had sufficient data to track the development of users'
117 moods. We removed 14 users who contributed less than 20 posts during that time. Finally we removed four
118 users who were under 18 year old and three users with more than 20% of the posts written in a language
119 other than English, because English was the common language of the annotation team. The final sample
120 consisted of 4086 posts from 70 users.

121 3.3 Corpus Annotation

122 3.3.1 Social Media Affect

123 For the purpose of this study, we refer to the affect shown in social media posts as *social media affect*.
124 In this study, we operationalize valence as the post author's attitude towards a primary target of opinion,
125 following (Mohammad, 2016). We refer to the 'dialectic' affective state as *mixed valence*. If there is no
126 clear trend towards positive or negative affect, the associated valence is *neutral*.

127 After extensive piloting, we created an annotation guideline (available as part of the supplementary
128 material) that was largely based on Mohammad (2016)'s work on defining the valence of a social media
129 post. Each post is assigned one of four affect polarities: + (positive), - (negative), \pm (mixed), or 0 (neutral).
130 We used manual annotation since this is commonly used in computational linguistics to create a baseline
131 gold standard data set for further analysis (Teufel, 1999).

132 Out of the 4086 posts, 2698 (66%) were annotated by a team of six trained annotators and 1185 (29%) by
133 the first author. 5% of all posts were annotated by all seven annotators to establish inter-rater reliability,
134 which was measured using Cohen's κ (Gamer et al., 2012). Average inter-rater reliability between the first
135 author and the annotators is 0.88, and 0.78 among the six annotators.

136 After annotation, most of the posts were of positive valence (N= 1588, 39%), followed by negative
137 valence (N=1164, 28%), neutral valence (N=982, 24%) and mixed (N=312, 8%). 40 posts were excluded
138 from analysis, since they did not contain English text.

139 3.3.2 Originality

140 We define posts that consist of quotes from sources such as song lyrics, books, or movies as non-original
141 content; all other content was defined as original. Since non-original content might not directly reflect
142 the user's moods or emotions, annotators were instructed to annotate such posts according to the likely
143 emotions of the author. For example, if a post consists of an uplifting motivational quote, annotators
144 considered the underlying valence to be positive.

145 In order to establish the originality of a post, we retrieved the first page of results obtained by searching
146 for the post text using the Google API. For each web page on the first page of results, we computed the
147 cosine similarity between the the post content and the page content. Posts with a cosine similarity greater
148 than 0.96 were labeled as non-original, and posts with a cosine similarity between 0.92 and 0.96, where

Table 1. Affect and originality representation for a sample week

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
<i>Affect</i>							
Post-Level	+ - -	+ - +	++	S	±	0	+ -
Day-Level	-	+	+	S	±	0	±
<i>Originality</i>							
Post-Level	O N O	O O N	N N	S	O	O	N N

Note: \leftrightarrow , negative valence: -, positive valence: +, mixed valence: \pm , S: silence day, original content: O, non-original content: N

149 the website links or website names included the words ‘lyrics’ or ‘quote’ were labeled as potentially
 150 non-original. Posts with a cosine similarity lower than 0.92 were labeled as original. The cutoff points were
 151 determined based on a sample of 300 posts manually annotated for originality by the first author. On these
 152 posts, the classifier yields 100% recall, 81% precision, and an F1-score of 0.89. In our data set, 287 (7%)
 153 of all posts were identified as non-original.

154 3.4 Modelling Affect Transitions

155 We examine two types of transitions:

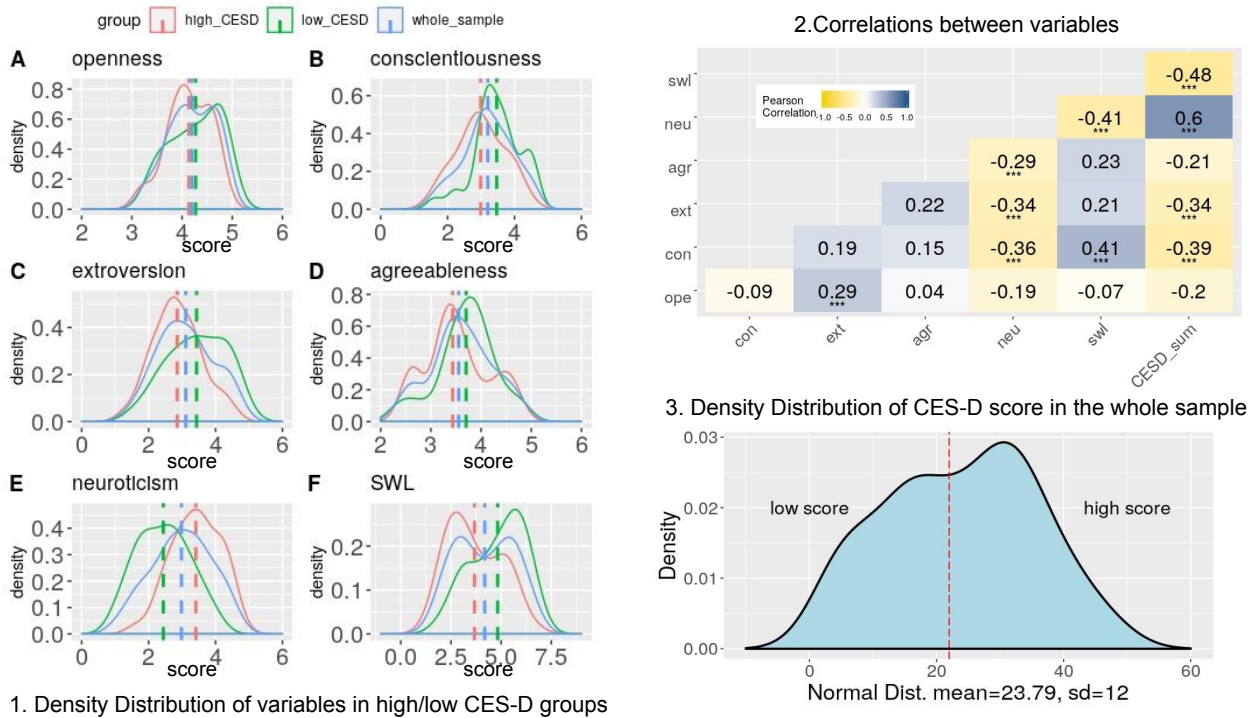
- 156 • **Post-Level versus Day-Level:** *Post-level* transitions focus on changes in affect between subsequent
 157 social media posts, whereas *day-level* transitions focus on changes in overall dominant affect between
 158 subsequent days.
- 159 • **Silence versus Non-Silence:** Not all users post every day. In our *default* models, these silent days are
 160 ignored, whereas in our *with-silence* models, days without posts are explicitly modelled as *Silence*.

161 The post-level social media affect is likely to be influenced by underlying emotions, which change
 162 more quickly, whereas the day-level social media affect is likely to be influenced by *underlying mood*
 163 during the day. Day-level affect was calculated as follows. If the majority of the posts p_{ij} on day d_j have the
 164 same affect a , then the affect of day d_j is set to a . If there is an equal number of positive (+) and negative
 165 (-) posts, or if the number of mixed affect (\pm) posts is equal to the number of posts with other types of
 166 affect, affect is set to \pm (mixed). For transitions between original and non-original posts, we only consider
 167 the post-level representation. Table 1 shows an example of the affect and originality representations.

168 3.5 Statistical Analysis

169 Demographic differences between users above and below the CES-D cut-off score for probable depression
 170 were assessed using Wilcoxon-Mann-Whitney tests (R-package ‘Stats’).

171 We used Pearson correlation coefficients to assess the significance of correlations between social media
 172 data on one hand and personality traits and mental wellbeing on the other hand. Due to the small sample size
 173 and the number of correlations computed, all correlation coefficients were estimated using a permutation
 174 approach (Higgins, 2003) as implemented in the R Package `jmuOutlier` (Garren, 2017). Correlations
 175 that reach $p < 0.01$ or better are reported as significant; correlations that reach $p < 0.05$ are reported as
 176 trends in the data. For all correlations reported in the paper, we give the estimated correlation coefficient,
 177 the bootstrap 95% confidence interval, and the corresponding coefficient of determination r^2 .



1. Density Distribution of variables in high/low CES-D groups

Figure 1. Basic statistics for personality trait scores, SWL and CES-D scores. Plot 1 shows density plots of the distribution of personality traits and SWL for all participants, participants with CES-D ≥ 22 (high CES-D), and participants with CES-D < 22 (low CES-D). The dotted line shows the median. Plot 2 is a heat map of correlations between personality traits, SWL and CES-D scores (***: $p < 0.001$). Plot 3 illustrates the distribution of the CES-score in the entire sample ($N = 70$). The dotted line indicates the cutoff score of 22.

4 RESULTS

178 4.1 Demographics and Baseline Statistics

179 Table 2 shows the basic statistics of our sample. Our data predominantly comes from single female
180 Caucasian young adults. The average CES-D score is above the cut-off for possible depressive disorder.

181 Thirty-nine (56%) participants had a CES-D score of 22 or higher (mean: 33, SD: 6.5), which means that
182 it is possible that they have depressive disorder, and 31 (44%) had a score of 21 or lower (mean: 12,
183 SD: 6). Figure 1 Plot 1 shows the density distributions of personality trait and SWL scores for three groups,
184 the full sample, people above the cut-off, people below the cut-off..

185 Participants with possible depressive disorder are less extroverted ($Z = 375$, $p < 0.005$), have higher levels
186 of neuroticism ($Z = 990$, $p < 0.001$), lower levels of conscientiousness ($Z = 375$, $p < 0.001$), and lower
187 satisfaction with life ($Z = 323$, $p < 0.001$). Detailed results are reported in Figure 1 Plot 2.

188 All scales are normally distributed (Shapiro-Wilks test), except for openness to experience ($W=0.96$,
189 $p < 0.05$), and satisfaction with life ($W=0.95$, $p < 0.05$), which are bimodal. Figure 1 Plot 2 shows the
190 correlations between different personality dimensions. As expected, the five personality dimensions are not
191 orthogonal.

Table 2. Demographics of the sample.

Variable	N (%)	Variable	Mean (SD)
<i>Gender</i>		<i>Age</i>	
- Female	49 (70%)	- Female	23.52 (6.56)
- Male	21 (30%)	- Male	22.84 (7.13)
<i>Ethnicity</i>		<i>Personality</i>	
- Caucasian	54 (75%)	- Openness to Experience	4.19 (0.46)
- Black	3 (4%)	- Conscientiousness	3.20 (0.75)
- Asian	5 (7%)	- Extraversion	3.11 (3.83)
- other	8 (14%)	- Agreeableness	3.55 (0.68)
		- Neuroticism	2.98 (0.89)
<i>Living Status</i>		<i>Mental Wellbeing</i>	
- Living with partner	8 (10%)	- SWL	4.18 (1.44)
- Single	54 (77%)	- CES-D	23.79 (11.86)
- Married	5 (7%)		
- Unknown	3 (4%)		

Note: Caucasian includes White people of American, British, and other origins; Black includes African Americans and Black people from Europe. SWL: score for Satisfaction with Life Scale. CES-D: Center for Epidemiologic Studies Depression Scale,

192 4.2 Social Media Affect: Frequencies versus Transitions

193 For **overall frequencies of affect category**, the only clear correlation is between extroversion and
 194 positive content. Overall, more extroverted participants are more likely to have days where they make
 195 predominantly positive posts ($r=0.29$, $p < 0.01$, $95\%CI = (-0.15, 0.32)$, $r^2 = 0.08$). In addition, participants
 196 who score higher on agreeableness tend to post fewer negative posts and have fewer days with predominantly
 197 negative posts (both $r=-0.26$, $p < 0.05$, $95\%CI = (-0.48, -0.04)$, $r^2 = 0.07$).

198 When we look at **transitions between affect categories**, however, a more nuanced picture emerges.
 199 Table 3 summarises the correlations between personality, well being and transition types. Significant
 200 correlations are summarised in Table 4. Due to the number of correlations presented, we choose a cut-off
 201 of $p < 0.01$, which is stricter than the normal $p < 0.05$.

202 Several transition types are correlated positively and negatively with Extroversion and Agreeableness.
 203 Neuroticism, conscientiousness, and SWL show interesting trends ($p < 0.05$) that do not reach significance
 204 (c.f. Table 3)

205 More extroverted participants are more likely to post predominantly positive content several days in a
 206 row (*day-level*, $++\leftrightarrow++$, $r=0.30$, $p < 0.001$, $95\% CI = (0.06, 0.54)$, $r^2=0.09$). They have more transitions to or
 207 from a silence day with a positive post (*post-level with-silence*, $S\leftrightarrow++$, $r=0.29$, $p < 0.01$, $95\% CI = (-0.01,$
 208 $0.46)$, $r^2=0.08$). This pattern fits well the overall predominance of posts with positive affect. Extroverts are
 209 also less likely to alternate between days with neutral and days with non-neutral content (*day-level*, for
 210 both $0\leftrightarrow++$ and $0\leftrightarrow-$, $r=-0.28$, $p < 0.01$, $95\% CI = (-0.52, -0.09)$, $r^2=0.08$).

211 People who score higher on agreeableness are less likely to follow a post with negative affect with
 212 another negative affect post ($-\leftrightarrow-$, *post-level with-silence*: $r=-0.37$, $p < 0.001$, $95\% CI = (-0.50, -0.06)$
 213 $, r^2=0.14$); This tendency is much less pronounced on the day-level ($-\leftrightarrow-$, $r=-0.22$, $p < 0.1$, $95\% CI =$
 214 $(-0.44, -0.02)$, $r^2=0.04$). On top of that, they are more likely to alternate between days with mixed valence

Table 3. Correlations between personality, SWL, and CES-D scores and affect transitions. Number of participants N=70

	Post-level representation (Post Plus Silence)														
	S↔S	-↔-	+↔+	±↔±	0↔0	+↔-	±↔+	±↔-	±↔0	0↔+	0↔-	S↔+	S↔-	±↔S	S↔0
<i>N_{Occ}</i>	1238	346	542	29	230	599	143	134	100	424	414	641	384	137	211
ope	0.09	-0.17	-0.17	-0.16	-0.05	-0.14	-0.07	-0.08	0.11	0.01	0.03	0.17	0.00	0.13	0.03
con	-0.06	0.01	0.09	-0.09	-0.15	0.11	0.00	-0.01	-0.14	-0.07	-0.08	0.16	0.00	0.15	-0.15
ext	0.04	-0.12	0.16	-0.10	-0.19	-0.06	-0.03	-0.12	-0.09	-0.09	-0.17	0.29**	-0.04	0.00	-0.18
agr	0.14	-0.37***	0.03	0.02	-0.15	-0.22*	0.08	0.04	0.04	-0.04	-0.23*	0.23*	-0.04	0.29**	-0.13
neu	-0.07	0.19	0.18	0.18	-0.03	0.23*	0.11	0.04	0.02	0.05	-0.05	-0.22*	-0.03	-0.23*	-0.13
swl	0.04	-0.10	-0.13	-0.10	0.06	-0.03	0.02	-0.05	-0.04	0.02	-0.08	0.02	0.16	-0.02	0.18
CESD	-0.04	0.19	0.08	0.09	0.00	0.04	0.15	0.07	0.03	-0.06	0.11	-0.20	0.00	-0.11	-0.03
Post-level representation (Post only), N = 70															
<i>N_{Occ}</i>	396	694	34	313	728	188	166	142	547	502					
ope	-0.16	-0.05	-0.06	-0.02	-0.05	0.06	-0.01	0.14	0.09	0.13					
con	-0.07	0.18	-0.07	-0.23*	0.08	0.14	0.10	-0.11	-0.13	-0.12					
ext	-0.04	0.33***	0.04	-0.24*	0.05	0.08	-0.10	-0.15	-0.16	-0.20					
agr	-0.28**	0.18	0.00	-0.16	-0.10	0.26*	0.28**	0.13	0.03	-0.26*					
neu	0.14	0.00	0.11	-0.02	0.16	-0.14	-0.09	-0.08	0.01	-0.12					
swl	0.00	-0.12	-0.11	0.11	0.02	0.09	0.09	-0.02	0.08	-0.04					
CESD	0.14	-0.04	0.03	0.04	-0.03	-0.06	-0.11	0.04	-0.11	0.13					
Day-level representation, N = 70															
<i>N_{Occ}</i>	228	281	271	267	304	287	303	296	298	261	311	242	259	261	261
ope	0.12	-0.17	-0.11	-0.05	-0.02	-0.08	0.00	-0.14	-0.07	-0.01	-0.02	0.12	-0.02	0.19	0.13
con	-0.06	-0.03	0.25*	0.05	-0.01	0.03	-0.03	-0.04	-0.16	-0.19	-0.12	0.08	0.10	0.06	-0.07
ext	0.06	-0.11	0.30***	-0.03	-0.14	0.04	0.14	-0.13	0.01	-0.28**	-0.28**	0.24*	-0.08	0.02	-0.17
agr	0.11	-0.22	0.15	-0.05	0.08	-0.12	0.16	-0.06	0.11	-0.08	-0.17	0.15	-0.07	0.28**	-0.09
neu	-0.08	0.16	0.00	0.19	-0.17	0.21*	0.09	0.11	-0.01	0.12	0.08	-0.14	-0.12	-0.26*	-0.03
swl	0.02	-0.08	-0.01	-0.08	0.25*	-0.03	-0.06	-0.10	0.03	-0.06	-0.04	-0.02	0.12	0.06	0.08
CESD	-0.03	0.11	-0.10	0.08	-0.18	0.02	0.10	0.08	0.08	-0.01	0.21	-0.18	0.03	-0.16	0.05

Note: Pearson correlation P-value (permutation testing): · < 0.1, * < .05, ** < .01, *** < .001, bidirectional transition types: ↔, negative valence: −, positive valence: +, mixed valence: ±, neutral: 0, silence day: S, *N_{Occ}*: number of occurrences of each transition type, ope: openness, con: conscientiousness, ext: extraversion, agr: agreeableness, neu: neuroticism, swl: Satisfaction with Life Scale, CESD: Center for Epidemiologic Studies Depression Scale

Table 4. Summary of the significant correlations between transition states and the five personality traits (*p* < 0.01)

	Transitions	Post-Level (with-silence)	Post-Level (without-silence)	Day-Level
Extraversion	S ↔ +	↑	—	—
	0 ↔ +	—	—	↓
	0 ↔ -	—	—	↓
	+ ↔ +	—	↑	↑
Agreeableness	- ↔ -	↓	↓	—
	± ↔ S	↑	—	↑
	± ↔ -	—	↑	—

Note: ↓ indicates a significant negative correlation at *p* < 0.01 or better, ↑ indicates a significant positive correlation at *p* < 0.01 or better. — indicates that the correlation is not significant at this level. Bidirectional transition types: ↔, negative valence: −, positive valence: +, mixed valence: ±, neutral: 0, silence day: S.

215 and silence (*day-level*, $\pm \leftrightarrow S$, $r=0.28$, $p < 0.01$, 95% CI = (-0.01, 0.46), $r^2=0.08$, *post-level with-silence*,
216 $\pm \leftrightarrow S$, $r=0.29$, $p < 0.01$, 95% CI = (0.08, 0.52), $r^2=0.08$).

217 Participants with higher neuroticism tend to alternate between positive and negative content, but this is
218 only evident when we take silence into account ($\pm \leftrightarrow -$, *post-level with-silence*: $r=0.23$, $p < 0.05$, 95% CI
219 = (0.00, 0.47), $r^2=0.04$, *post-level without-silence*: $r=0.16$, 95% CI = (-0.08, 0.41), $r^2=0.025$, *day-level*:
220 $r=0.21$, $p < 0.05$, 95% CI = (-0.46, -0.10), $r^2=0.04$).

221 There are interesting differences in transition patterns that incorporate information about silence days and
222 those that do not. When disregarding silence days, we observe that people with higher conscientiousness or
223 extroversion are slightly less likely to follow a neutral post with another neutral post (*post-level without-*
224 *silence*, conscientiousness, $0 \leftrightarrow 0$, $r = -0.23$, $p < 0.05$, 95% CI = (-0.41, -0.04), $r^2=0.07$; extroversion, $0 \leftrightarrow 0$,
225 $r = -0.24$, $p < 0.05$, 95% CI = (-0.41, -0.04), $r^2=0.07$).

226 When we take into account silence days for computing transitions, we find several more interesting
227 trends. People who are more satisfied with life are more likely to follow a neutral post with another neutral
228 post ($0 \leftrightarrow 0$, *day-level*: $r=0.25$, $p < 0.05$, 95% CI = (-0.01, 0.44), $r^2=0.06$). In addition, people with higher
229 neuroticism are more likely to alternate between positive and negative posts ($0 \leftrightarrow -$, *day-level*: $r=0.21$,
230 $p < 0.05$, 95% CI = (-0.01, 0.40), $r^2=0.04$), but less likely to make a positive post after a period of one or
231 more silence days ($S \leftrightarrow +$, *post-level with-silence*: $r=-0.22$, $p < 0.05$, 95% CI = (-0.48, 0.00), $r^2=0.04$). We
232 found that silence to silence transitions are not correlated with personality or mental health.

233 4.3 Post Originality

234 High CES-D scores are significantly correlated with posting non-original content ($r=0.29$, $p < 0.01$, 95%
235 CI = (0.10, 0.46), $r^2=0.08$). There is a similar tendency for participants with higher neuroticism scores
236 ($r=0.25$, $p < 0.05$, 95% CI = (0.06, 0.43), $r^2=0.07$). Examining transitions between post originality shows
237 that these effects stem from slightly different posting patterns. Users with higher CES-D scores tend to
238 follow non-original content with non-original content ($N \leftrightarrow N$, *post-level with-silence*, $r=0.26$, $p < 0.05$,
239 95% CI = (0.07, 0.43), $r^2=0.07$) or to alternate between original and non-original content ($N \leftrightarrow O$ *post-level*
240 *with-silence*, $r=0.27$, $p < 0.05$, 95% CI = (0.08, 0.44), $r^2=0.07$). Users with higher neuroticism scores
241 tend to post sequences of non-original content ($N \leftrightarrow N$, *post-level with-silence*, $r=0.25$, $p < 0.05$, 95% CI =
242 (0.06, 0.43), $r^2=0.05$), and are less likely to post original content before or after a period of silence ($O \leftrightarrow S$,
243 *post-level with-silence*, $r=0.28$, $p < 0.05$, 95% CI = (0.09, 0.45), $r^2=0.08$).

244 Since neuroticism is closely linked to depression symptoms, we also computed a partial correlation
245 between content originality and CES-D while controlling for neuroticism. The resulting correlation was no
246 longer significant ($r=0.14$, $p = 0.22$, $r^2=0.02$). Therefore, the association between content originality and
247 depression symptoms might be moderated by neuroticism.

5 DISCUSSION

248 5.1 Main Findings

249 Many studies have found associations between the frequency of affective words used in social media
250 text and personality. However, existing studies often see affect as static and only focused on the strength
251 of bipolar valence (positive/negative). Instead, our work focuses on affect patterns. We encode posting
252 behaviour, transitions between affect states, and content originality. From a practical point of view, our
253 technique can supplement experience sampling techniques (Myin-Germeys et al., 2018) to help clinicians

254 and patients develop a more comprehensive view of a person's affect patterns, arrive at a better substantiated
255 diagnosis, and make improved treatment decisions. However, this depends on whether the patient is willing
256 to share information from their social media feed with their therapist.

257 Overall, the correlations seen between affect transitions and personality traits are in line with the
258 consensus in the early literature (Gross et al., 1998). Extroverts tend to produce sequences of positive posts.
259 This behaviour fits well with the positive emotional core in extroverts stipulated in (Watson and Clark,
260 1997). Participants with higher agreeableness are less likely to post sequences of negative posts. This could
261 be due to their ability to regulate negative affect (Meier et al., 2006; Haas et al., 2007).

262 Although the psychology literature suggests a strong association between negative mood states and
263 neuroticism (Rusting and Larsen, 1995), we did not find this in our data. Our results are in line with
264 previous studies of verbal cues to personality traits in social media (Schwartz et al., 2013; Yarkoni, 2010;
265 Golbeck et al., 2011; Park et al., 2015). Golbeck et al. (2011) found social media users who were more likely
266 to talk about anxiety were on the higher end of the neuroticism scale. We speculate that self-presentation
267 bias may influence how social media users regulate their expression of negative emotions in their public
268 posts. The only relevant association we found was that social media users on the high end of neuroticism
269 are more likely to switch between posting positive and negative affective content. This finding aligns well
270 with the fact that high neuroticism is associated with high emotional instability (Costa and McCrae, 1992).

271 The link between posting non-original content and elevated depression symptoms appears to be moderated
272 by neuroticism. This suggests that high levels of neuroticism predispose users both to depressive symptoms,
273 and to an indirect disclosure of emotions through quotes and lyrics.

274 In our sample, the prevalence of depressive symptoms is higher than what would be expected in general
275 population. In the original CES-D paper, Radloff (1977) proposed three levels of depression severity: low
276 (0-15), mild-to-moderate (16-22), and high (23-60). They found that only 21% of the general population
277 scored above the low symptom level. In contrast, in our sample, nearly half of the participants exhibit a
278 high level of symptoms (>22). Within the context social media studies of depression, however, our data set
279 is not exceptional. For many studies in the area, high symptom individuals account for nearly half of the
280 data set (Reece et al., 2017; Tsugawa et al., 2015; Nadeem, 2016; De Choudhury et al., 2013; Orabi et al.,
281 2018).

282 Our results support the claim that affect expressed in social media data text is associated with social
283 media users' affect patterns in real life. However, the data set used in this study is from the early 2010's,
284 and only covers the well established social media platform Facebook. The associations found in this study
285 are likely to be slightly different from those found in another social network (e.g., Instagram) or in a new
286 data set collected ten years later.

287 **5.2 Limitations.**

288 Due to the restrictions imposed by the need for sufficient Facebook updates to allow analysis, our final
289 sample is relatively small. Given the size of the significant effects we found in the data, power calculations
290 indicate that a well-powered study should include data from around 200 users (Schönbrodt and Perugini,
291 2013). It also skews heavily towards younger female Caucasians with relatively low satisfaction with life
292 and strong depression symptoms. It is possible that other groups of users (e.g., non-Caucasians, males) are
293 less likely to disclose personal information about mood and emotions on their public Facebook (Dosono
294 et al., 2017; McDonald et al., 2019).

6 CONCLUSION AND FUTURE WORK

295 In this pilot study, we demonstrated the benefits of detailed representations of social media affect for
296 unpacking the relationship between personality, mental wellbeing, and the content posted on social media.
297 Importantly, our representations include non-binary affect categories (positive, negative, mixed, neutral),
298 and take into account content originality. As a consequence, we were able to obtain a more detailed picture
299 of the link between patterns of affect and depressive symptoms.

300 In future work, we plan to enrich our data set with more in-depth analyses of original versus non-original
301 content, extend coverage by including a larger sample of the myPersonality data set, and construct statistical
302 models that allow us to observe long-term trends in posting patterns. Future studies should also examine
303 the extent to which affect expressed in non-original content is aligned with the users' affect when they post
304 the material.

ACKNOWLEDGEMENTS

305 We thank Michael Kosinski and David Stillwell for permission to use myPersonality, and our six
306 undergraduate Research Assistants from the Psychology Department of the University of Edinburgh
307 for their hard annotation work. Walid Magdy's and Maria Wolters' work on this paper was partly funded
308 by The Alan Turing Institute (EPSRC, EP/N510129/1).

REFERENCES

- 309 Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. 2011. Sentiment
310 analysis of twitter data. In *Proceedings of the Workshop on Language in Social Media (LSM 2011)*
311 (Portland, Oregon). 30–38.
- 312 Yoram Bachrach, Michal Kosinski, Thore Graepel, Pushmeet Kohli, and David Stillwell. 2012. Personality
313 and patterns of Facebook usage. In *Proceedings of the 4th annual ACM web science conference* (New
314 York). ACM, 24–32.
- 315 Alexandra Balahur, Ralf Steinberger, Erik van der Goot, Bruno Pouliquen, and Mijail Kabadjov. 2009.
316 Opinion mining on newspaper quotations. In *Proceedings of the 2009 IEEE/WIC/ACM International*
317 *Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 03* (Milano, Italy). IEEE
318 Computer Society, 523–526.
- 319 Natalya N Bazarova, Yoon Hyung Choi, Victoria Schwanda Sosik, Dan Cosley, and Janis Whitlock. 2015.
320 Social sharing of emotions on Facebook: Channel differences, satisfaction, and replies. In *Proceedings*
321 *of the 18th ACM conference on computer supported cooperative work & social computing* (Vancouver,
322 BC, Canada). ACM, 154–164.
- 323 Lauren M Bylsma, Bethany H Morris, and Jonathan Rottenberg. 2008. A meta-analysis of emotional
324 reactivity in major depressive disorder. *Clinical psychology review* 28, 4 (2008), 676–691.
- 325 Lauren M Bylsma, April Taylor-Clift, and Jonathan Rottenberg. 2011. Emotional reactivity to daily events
326 in major and minor depression. *Journal of abnormal psychology* 120, 1 (2011), 155.
- 327 Gustavo Carlo, Maria Vicenta Mestre, Meredith M McGinley, Paula Samper, Ana Tur, and Deanna
328 Sandman. 2012. The interplay of emotional instability, empathy, and coping on prosocial and aggressive
329 behaviors. *Personality and Individual Differences* 53, 5 (2012), 675–680.
- 330 Lushi Chen, Christopher Hon Kwong Cheng, and Tao Gong. 2020. Inspecting Vulnerability to Depression
331 from Social Media Affect. *Frontiers in Psychiatry* 11 (2020), 54.

- 332 Lushi Chen, Tao Gong, Michal Kosinski, David Stillwell, and Robert L Davidson. 2017. Building a profile
333 of subjective well-being for social media users. *PloS one* 12, 11 (2017), e0187278.
- 334 Glen Coppersmith, Mark Dredze, and Craig Harman. 2014. Quantifying mental health signals in Twitter.
335 In *Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic*
336 *signal to clinical reality* (Denver, Colorado). 51–60.
- 337 Glen Coppersmith, Mark Dredze, Craig Harman, Kristy Hollingshead, and Margaret Mitchell. 2015.
338 CLPsych 2015 shared task: Depression and PTSD on Twitter. In *Proceedings of the 2nd Workshop on*
339 *Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. 31–39.
- 340 Paul T Costa and Robert R McCrae. 1992. *Neo Pi-R*. Psychological Assessment Resources Odessa, FL.
- 341 Munmun De Choudhury and Sushovan De. 2014. Mental health discourse on reddit: Self-disclosure, social
342 support, and anonymity. In *Eighth International AAAI Conference on Weblogs and Social Media* (Ann
343 Arbor, Michigan,). 21–30.
- 344 Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression
345 via social media. In *Seventh international AAAI conference on weblogs and social media* (Cambridge,
346 Massachusetts, USA). 170–185.
- 347 Ed Diener and Robert A Emmons. 1984. The independence of positive and negative affect. *Journal of*
348 *personality and social psychology* 47, 5 (1984), 1105.
- 349 Ed Diener, Eunkook M Suh, Richard E Lucas, and Heidi L Smith. 1999. Subjective well-being: Three
350 decades of progress. *Psychological bulletin* 125, 2 (1999), 276.
- 351 E. D. Diener, Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. 1985. The satisfaction with life
352 scale. *Journal of personality assessment* 49, 1 (1985), 71–75.
- 353 John M Digman. 1990. Personality structure: Emergence of the five-factor model. *Annual review of*
354 *psychology* 41, 1 (1990), 417–440.
- 355 Bryan Dosono, Yasmeen Rashidi, Taslima Akter, Bryan Semaan, and Apu Kapadia. 2017. Challenges
356 in Transitioning from Civil to Military Culture: Hyper-Selective Disclosure Through ICTs. *Proc.*
357 *ACM Hum.-Comput. Interact.* 1, CSCW (Dec. 2017), 41:1–41:23. [https://doi.org/10.1145/](https://doi.org/10.1145/3134676)
358 [3134676](https://doi.org/10.1145/3134676)
- 359 Golnoosh Farnadi, Susana Zoghbi, Marie-Francine Moens, and Martine De Cock. 2013. Recognising
360 personality traits using Facebook status updates. In *Seventh International AAAI Conference on Weblogs*
361 *and Social Media* (Cambridge, Massachusetts, USA). 154–164.
- 362 Barbara L Fredrickson. 1998. What good are positive emotions? *Review of general psychology* 2, 3 (1998),
363 300–319.
- 364 Nico H Frijda. 1993. Moods, emotion episodes, and emotions. *Handbook of emotions* 12, 2 (1993), 155.
- 365 Frank Fujita, Ed Diener, and Ed Sandvik. 1991. Gender differences in negative affect and well-being: the
366 case for emotional intensity. *Journal of personality and social psychology* 61, 3 (1991), 427.
- 367 Matthias Gamer, Jim Lemon, Ian Fellows, and Puspendra Singh. 2012. irr: Various Coefficients of
368 Interrater Reliability and Agreement. R package version 0.84. *Internet resource: http://CRAN.R-project.*
369 *org/package= irr](Verified April 10, 2013)* (2012).
- 370 Steven T Garren. 2017. Permutation Tests for Nonparametric Statistics Using R. *Asian Research Journal*
371 *of Mathematics* (2017), 1–8.
- 372 Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. 2011. Predicting personality
373 from twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011*
374 *IEEE third international conference on social computing*. IEEE, 149–156.

- 375 L R Goldberg, J A Johnson, H W Eber, R Hogan, M C Ashton, C R Cloninger, and H G Gough. 2006. The
376 international personality item pool and the future of public-domain personality measures. *Journal of*
377 *Research in Personality* 40 (2006), 84–96.
- 378 James J Gross, Steven K Sutton, and Timothy Ketelaar. 1998. Relations between affect and personality:
379 Support for the affect-level and affective-reactivity views. *Personality and social psychology bulletin* 24,
380 3 (1998), 279–288.
- 381 Brian W Haas, Kazufumi Omura, R Todd Constable, and Turhan Canli. 2007. Is automatic emotion
382 regulation associated with agreeableness? A perspective using a social neuroscience approach.
383 *Psychological Science* 18, 2 (2007), 130–132.
- 384 Bruce Headey, Jonathan Kelley, and Alex Wearing. 1993. Dimensions of mental health: Life satisfaction,
385 positive affect, anxiety and depression. *Social indicators research* 29, 1 (1993), 63–82.
- 386 J.J. Higgins. 2003. *Introduction to Modern Non-Parametric Statistics*.
- 387 Marlies Houben, Wim Van Den Noortgate, and Peter Kuppens. 2015. The relation between short-term
388 emotion dynamics and psychological well-being: A meta-analysis. *Psychological bulletin* 141, 4 (2015),
389 901.
- 390 Patrik N Juslin and Petri Laukka. 2004. Expression, perception, and induction of musical emotions: A
391 review and a questionnaire study of everyday listening. *Journal of new music research* 33, 3 (2004),
392 217–238.
- 393 Junjie Lin, Wenji Mao, and Daniel D Zeng. 2017. Personality-based refinement for sentiment classification
394 in microblog. *Knowledge-Based Systems* 132 (2017), 204–214.
- 395 Robert R. McCrae and Juri Allik (Eds.). 2002. *The five-factor model of personality across cultures*.
396 Springer Science and Business Media.
- 397 Robert R. McCrae and Oliver P. John. 1992. An introduction to the five-factor model and its applications.
398 *Journal of personality* 60, 2 (1992), 175–215.
- 399 James McDonald, Kate Lockwood Harris, and Jessica Ramirez. 2019. Revealing and Concealing Difference:
400 A Critical Approach to Disclosure and an Intersectional Theory of “Closeting”. *Communication Theory*
401 (2019). <https://doi.org/10.1093/ct/qtz017>
- 402 Brian P Meier, Michael D Robinson, and Benjamin M Wilkowski. 2006. Turning the other cheek:
403 Agreeableness and the regulation of aggression-related primes. *Psychological Science* 17, 2 (2006),
404 136–142.
- 405 Ivan Mervielde, Veerle Buyst, and Filip De Fruyt. 1995. The validity of the Big-Five as a model for
406 teachers’ ratings of individual differences among children aged 4–12 years. *Personality and Individual*
407 *Differences* 18, 4 (1995), 525–534.
- 408 Saif Mohammad. 2016. A practical guide to sentiment annotation: Challenges and solutions. In *Proceedings*
409 *of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*
410 (San Diego, California). 174–179.
- 411 Karo Moilanen and Stephen Pulman. 2007. Sentiment composition. In *Proceedings of the Recent Advances*
412 *in Natural Language Processing International Conference* (Borovets, Bulgaria). 378–382.
- 413 Inez Myin-Germeys, Zuzana Kasanova, Thomas Vaessen, Hugo Vachon, Olivia Kirtley, Wolfgang
414 Viechtbauer, and Ulrich Reininghaus. 2018. Experience sampling methodology in mental health research:
415 new insights and technical developments. *World Psychiatry* 17, 2 (2018), 123–132. <https://doi.org/10.1002/wps.20513> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/wps.20513>
- 416
417 Moin Nadeem. 2016. Identifying depression on Twitter. *arXiv preprint arXiv:1607.07384* (2016).

- 418 Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. 2018. Deep
419 learning for depression detection of twitter users. In *Proceedings of the Fifth Workshop on Computational*
420 *Linguistics and Clinical Psychology: From Keyboard to Clinic* (New Orleans, LA). 88–97.
- 421 John G. Orme, Janet Reis, and Elicia J. Herz. 1986. Factorial and discriminant validity of the center for
422 epidemiological studies depression (CES-D) scale. *Journal of clinical psychology* 42, 1 (1986), 28–33.
- 423 Gregory Park, H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Michal Kosinski, David J
424 Stillwell, Lyle H Ungar, and Martin EP Seligman. 2015. Automatic personality assessment through
425 social media language. *Journal of personality and social psychology* 108, 6 (2015), 934.
- 426 Minsu Park, Chiyong Cha, and Meeyoung Cha. 2012. Depressive moods of users portrayed in Twitter.
427 In *Proceedings of the ACM SIGKDD Workshop on healthcare informatics (HI-KDD)*, Vol. 2012. ACM
428 New York, NY, 1–8.
- 429 William Pavot and Ed Diener. 2009. Review of the satisfaction with life scale. *Psychological assessment* 5,
430 2 (2009), 164.
- 431 James W Pennebaker and Laura A King. 1999. Linguistic styles: Language use as an individual difference.
432 *Journal of personality and social psychology* 77, 6 (1999), 1296.
- 433 Nooshin Pishva, Maryam Ghalehban, Afsane Moradi, and Leila Hoseini. 2011. Personality and happiness.
434 *Procedia-Social and Behavioral Sciences* 30 (2011), 429–432.
- 435 Lenore Sawyer Radloff. 1977. The CES-D scale: a self-report depression scale for research in the general
436 population. *Applied psychological measurement* 1, 3 (1977), 385–401.
- 437 Andrew G Reece, Andrew J Reagan, Katharina LM Lix, Peter Sheridan Dodds, Christopher M Danforth,
438 and Ellen J Langer. 2017. Forecasting the onset and course of mental illness with Twitter data. *Scientific*
439 *reports* 7, 1 (2017), 13006.
- 440 Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen, and Jordan
441 Boyd-Graber. 2015. Beyond LDA: Exploring supervised topic modeling for depression-related language
442 in Twitter. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology:*
443 *From Linguistic Signal to Clinical Reality* (Denver, Colorado). 99–107.
- 444 Robert E. Roberts. 1980. Reliability of the CES-D scale in different ethnic contexts. *Psychiatry research* 2,
445 2 (1980), 125–134.
- 446 Sara Rosenthal, Preslav Nakov, Svetlana Kiritchenko, Saif Mohammad, Alan Ritter, and Veselin Stoyanov.
447 2015. Semeval-2015 task 10: Sentiment analysis in twitter. In *Proceedings of the 9th international*
448 *workshop on semantic evaluation (SemEval 2015)* (Denver, Colorado). 451–463.
- 449 Mary K Rothbart, Stephan A Ahadi, and David E Evans. 2000. Temperament and personality: origins and
450 outcomes. *Journal of personality and social psychology* 78, 1 (2000), 122.
- 451 Jonathan Rottenberg. 2005. Mood and emotion in major depression. *Current Directions in Psychological*
452 *Science* 14, 3 (2005), 167–170.
- 453 Jonathan Rottenberg and James J Gross. 2003. When emotion goes wrong: Realizing the promise of
454 affective science. *Clinical Psychology: Science and Practice* 10, 2 (2003), 227–232.
- 455 James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6
456 (1980), 1161.
- 457 James A Russell. 2003. Core affect and the psychological construction of emotion. *Psychological review*
458 110, 1 (2003), 145.
- 459 Cheryl L Rusting. 1998. Personality, mood, and cognitive processing of emotional information: three
460 conceptual frameworks. *Psychological bulletin* 124, 2 (1998), 165.
- 461 Cheryl L Rusting and Randy J Larsen. 1995. Moods as sources of stimulation: Relationships between
462 personality and desired mood states. *Personality and individual differences* 18, 3 (1995), 321–329.

- 463 Richard M Ryan and Edward L Deci. 2001. On happiness and human potentials: A review of research on
464 hedonic and eudaimonic well-being. *Annual review of psychology* 52, 1 (2001), 141–166.
- 465 Klaus R Scherer. 2004. Which emotions can be induced by music? What are the underlying mechanisms?
466 And how can we measure them? *Journal of new music research* 33, 3 (2004), 239–251.
- 467 Klaus R Scherer, Marcel R Zentner, et al. 2001. Emotional effects of music: Production rules. *Music and*
468 *emotion: Theory and research* 361, 2001 (2001), 392.
- 469 Ulrich Schimmack, Shigehiro Oishi, and Ed Diener. 2002. Cultural influences on the relation between
470 pleasant emotions and unpleasant emotions: Asian dialectic philosophies or individualism-collectivism?
471 *Cognition & Emotion* 16, 6 (2002), 705–719.
- 472 Felix D Schönbrodt and Marco Perugini. 2013. At what sample size do correlations stabilize? *Journal of*
473 *Research in Personality* 47, 5 (2013), 609–612.
- 474 H Andrew Schwartz, Johannes C Eichstaedt, Margaret L Kern, Lukasz Dziurzynski, Stephanie M Ramones,
475 Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin EP Seligman, et al. 2013.
476 Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one* 8,
477 9 (2013), 234.
- 478 Shalom H Schwartz. 1992. Universals in the content and structure of values: Theoretical advances and
479 empirical tests in 20 countries. In *Advances in experimental social psychology*. Vol. 25. Elsevier, 1–65.
- 480 Gal Sheppes, Gaurav Suri, and James J Gross. 2015. Emotion regulation and psychopathology. *Annual*
481 *review of clinical psychology* 11 (2015), 379–405.
- 482 David H Silvera, Anne M Lavack, and Fredric Kropp. 2008. Impulse buying: the role of affect, social
483 influence, and subjective wellbeing. *Journal of Consumer Marketing* 25, 1 (2008), 23–33.
- 484 Kamlesh Singh and Shalini Duggal Jha. 2008. Positive and negative affect, and grit as predictors of
485 happiness and life satisfaction. *Journal of the Indian Academy of Applied Psychology* 34, 2 (2008),
486 40–45. <https://doi.org/10.1038/s41746-020-0233-7>
- 487 Simone Teufel. 1999. *Argumentative Zoning: Information Extraction from Scientific Articles*. Ph.D.
488 Dissertation. Centre for Cognitive Science, University of Edinburgh.
- 489 Renee J Thompson, Jutta Mata, Susanne M Jaeggi, Martin Buschkuhl, John Jonides, and Ian H Gotlib.
490 2012. The everyday emotional experience of adults with major depressive disorder: Examining emotional
491 instability, inertia, and reactivity. *Journal of abnormal psychology* 121, 4 (2012), 819.
- 492 Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, and Hiroyuki Ohsaki. 2015.
493 Recognizing depression from twitter activity. In *Proceedings of the 33rd annual ACM conference on*
494 *human factors in computing systems*. ACM, 3187–3196.
- 495 David Watson. 2000. *Mood and temperament*. Guilford Press.
- 496 David Watson and Lee Anna Clark. 1997. Extraversion and its positive emotional core. In *Handbook of*
497 *personality psychology*. Elsevier, 767–793.
- 498 David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures
499 of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6
500 (1988), 1063.
- 501 Tal Yarkoni. 2010. Personality in 100,000 words: A large-scale analysis of personality and word use among
502 bloggers. *Journal of research in personality* 44, 3 (2010), 363–373.
- 503 Wu Youyou, Michal Kosinski, and David Stillwell. 2015. Computer-based personality judgments are more
504 accurate than those made by humans. *Proceedings of the National Academy of Sciences* 112, 4 (2015),
505 1036–1040.