



Living in the Past, Present, and Future: Measuring Temporal Orientation with Language

Journal:	<i>Journal of Personality</i>
Manuscript ID	JOPY-15-0145.R1
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Park, Gregory; University of Pennsylvania, Psychology Schwartz, H.; University of Pennsylvania, Computer & Information Sciences; Stony Brook University, Computer Science Kern, Margaret; University of Melbourne, Graduate School of Education Sap, Maarten; University of Pennsylvania, Computer & Information Sciences Weingarten, Evan; University of Pennsylvania, The Wharton School Eichstaedt, Johannes; University of Pennsylvania, Department of Psychology Berger, Jonah; University of Pennsylvania, The Wharton School, Marketing; Stillwell, David; University of Cambridge, Department of Psychology Kosinski, Michal; University of Cambridge, Department of Psychology Ungar, Lyle; University of Pennsylvania, Computer & Information Sciences; University of Pennsylvania, Psychology Seligman, Martin; University of Pennsylvania, Department of Psychology</p>
Keywords:	temporal orientation, language, computational social science, social media, big data

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13

**Living in the Past, Present, and Future:
Measuring Temporal Orientation with Language**

Gregory Park¹, H. Andrew Schwartz^{1,2}, Maarten Sap¹, Margaret L. Kern³,

Evan Weingarten¹, Johannes C. Eichstaedt¹, Jonah Berger¹, David J. Stillwell⁴,

Michal Kosinski⁵, Lyle H. Ungar¹, Martin E. P. Seligman¹

¹ University of Pennsylvania, ² Stony Brook University, ³ The University of Melbourne,

⁴ University of Cambridge, ⁵ Stanford University

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42

Author Note

Gregory Park, Department of Psychology, University of Pennsylvania; H. Andrew Schwartz, Computer & Information Science, University of Pennsylvania and Computer Science, Stony Brook University; Margaret L. Kern, Melbourne Graduate School of Education, The University of Melbourne, Australia; Maarten Sap, Department of Psychology, University of Pennsylvania; Evan Weingarten, The Wharton School, University of Pennsylvania; Johannes C. Eichstaedt, Department of Psychology, University of Pennsylvania; David J. Stillwell, Psychometrics Centre, University of Cambridge, UK; Michal Kosinski, Graduate School of Business, Stanford University; Lyle H. Ungar, Computer & Information Science, University of Pennsylvania; Martin E. P. Seligman, Department of Psychology, University of Pennsylvania

43
44
45
46
47
48
49

Support for this article was provided by grant #63597 from the Robert Wood Johnson Foundation (M. E. P. Seligman, PI) and by a grant from the Templeton Religion Trust (M.E.P. Seligman, H. A. Schwartz, L. H. Ungar, co-PIs).

50
51
52
53
54
55
56
57
58
59
60

Correspondence concerning this article should be addressed to Gregory Park, Department of Psychology, University of Pennsylvania, 3701 Market Street, 2nd floor, Philadelphia, PA 19104. Email: gregpark@sas.upenn.edu

Abstract

Objective: Temporal orientation refers to individual differences in the relative emphasis one places on the past, present, or future, and is related to academic, financial, and health outcomes.

We propose and evaluate a method for automatically measuring temporal orientation through language expressed on social media.

Method: Judges rated the temporal orientation of 4,302 social media messages. We trained a classifier based on these ratings, which could accurately predict the temporal orientation of new messages in a separate validation set (accuracy/mean sensitivity = .72; mean specificity = .77).

We used the classifier to automatically classify 1.3 million messages written by 5,372 participants (50% female, aged 13-48). Finally, we tested whether individual differences in past, present, and future orientation differentially related to gender, age, Big Five personality, satisfaction with life, and depressive symptoms.

Results: Temporal orientations exhibit several expected correlations with age, gender, and Big Five personality. More future-oriented people were older, more likely to be female, more conscientious, less impulsive, less depressed, and more satisfied with life; present orientation showed the opposite pattern.

Conclusion: Language-based assessments can complement and extend existing measures of temporal orientation, providing an alternative approach and additional insights into language and personality relationships.

Keywords: temporal orientation, language, computational social science, social media, big data

Living in the Past, Present, or Future: Measuring Temporal Orientation with Language

Consider three pairs of emotions: (a) regret and nostalgia, (b) boredom and joy, and (c) dread and hope. In each pair, emotions are opposed in valence but similar in orientation towards the past (a), present (b), or future (c). Psychological research has mostly concentrated on understanding people's tendencies to express positive or negative emotions, but less attention has been given to their relative focus on the past, present, or future. One reason may be that these *temporal orientations* are hard to measure with traditional self-report methods. We introduce a method for automatically assessing temporal orientation through language expressed in social media. In addition, we explore differences across age and gender, and connections to personality, subjective well-being, and depressive symptoms.

Studies on Temporal Orientation

Most studies of temporal orientation have focused on *future-oriented* thinking and its relation to educational, health, and financial outcomes. For example, students with higher future orientation study longer and earn better grades (Horstmanshof & Zimitat, 2007; Zimbardo & Boyd, 1999), and more future-oriented adults use less alcohol and tobacco (Adams & Nettle, 2009; Daughterty & Brase, 2010; Keough, Zimbardo, & Boyd, 1999), practice safer sex (Rothspan & Read, 1996), exercise more frequently (Ouellette, Hessling, Gibbons, Reis-Bergan, & Gerrard, 2005), hold more positive attitudes towards exercise (Joireman, Shaffer, Balliet, & Strathman, 2012), control diets better (Piko & Brassai, 2009), have lower body mass indexes, (Adams & Nettle, 2009; Adams & White, 2009), save more of their income (Webley & Nyhus, 2006), and plan their finances further into the future (Adams & Nettle, 2009).

Present and future orientations also have well-established age differences. As people grow older, they report thinking less about the present and more about the future (Casey, Jones,

1
2
3 & Hare, 2008; Nurmi, 2005; Steinberg et al., 2009). Early childhood is characterized by a
4
5 preoccupation with the immediate present, whereas weighing the consequences of today's
6
7 decisions is a hallmark of maturity. According to questionnaire measures, future-oriented
8
9 thinking begins in early adolescence, becomes more common throughout adolescence, and levels
10
11 off in young adulthood (Steinberg et al., 2009).
12
13
14

15 Studies have also found smaller but consistent gender differences in temporal orientation.
16
17 Across eight samples, Keough et al. (1999) found women were more future-oriented and men
18
19 were more present-oriented. Steinberg et al. (2009) reported that women scored significantly
20
21 higher than men on three measures of future orientation.
22
23

24 **Measuring Temporal Orientation**

25
26
27 Temporal orientation is typically measured by self-reports, such as the Zimbardo Time
28
29 Preference Inventory (ZPTI; Zimbardo & Boyd, 1999) and the Consideration of Future
30
31 Consequences scale (CFC; Joireman et al., 2012; Strathman, Gleicher, Boninger, & Edwards,
32
33 1994). Respondents rate statements about their thinking or planning style, and these items form
34
35 subscales measuring past ("It gives me great pleasure to think about my past"; ZPTI), present ("I
36
37 often follow my heart more than my head"; ZPTI), and future orientations ("When I make a
38
39 decision, I think about how it might affect me in the future"; CFC). These measures are easy to
40
41 administer and predict several outcomes, as noted above.
42
43
44

45
46 However, these self-reported items highly overlap with self-reported measures of
47
48 personality traits. For example, future orientation is strongly correlated with *conscientiousness*
49
50 (*rs* range from .50 to .60; Strathman et al., 1994; Zhang & Howell, 2011; Zimbardo & Boyd,
51
52 1999). It may be that conscientiousness predisposes a person to be more future-oriented, but such
53
54 distinctions are complicated by the fact that questionnaire measures of conscientiousness and
55
56
57
58
59
60

1
2
3 future orientation are also very similar. For example, the ZPTI Future scale includes the item “I
4 make lists of things to do”, while conscientiousness scales include items such as “I do things
5 according to a plan” (Goldberg et al., 2006). A behavior-based measure of temporal orientation
6 could provide researchers with an alternative method that has less overlap with measures of
7 similar constructs.
8
9

10
11
12
13
14
15 Likewise, self-reports often have an implicit evaluative component, such as the ZPTI’s
16 *Past-Negative* (e.g., “Painful past experiences keep being replayed in my mind) and *Past-*
17 *Positive* (e.g., “It gives me pleasure to think about my past”) subscales. These two subscales
18 correlate with measures related to subjective well-being (neuroticism, depression, and self-
19 esteem; Zimbardo & Boyd, 1999). The evaluative aspect—the tendencies to rate experiences and
20 memories as positive or negative—may be driving these correlations, rather than a true
21 association with temporal orientation. If so, these measures cannot assess the unique contribution
22 of temporal orientation on well-being.
23
24
25
26
27
28
29
30
31
32
33

34
35 These measurement confounds prevent researchers from clearly separating temporal
36 orientation from other related traits. One solution lies in behavior-based measures (Roberts,
37 Harms, Smith, Wood, & Webb, 2006). Behavior-based measures remove the shared method
38 variance with self-reports (i.e., overlapping, similar items), reduce the influence of a
39 respondent’s evaluative style, and enable multi-method designs. Language use provides one
40 psychologically rich and practical source of behavioral data (Kern et al., 2014; Pennebaker,
41 Mehl, & Niederhoffer, 2003). When combined with techniques from natural language
42 processing, statistical models can accurately predict several individual characteristics—age,
43 gender, and personality—from language alone (Park et al., 2015; Schwartz et al., 2013b).
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 In the current study, we created a new language-based measure of temporal orientation.
4
5 First, we developed a model to classify text as oriented towards the past, present, or future. We
6
7 used this model to classify millions of Facebook *status updates* (i.e., short text messages used to
8
9 describe someone's current mood, thoughts, activities, or plans), creating a person-level measure
10
11 of past, present, and future orientation. We then compared orientations to age, gender, and
12
13 personality—checking for consistency with patterns found using self-reports— and then
14
15 extended these comparison to life satisfaction and depression.
16
17
18
19

20 **Part 1: Message-level Temporal Classification Model**

21
22 We developed a classification model on one set of language data, with the goal of
23
24 automatically classifying a second set of data as past-, present-, or future-oriented on the basis of
25
26 several linguistic features (Schwartz et al., 2015). This process required that we (1) obtain a set
27
28 of text samples for training; (2) annotate these text samples as past-, present-, or future-oriented;
29
30 (3) extract linguistic features (e.g., words, phrases, number of words) from each text sample; (4)
31
32 train a statistical model to predict the text's temporal annotation based on its linguistic features;
33
34 and (5) evaluate the accuracy of this model on a new set of messages.
35
36
37
38

39 **Training Messages**

40
41 For our initial set of text samples, we used 6,000 messages from Twitter and Facebook.
42
43 From Twitter (a microblogging platform on which users can post short text messages, or
44
45 “tweets”, limited to 140 characters), we sampled 3,000 messages, drawn from a random feed
46
47 provided by Twitter during September 2012. From Facebook, we sampled 3,000 status updates,
48
49 drawn from users of the MyPersonality application (Kosinski, Stillwell, & Graepel, 2013)
50
51 between January 2009 and October 2011. MyPersonality is a third-party application through
52
53 which users can complete personality and other psychological measures and share results with
54
55
56
57
58
59
60

1
2
3 friends. Users voluntarily allowed the application to access all of their Facebook status updates
4
5 for research purposes. Of the 6,000 training messages, 1,489 were identified as song lyrics,
6
7 famous quotations, or posts by bot (i.e., automated) accounts, and these were removed from the
8
9 training sample.
10
11

12 **Message Annotation**

13
14
15 Three independent judges rated the temporal orientation of each of the remaining 4,511
16
17 messages, using fractions of the day in the past or future. For example, a message referring to the
18
19 immediate present was rated as 0, an hour in the future was +1/24, 1 day in the future was +1,
20
21 one week in the future was +7, and one day in the past was -1. Judges were instructed to mark
22
23 non-interpretable messages as 'NA'. We removed messages that were rated 'NA' by all three
24
25 raters, which excluded an additional 209 messages (125 Twitter messages and 84 Facebook
26
27 messages). Inter-rater agreement for the remaining 4,302 messages was high (intraclass
28
29 correlation coefficient, ICC = .85).
30
31
32
33

34 We used the mean rating to classify each message into three categories: past-oriented
35
36 (mean rating < 0), present-oriented (mean rating = 0), or future-oriented (mean rating > 0). Table
37
38 1 lists examples of messages, individual ratings, and final orientation classification. Of the 4,302
39
40 messages, 1,178 (27.4%) were classified as past-oriented, 2,043 (47.5%) as present-oriented, and
41
42 1,081 (25.1%) as future-oriented. Of the 2,293 Facebook messages, 659 (28.7%) were classified
43
44 as past-oriented, 990 (43.2%) as present-oriented, and 644 (28.1%) as future-oriented. Of the
45
46 2,009 Twitter messages, 519 (25.8%) were classified as past-oriented, 1,053 (52.4%) as present-
47
48 oriented, and 437 (21.8%) as future-oriented.
49
50
51
52

53 **Linguistic Feature Extraction**

We extracted five types of linguistic features from each message: words and phrases, time expressions, parts of speech, word categories, and length of message.

Words and phrases. We used an emoticon-aware tokenizer (*happierfuntokenizing*; Potts; 2011) to divide messages into smaller word-like units, or *tokens*. The tokenizer was sensitive to single words, punctuation, non-conventional usages and spellings (e.g., *omg*, *lol*) and emoticons (e.g., *:-)*), which are common on social media. We represented a message's constituent words, phrases, and similar features using a binary encoding. That is, for each message, if a given word or phrase appeared at least once, it was coded as 1, otherwise it was coded as 0.

Time expressions. We used the Stanford SUTime annotator (Chang & Manning, 2012) to identify time expressions (e.g., "yesterday", "next September") within each message. Once identified, time expressions were used to derive six features: the mean temporal difference (in days) between all time expressions in the message and the time of the message's creation, the log (base 2) of this difference, the absolute value of the difference, and three binary variables encoding whether any time expressions in the messages referred to the past, present, or future. We also added a feature coding that indicated the total number of time expressions that occurred in the message.

Parts of speech. We used Stanford's part-of-speech tagger (Toutanova, Klein, Manning, & Singer, 2003) to identify each token's corresponding part of speech. For each possible part of speech tag, we calculated the frequency of the tag within each message and divided the frequency by the total number of tokens in each message.

Word categories. We used the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007) dictionaries to count the frequency of words in 64 pre-defined categories, including temporally-oriented categories such as *future* words (e.g.,

1
2
3 “will”, “gonna”, “might”). The frequency of words within each LIWC category was divided by
4
5 the total number of tokens in the message, resulting in 64 separate features.
6
7

8 **Message length.** Two features captured message length: the mean length (i.e., number of
9 characters) of all tokens in the message, and the total number of tokens in the message.
10
11

12 **Temporal Classification Model**

13
14
15 After extracting linguistic features from each message, we fit a statistical model over the
16 set of training messages to predict their rated temporal orientation from the features. Because this
17 task requires classification into three categories (past, present, and future), we explored four
18 classification techniques, implemented in the *scikit-learn* Python module (Pedregosa et al.,
19 2011): logistic regression (LR) with Lasso regularization, support vector classification with a
20 linear kernel (LSVC), support vector classification with a radial basis kernel (rSVC), and a forest
21 of extremely randomized trees (ERT).
22
23
24
25
26
27
28
29
30

31
32 ERT fits many (hundreds or more) single trees to random portions of the training data,
33 and then combines the individual predictions to form a more stable ensemble prediction.
34
35 Traditional decision tree models naturally handle non-linear relationships and interactions
36 between predictors, but single trees are unstable and prone to overfitting (Berk, 2008). In our
37 case, each decision tree was fit to a random subset of messages from the training data and a
38 random subset of features. Splits at each node in the decision tree were also randomly chosen.
39
40 We used the following ERT parameters: we built 1,000 trees, chose node splits using the Gini
41 impurity measure, and used the square-root of the total number of features as the amount of
42 randomly selected features when building each tree. To classify a new message’s temporal
43 orientation, we applied the 1,000 fitted trees to the new message (i.e., its corresponding features)
44 and used the most frequent class as the predicted class.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
Model evaluation. We evaluated the performance of all four techniques by applying it to a new independent set of messages. We randomly sampled 500 Facebook status updates from the MyPersonality data set (not included in the training set), and three independent judges rated the message orientation as either past, present, or future.¹ Agreement between raters was high (ICC = 0.83). We used the majority rating as each message's temporal orientation. The resulting orientations of the messages were 131 (26.2%) past-oriented, 250 (50.0%) present-oriented, and 105 (21.0%) future-oriented. Fourteen messages were three-way ties (one past, one present, one future), and these messages were coded as present (the most frequent class). We then applied each classification technique to these messages, comparing the agreements between model prediction and human ratings.

27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
As benchmarks, a random classifier would have an accuracy of 0.33, and predicting the most frequent class (present) would yield an accuracy of 0.53. The resulting accuracies of the four techniques were logR = .69, ISVC = .71, rSVC = .68, and ERT = .72. We concluded that the ERT model was best for automatically classifying new messages.² Mean specificity of the ERT model, or how often a message was correctly *not* classified as an incorrect class, was 0.77. Of the 131 messages that were truly past (based on human judgments), 79 were predicted as past, 42 as present, and 10 as future. Of the 264 messages that were truly present, 15 were predicted as past, 232 as present, and 17 as future. Of the 105 messages that were truly future, 9 were predicted as past, 46 as present, and 50 as future.

48
49
50
51
52
53
54
55
56
57
58
59
60
To evaluate the relative importance of each feature type in the ERT model, we examined how model performance changed across different combinations of features. We started by using only one feature type to classify messages, resulting in the following accuracies: only message lengths (.54), only time expressions (.59), only parts of speech (.61), only word categories (.68),

1
2
3 and only words and phrases (.69). We then tested the model performance using all *except* one
4
5 feature type, resulting in the following accuracies: all except words and phrases (.67), all except
6
7 word categories (.70), all except time expressions (.71), all except parts of speech (.71), and all
8
9 except message lengths (.72). We concluded that all feature types but message lengths add useful
10
11 information and improve performance. However, the inclusion of message length features does
12
13 not reduce model performance, so we used all five feature types in the final model.
14
15

16 17 18 **Part 2: Assessment of Person-level Temporal Orientation**

19
20 After developing an accurate model, we then applied the model to a much larger set of
21
22 messages from Facebook users, and compared their aggregated temporal patterns to several self-
23
24 reported individual characteristics.
25

26 27 **Participants**

28
29 Participants were drawn from a pool of 72,559 users of the MyPersonality Facebook
30
31 application who were not a part of the training set, who also granted access to all status
32
33 messages, written between June 2009 and November 2011. This pool of users was 62% female
34
35 with an average age of 23.3 years old ($SD = 8.9$; median = 20). For practical purposes, we
36
37 sampled a smaller subset of users, rather than use the full pool. The pool of users wrote over 20
38
39 million messages, and extracting linguistic features, particularly the syntactic parsing needed to
40
41 extract time-expressions from all of these messages is a very time-intensive process. We
42
43 reasoned that a smaller sample of participants and messages (i.e., about 5,000 participants with
44
45 roughly one million messages) would still yield stable estimates but also allow a much shorter
46
47 development cycle (i.e., days instead of weeks).
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The full MyPersonality sample had a high concentration of users between the ages of 18-
4
5 22 (36% of users) and more women than men (61% of users). To ensure the subsample included
6
7 adults from a large age span, we stratified our sample across age and gender, which resulted in a
8
9 much more balanced sample. We also wanted to ensure that the participants in our sample had
10
11 completed other relevant psychological measures. To satisfy these requirements, we sampled two
12
13 subsets of participants.³
14
15

16
17 Subset 1 was an age- and gender-balanced sample, which was created by randomly
18
19 sampling 180 participants (90 men, 90 women) from two-year age bins ranging from 13 to 48
20
21 ([13, 14], [15, 16] ... [47, 48]), resulting in a sample of 3,240 participants. All participants in this
22
23 stratified sample reported their age, gender, completed a self-report measure of Big Five
24
25 personality factors (detailed below), and wrote at least 100 status updates. The mean and median
26
27 age of the resulting subsample were 30.5.
28
29

30
31 Subset 2 included 2,132 participants who reported age, gender, wrote at least 100 status
32
33 updates, and completed at least one measure of impulsivity, life satisfaction, or depressive
34
35 symptoms. The subset included 754 men and 1,378 women, and had a mean age of 21.7 ($SD =$
36
37 7.6, median = 19.0).
38
39

40 41 **Measures**

42
43 **Big Five Personality.** All participants from subset 1 completed items assessing Big Five
44
45 personality (openness to experience, conscientiousness, extraversion, agreeableness, and
46
47 neuroticism) from the International Personality Item Pool (IPIP; Goldberg et al., 2006). All
48
49 participants completed at least the 20-item version of this measure. Participants could optionally
50
51 complete additional IPIP items; 636 participants completed the full 100-item version of measure.
52
53
54
55
56
57
58
59
60

1
2
3 **Barratt Impulsiveness Scale.** From subset two, 762 participants completed the Barratt
4 Impulsiveness Scale (BIS; Stanford et al., 2009), a 30-item assessment of general impulsiveness.
5 Each BIS item states a manner of acting or thinking (e.g., “I do things without thinking”, “I buy
6 things on impulse”), and participants indicate how accurately each statement describes
7 themselves on a 4-point scale (1 = rarely/never; 4 = almost always/always). For 76 participants
8 who were missing responses for a single item, we imputed the single missing value with the
9 mean of the remaining items. We excluded 18 participants who were missing scores on more
10 than one item, leaving 744 participants with BIS scores. We calculated the full-scale score as the
11 mean across all 30 items (Cronbach’s $\alpha = .83$).
12
13
14
15
16
17
18
19
20
21
22
23

24 **Satisfaction with Life.** From subset two, 1,369 participants completed the Satisfaction
25 with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985), a five-item assessment of
26 life satisfaction. Participants indicate their agreement with five statements (e.g., “I am satisfied
27 with my life”, “The conditions of my life are excellent”) on a 7-point scale (1 = strongly
28 disagree; 7 = strongly agree). There were no missing responses across the participants who met
29 the inclusion criteria for subset 2. For 79 participants that completed the SWLS more than once,
30 we only used data from the first administration. We calculated the full-scale score as the mean
31 across the five items ($\alpha = .87$).
32
33
34
35
36
37
38
39
40
41
42
43

44 **Center for Epidemiologic Studies Depression Scale.** From subset two, 420 participants
45 completed the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977), a 20-
46 item measure of self-reported depressive symptoms. Each item describes a symptom (e.g., “I felt
47 depressed”, “I had crying spells”), and participants indicated the frequency of experiencing each
48 symptom on a 4-point scale (1 = rarely or none of the time; 4 = most or all of the time). For 42
49 participants who were missing responses for a single item, we imputed the missing item with the
50
51
52
53
54
55
56
57
58
59
60

mean of the remaining items. We excluded nine participants who were missing scores on more than one item. We calculated the mean across all items as the total scale score for the remaining 411 participants ($\alpha = .85$).

Person-Level Evaluation

In total, participants from the two subsets wrote 1,323,939 messages (each participant individually wrote at least 100 messages). We applied the temporal classifier developed in Part 1 to every message. For each participant, we calculated the number of his/her messages that were classified as past, present, or future, and then divided these three frequencies by their total number of messages, resulting in the proportions of a person's message that were past, present, and future-oriented. On average, 19% of participants' messages were past-oriented, 65% were present-oriented, and 16% were future-oriented.

Relevant language features. To better understand which language features were relevant to classification in this new set of messages, we examined which 1-grams (i.e., single words or tokens) were most strongly correlated with classifications of past, present, and future. We chose to examine 1-grams (as opposed to two or three word phrases) because they are more easily interpreted than other features used by the model. To calculate these correlations, we first recoded every message-level classification as three binary variables (e.g., past = 0/1; present = 0/1; future = 0/1), where a 1 indicated the message's orientation. For each orientation, we correlated the message-level relative frequency of single words with the corresponding binary variable. In the resulting correlations, high positive correlations indicate that greater frequency of a given word was correlated with that temporal orientation.

For each orientation, many of the most strongly correlated 1-grams included some clear temporal information, either in verb tense (e.g., *was*, *is*, or *will*) or as a part of a temporally-

1
2
3 relevant phrase. For example, the 20 1-grams most strongly correlated with past orientation were
4
5 (correlations shown in parentheses; all correlations are $p < .05$, Bonferroni-corrected) *was* (.37),
6
7 *had* (.28), *got* (.25), *did* (.16), *went* (.15), *just* (.13), *last* (.12), *made* (.12), *been* (.11), *saw* (.11), *a*
8
9 (.10), *were* (.10), *came* (.09), *said* (.09), *from* (.08), *found* (.08), *today* (.07), *didn't* (.07), *thought*
10
11 (.07), and *he* (.06). The 1-grams most correlated with present orientation included present-tense
12
13 verbs but also words likely used in interpersonal communication (e.g., second-person pronouns)
14
15 and questions: *is* (.13), *you* (.11), *love* (.09), *are* (.08), *?* (.07), *your* (.07), *happy* (.06), *don't*
16
17 (.05), *life* (.05), *like* (.05), *people* (.05), *why* (.04), *want* (.04), *can* (.04), quotation marks (“”;
18
19 .04), *know* (.04), ellipses (... ; .04), *you're* (.03), *right* (.03), and *do* (.03). The 1-grams correlated
20
21 with future orientation included future tense verbs and time-related words: *going* (.28), *to* (.22),
22
23 *tonight* (.21), *will* (.19), *wait* (.18), *be* (.12), *days* (.12), *get* (.11), *today* (.10), *go* (.10), *then* (.09),
24
25 *next* (.08), *for* (.08), *soon* (.08), *see* (.07), *until* (.06), *excited* (.06), *can't* (.05), *watch* (.05), and
26
27 *this* (.05).

28
29
30
31
32
33
34 **Age and gender.** Past and future orientation increased markedly with age; present
35
36 orientation decreased markedly. Table 2 summarizes Pearson correlations (r) between user-level
37
38 temporal orientations and age, calculated using the age-stratified subset 1. To illustrate, we
39
40 standardized user-level orientations and plotted the mean standard score of each age group for
41
42 each orientation (Figure 1; for an alternate display showing individual data points, see Figure A1
43
44 in Appendix A). Across all age groups, the rank order of past, present, and future orientation
45
46 remained the same: present-oriented messages were always the most frequent and future-oriented
47
48 were least frequent. However, there were large differences in the relative proportion of each
49
50 orientation across age.
51
52
53
54
55
56
57
58
59
60

We considered the possibility that younger users may write messages more frequently than older users, and therefore younger users would be more likely to write about the present, simply because less time has passed since writing their last message. To test whether message frequency accounted for age differences in temporal orientation, we recalculated correlations between age and orientations while adjusting for each user's total number of messages. These adjusted correlations ($r_{age \times past_adj} = .21$; $r_{age \times present_adj} = -.23$; $r_{age \times future_adj} = .16$) were virtually identical to the unadjusted correlations ($r_{age \times past} = .21$; $r_{age \times present} = -.23$; $r_{age \times future} = .16$), indicating that age differences count not be accounted for by younger users' higher message frequency.

Women were more past-oriented (overall Cohen's $d = .10$; 95% CI = [.03, .17]), less present-oriented ($d = -.27$; [-.20, -.34]), and more future-oriented than men across all ages ($d = .34$; [.27, .41]). We checked for changes in gender differences across age bins by calculating ds within each two-year age group and then regressing the ds on age. We found no significant trends in ds over age ($b_{past} = .006$, $p = .162$; $b_{present} = -.004$, $p = .397$; $b_{future} = -.001$, $p = .748$).

Personality. Temporal orientation was most strongly associated with conscientiousness and openness to experience. More future-oriented people were more conscientiousness ($r = .14$ [.10, .17]) but less open ($r = -.14$ [-.17, -.10]), while the opposite pattern occurred in more present-oriented people ($r_{conscientiousness} = -.11$, [-.14, -.07]; $r_{openness} = .09$ [.06, .12]). Table 2 lists all rs and 95% confidence intervals between orientations and Big Five personality factors, calculated within subset 1.

Impulsiveness, life satisfaction, and depressive symptoms. With subset 2, we calculated Pearson correlations between each temporal orientation and impulsiveness, satisfaction with life, and depressive symptoms. We controlled for participants' age and gender

1
2
3 by standardizing each outcome measure and temporal orientation, and then regressing temporal
4
5 orientation on each outcome, with age and gender as covariates. The resulting coefficient on
6
7 temporal orientation is equivalent to a Pearson correlation adjusted for age and gender. Higher
8
9 future orientation was significantly correlated with lower impulsiveness ($r = -.08 [-.16, -.01]$),
10
11 higher life satisfaction ($r = .07 [.02, .13]$), and fewer depressive symptoms ($r = -.16 [-.29, -.03]$).
12
13 In contrast, higher present orientation was significantly correlated with lower life satisfaction (r
14
15 = $-.08 [-.13, -.02]$) and more depressive symptoms ($r = .16 [.04, .29]$).
16
17
18
19

20 **Self-descriptions from personality items.** To complement Big Five correlations with
21
22 richer psychological descriptions, we examined IPIP personality items that were significantly
23
24 positively correlated with past, present, or future orientation for a subset of 636 participants who
25
26 completed the 100-item IPIP measure. Significant self-descriptions are listed in Table 3, and a
27
28 complete list of all items and correlations is available in Supplement 1.
29
30
31
32

33 **Discussion**

34
35 We developed a language-based measure of temporal orientation, and we applied this
36
37 method to a large sample to explore associations with age, gender, personality, and well-being.
38
39 This method may be a useful complement to existing methods, particularly when traditional self-
40
41 report measures would not be feasible.
42
43

44
45 At the message level, our temporal classifier accurately predicted the orientation of a
46
47 message, as rated by multiple human judges. At the person level, our measure of temporal
48
49 orientation converged with external correlates in theoretically expected ways. Future orientation
50
51 increased with age, whereas present orientation decreased with age. Women were more future-
52
53 oriented than men. Future orientation correlated with higher conscientiousness, and the self-
54
55
56
57
58
59
60

1
2
3 descriptions from personality items aligned with several characteristics related to different
4
5 orientations.
6
7

8 We found several small correlations between temporal orientation and Big Five
9
10 personality dimensions, but the largest were with conscientiousness; conscientious people were
11
12 more future-oriented and less present-oriented. This aligns well with characterizations of the
13
14 highly conscientious person, who plans, delays gratification, and controls impulses better than
15
16 most (Roberts, Lejuez, Krueger, Richards, & Hill, 2014). However, the correlations between
17
18 temporal orientation and the Big Five were smaller than those seen in previous mono-method,
19
20 questionnaire-based studies (absolute mean $r = .06$, versus absolute mean $r = .17$ in Zimbardo &
21
22 Boyd, 1999). One explanation for this attenuation is that the use of two different measurement
23
24 methods (language-based and questionnaire-based) prevents shared method variance from
25
26 inflating correlations (Roberts et al., 2006).
27
28
29
30
31

32 This method did replicate the expected patterns with age and gender seen in prior self-
33
34 report studies. Across ages 13 to 48, people were substantially more past- and future-oriented
35
36 and less present-oriented (Figure 1). This is consistent with trends found in studies of adolescents
37
38 and young adults (Casey et al., 2008; Steinberg et al., 2009). Age trends were similar in women
39
40 and men, but we did find a significant gender differences across all ages; women were more
41
42 future-oriented and only slightly more past-oriented, while men were more present-oriented. The
43
44 size of the gender difference was consistent with studies using self-reports (e.g., Keough et al.,
45
46 1999).
47
48
49

50 By analyzing responses to individual personality items, we found that temporal
51
52 orientation corresponded to differences in how individuals described themselves (Table 2),
53
54 particularly when contrasting present and future orientation. Highly present-oriented people may
55
56
57
58
59
60

1
2
3 be best characterized as impulsive across many domains—socially (“I cut others to pieces”),
4
5 emotionally (“I have frequent mood swings”), and motivationally (“I don’t put my mind on the
6
7 task at hand”)—but also more open to aesthetic experiences (“I believe in the importance of art”)
8
9 and fantasy (“I enjoy wild flights of fantasy”). Highly future-oriented described a much narrower
10
11 focus on practical planning (“I carry out my plans”) and getting things done (“I complete tasks
12
13 successfully”), with little interest in abstract matters (“I avoid philosophical discussions” and “I
14
15 am not interested in abstract ideas”).
16
17
18

19
20 Overall, the contrasting self-descriptions of the present-oriented and the future-oriented
21
22 are similar to *stability* and *plasticity*, two higher-order traits that describe tendencies to maintain
23
24 goals or engage with the world (Hirsh, DeYoung, & Peterson, 2009). Whereas stability is the
25
26 capacity to resist disruption and maintain action towards future goals, plasticity is the capacity
27
28 for emotional, cognitive, and environmental exploration (DeYoung, 2015). Overemphasis on the
29
30 present or the future may reflect different trade-offs between these two fundamental motivations.
31
32 In this framing, highly present-oriented people may be highly exploratory and engaged with the
33
34 environment (high plasticity) at the cost of more stable long-term goals (low stability, or
35
36 instability), while highly future-oriented people maintain a strong focus on distant goals (high
37
38 stability) at the cost of exploration and information gathering from their inner and outer worlds
39
40 (low plasticity, or rigidity).
41
42
43
44

45
46 More future-oriented people, however, were more satisfied with life and less depressed.
47
48 Because future orientation predicts favorable educational, financial, and health outcomes (Adams
49
50 & Nettle, 2009; Keough et al., 1999), it may not seem surprising that it correlates with positive
51
52 evaluations of one’s life and alleviation from psychological distress. However, this pattern was
53
54 not clear from prior research on orientations and well-being (Boniwell & Zimbardo, 2004; Zhang
55
56
57
58
59
60

1
2
3 & Howell, 2011), and our method enabled a larger study than typically possible, while removing
4
5 the evaluative confounds inherent in relying solely on self-report measures.
6
7

8 **Applications**

9
10 Our method may be most valuable as a complement to ongoing studies or existing
11 samples. Participants in a research study might be asked to voluntarily provide access to their
12 social media language (e.g., Facebook status updates or Twitter tweets), and then the classifier
13 can be applied to their posts, quickly adding a measure of temporal orientation or other
14 characteristics. Given the growing popularity of social media platforms (Duggan, Ellison,
15 Lampe, Lenhart, & Madden, 2014), language-based methods can collect large samples much
16 faster than is feasible through other approaches. For instance, human ratings of temporal
17 orientation requires about 90 seconds per message; at this rate, a single human judge would need
18 to rate continuously for over three years to annotate our collection of 1.3 million messages. Our
19 automatic classifier rated this entire set in minutes.
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 While our method annotated messages to characterize individuals, it can also potentially
35 be adapted to characterize entire geographic regions. Because social media messages often
36 contain fine-grained geographic metadata, messages from well-defined areas (e.g., U.S. counties)
37 can be aggregated, annotated, and compared by orientation. Perceptions of time and the daily
38 tempo of life vary substantially across regions and cultures (Banfield, 1974; Levine, 1997), and
39 these differences may be embedded in language and related to other important outcomes. For
40 example, a recent study of search queries found that countries differ in how much their users
41 search for information about future dates, and that more future-oriented countries have larger per
42 capita gross domestic product (Preis, Moat, Stanley, & Bishop, 2012). Similar social media
43 methods have already been used to characterize regions along psychological dimensions, such as
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 consumer confidence (O'Connor, Balasubramanian, Routledge, & Smith, 2010), life satisfaction
4
5 (Schwartz et al., 2013), and hostility (Eichstaedt et al., 2015).
6
7

8 Because we developed the model using a blend of Facebook and Twitter messages, it
9
10 may generalize to messages written on either platform, but explicit evaluations over Twitter
11
12 messages are still needed (see Sap et al., 2014 for a successful example of model building across
13
14 both platforms). However, because both Facebook and Twitter are designed to elicit descriptions
15
16 of a user's current status, they may be biased toward the present, and the relative proportions of
17
18 past-, present-, and future-oriented messages may not hold for other online social media
19
20 platforms. As users shift to other platforms, the extent to which the models need to be adjusted
21
22 should be considered.
23
24
25

26 **Limitations**

27
28 Our study also had several limitations. We used a very coarse representation of time,
29
30 splitting messages into past, present, and future categories. A fine-grained approach that
31
32 distinguishes near future from the distant future would be more sensitive to the depth of one's
33
34 temporal horizon. For example, thinking about the distant future may be a better predictor of
35
36 health and financial behaviors than only thinking about the short-term future.
37
38
39

40
41 Second, we focused only on the temporal orientation of a message and ignored other
42
43 qualities like emotional valence. Incorporating valence may allow distinctions between similarly-
44
45 oriented emotions, such as regret or positive nostalgia, which have opposite associations with
46
47 well-being (Sedikides, Wildschut, Arndt, & Routledge, 2008).
48
49

50
51 Third, our sample consisted of selected sets of social media users, who are not fully
52
53 representative of the general population. However, the representativeness of social media
54
55 continues to increase every year. Currently, 58% of all American adults use Facebook, and usage
56
57
58
59
60

1
2
3 is spread evenly across demographic and socioeconomic lines (Duggan et al., 2014). Even if the
4
5 findings only apply to the population of social media users, it still represents a considerably
6
7 larger portion of the general population than small studies with U.S. undergraduates.
8
9

10 Fourth, while our sample spanned a large age range, it did not include adults older than
11
12 48 years old. Social media use among older adults is growing every year (31% of adults over 65
13
14 use Facebook; Duggan et al., 2014), but this demographic is still underrepresented. This is
15
16 particularly limiting given our age-related findings, which contrast with the finding that “older
17
18 people are mostly present-oriented” (Carstensen, Isaacowitz, & Charles, 1999, p. 168). Our
19
20 sample may have been too young to detect such patterns.
21
22
23

24 **Conclusion**

25
26
27 Temporal orientation can be measured through everyday language on social media. Our
28
29 language-based measure of temporal orientation replicated several theoretically expected
30
31 patterns with age, gender, and personality, and allowed the discovery of new connections with
32
33 well-being. As social media expands, our approach complements other measures and can help
34
35 researchers study temporal orientation at large scale.
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this manuscript was supported by grants from the Robert Wood Johnson Foundation (#63597) and the Templeton Religion Trust (#TRT0048).

References

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Adams, J., & Nettle, D. (2009). Time perspective, personality and smoking, body mass, and physical activity: An empirical study. *British Journal of Health Psychology, 14*, 83-105.
- Adams, J., & White, M. (2009). Time perspective in socioeconomic inequalities in smoking and body mass index. *Health Psychology, 28*, 83-90.
- Banfield, E. C. (1974). *The unheavenly city revisited*. Boston: Little, Brown and Company.
- Berk, R. A. (2008). *Statistical learning from a regression perspective*. New York: Springer.
- Boniwell, I., & Zimbardo, P. G. (2004). Balancing time perspective in pursuit of optimal functioning. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 165-178). Hoboken, NJ: Wiley and Sons.
- Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist, 54*, 165-181.
- Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences, 1124*, 111-126.
- Chang, A. X., & Manning, C. D. (2012, May). SUTime: A library for recognizing and normalizing time expressions. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation*. Istanbul, Turkey. Retrieved from <http://nlp.stanford.edu/pubs/lrec2012-sutime.pdf>
- Daugherty, J. R., & Brase, G. L. (2010). Taking time to be healthy: Predicting health behaviors with delay discounting and time perspective. *Personality and Individual Differences, 48*, 202-207.
- DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality, 56*, 33-58. <http://dx.doi.org/10.1016/j.jrp.2014.07.004>

1
2
3 Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale.

4
5
6 *Journal of Personality Assessment*, 49, 71-75.

7
8 http://dx.doi.org/10.1207/s15327752jpa4901_13

9
10 Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2014). Pew Internet: Social

11
12 Media Update 2014. Retrieved from: <http://www.pewinternet.org/2015/01/09/social->

13
14
15 [media-update-2014/](http://www.pewinternet.org/2015/01/09/social-media-update-2014/)

16
17 Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., ... &

18
19 Seligman, M. E. P. (2015). Psychological language on Twitter predicts county-level heart

20
21 disease mortality. *Psychological Science*, 26, 159-169. doi:10.1177/0956797614557867

22
23
24 Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., &

25
26
27 Gough, H. G. (2006). The international personality item pool and the future of public

28
29 domain personality measures. *Journal of Research in Personality*, 40, 84-96.

30
31
32 <http://dx.doi.org/10.1016/j.jrp.2005.08.007>

33
34 Keough, K. A., Zimbardo, P. G., & Boyd, J. N. (1999). Who's smoking, drinking, and using

35
36 drugs? Time perspective as a predictor of substance use. *Basic and Applied Social*

37
38
39 *Psychology*, 21, 149-164.

40
41 Kern, M. L., Eichstaedt, J. C., Schwartz, H. A., Dziurzynski, L., Ungar, L. H., Stillwell, D. J., ...

42
43 & Seligman, M. E. (2014). The online social self: An open vocabulary approach to

44
45
46 personality. *Assessment*, 21, 386-397. <http://dx.doi.org/10.1177/1073191113514104>

47
48 Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable

49
50 from digital records of human behavior. *Proceedings of the National Academy of*

51
52
53 *Sciences*, 110, 5802-5805.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Hirsh, J. B., DeYoung, C. G., & Peterson, J. B. (2009). Metatraits of the Big Five differentially predict engagement and restraint of behavior. *Journal of Personality*, *77*, 1085-1102.
- Horstmanshof, L., & Zimitat, C. (2007). Future time orientation predicts academic engagement among first-year university students. *British Journal of Educational Psychology*, *77*, 703-718.
- Joireman, J., Shaffer, M. J., Balliet, D., & Strathman, A. (2012). Promotion orientation explains why future-oriented people exercise and eat healthy: Evidence from the two-factor Consideration of Future Consequences-14 scale. *Personality and Social Psychology Bulletin*, *38*, 1272-1287.
- Nurmi, J. E. (2005). Thinking about and acting upon the future: Development of future orientation across the life span. In A. J. Strathman & J. A. Joireman (Eds.), *Understanding behavior in the context of time: Theory, research, and application* (pp. 31-57). Mahwah, NJ: Lawrence Erlbaum.
- O'Connor, B., Balasubramanian, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *Fourth International AAAI Conference on Weblogs and Social Media*, *11*, 122-129.
- Ouellette, J. A., Hessling, R., Gibbons, F. X., Reis-Bergan, M., & Gerrard, M. (2005). Using images to increase exercise behavior: Prototypes versus possible selves. *Personality and Social Psychology Bulletin*, *31*, 610-620.
- Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., & Schneider, N. (2012). Part-of-speech tagging for Twitter: Word clusters and other advances. Retrieved from <http://www.ark.cs.cmu.edu/TweetNLP/owoputi+etal.tr12.pdf>

- 1
2
3 Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L.
4
5 H., & Seligman, M. E. P. (2015). Automatic personality assessment through social media
6
7 language. *Journal of Personality and Social Psychology*, *108*, 934-952.
8
9 <http://dx.doi.org/10.1037/pspp0000020>
10
11
12 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.,... Duchesnay, E.
13
14 (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*,
15
16 *12*, 2825–2830.
17
18
19 Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). *The*
20
21 *development and psychometric properties of LIWC2007*. Austin, TX: LIWC.net.
22
23
24 Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural
25
26 language use: Our words, our selves. *Annual Review of Psychology*, *54*, 547-577.
27
28
29 <http://dx.doi.org/10.1146/annurev.psych.54.101601.145041>
30
31
32 Piko, B. F., & Brassai, L. (2009). The role of individual and familial protective factors in
33
34 adolescents' diet control. *Journal of Health Psychology*, *14*, 810-819.
35
36
37 Potts, C. (2011). *happyfuntokenizer* (Version 1.0) [computer software]. Retrieved from:
38
39 <http://sentiment.christopherpotts.net/code-data/happyfuntokenizing.py>
40
41
42 Preis, T., Moat, H. S., Stanley, H. E., & Bishop, S. R. (2012). Quantifying the advantage of
43
44 looking forward. *Nature: Scientific Reports*, *2*, 350.
45
46
47 Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the
48
49 general population. *Applied Psychological Measurement*, *1*, 385-401
50
51
52 Roberts, B. W., Harms, P. D., Smith, J. L., Wood, D., & Webb, M. (2006). Using multiple
53
54 methods in personality psychology. In M. Eid & E. Diener (Eds.), *Handbook of*
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

psychological assessment: A multimethod perspective (pp. 321–335). Washington, DC:
American Psychological Association.

Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014). What is
conscientiousness and how can it be assessed? *Developmental Psychology, 50*, 1315-1330.
doi: [10.1037/a0031109](https://doi.org/10.1037/a0031109)

Rothspan, S., & Read, S. J. (1996). Present versus future time perspective and HIV risk among
heterosexual college students. *Health Psychology, 15*, 131-134.

Sap, M., Park, G., Eichstaedt, J. C., Kern, M. L., Stillwell, D. J., Kosinski, M., ... & Schwartz,
H. A. (2014). Developing age and gender predictive lexica over social media. *Proceedings
of the 2014 Conference on Empirical Methods in Natural Language Processing*, 1146-
1151.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Agrawal, M., Park, G., ... &
Lucas, R. E. (2013a, June). Characterizing geographic variation in well-being using tweets.
In *Seventh International AAAI Conference on Weblogs and Social Media*, Boston, MA.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M.,
..., & Ungar, L. H. (2013b). Personality, gender, and age in the language of social media:
The open vocabulary approach. *PLOS ONE, 8*, e73791.

<http://dx.doi.org/10.1371/journal.pone.0073791>

Schwartz, H. A., Park, G., Sap, M., Weingarten, E., Eichsteadt, J. C., Kern, M. L., ... & Ungar,
L. H. (2015). Extracting human temporal orientation from Facebook language. *North
American Chapter of the Association for Computational Linguistics (NAACL), Human
Language Technologies*, Denver, CO.

- 1
2
3 Sedikides, C., Wildschut, T., Arndt, J., & Routledge, C. (2008). Nostalgia: Past, present, and
4
5 future. *Current Directions in Psychological Science*, *17*, 304-307.
6
7
8 Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., & Patton, J. H.
9
10 (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality*
11
12 *and Individual Differences*, *47*, 385-395. doi: <http://dx.doi.org/10.1016/j.paid.2009.04.008>
13
14
15 Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age
16
17 differences in future orientation and delay discounting. *Child Development*, *80*, 28-44.
18
19
20 Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, C. S. (1994). The consideration of
21
22 future consequences: Weighing immediate and distant outcomes of behavior. *Journal of*
23
24 *Personality and Social Psychology*, *66*, 742-752. doi: [10.1037/0022-3514.66.4.742](http://dx.doi.org/10.1037/0022-3514.66.4.742)
25
26
27 Toutanova, K., Klein, D., Manning, C. D., & Singer, Y. (2003, May). Feature-rich part-of-speech
28
29 tagging with a cyclic dependency network. In *Proceedings of the 2003 Conference of the*
30
31 *North American Chapter of the Association for Computational Linguistics on Human*
32
33 *Language Technology*, Edmonton, Canada.
34
35
36 Webley, P., & Nyhus, E. K. (2006). Parents' influence on children's future orientation and
37
38 saving. *Journal of Economic Psychology*, *27*, 140-164.
39
40
41 Zhang, J. W., & Howell, R. T. (2011). Do time perspectives predict unique variance in life
42
43 satisfaction beyond personality traits? *Personality and Individual Differences*, *50*, 1261-
44
45 1266.
46
47
48 Zimbardo, P. G., & Boyd, J. N. (1999). Putting time in perspective: A valid, reliable individual-
49
50 differences metric. *Journal of Personality and Social Psychology*, *77*, 1271-1288.
51
52
53
54
55
56
57
58
59
60

Footnotes

¹In order to replicate how the model would be applied to the final test set, judges were not given the option to rate something as non-interpretable. We used forced-choice here because, when applying the model to messages, we cannot remove or exclude messages from classification, and this gives a more realistic assessment of how the classifier functions on a new set of text.

²While we selected the ERT model on the basis of test set performance, we also checked the ERT model accuracy in the training sample using 10-fold cross-validation. The average accuracy of the full ERT model over the training sample was 0.68.

³Although the MyPersonality sample includes participants older than 48, the sample size drops steeply with every year, and many of these users do not meet the other requirements (e.g., wrote at least 100 messages). Thus, we only included bins up to age 48.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1
Example social media messages and ratings

Message	Rating 1	Rating 2	Rating 3	Mean rating	final classification
My phone broke last weekend :(-3	-3	-3	-3	past
did nothing this morning but watch TV and it was fantastic =)	-.67	-.50	-.50	-.55	past
love this weather!!!	0	0	0	0	present
dislikes being sick... and misses her bf	0	0	0	0	present
can't wait to get a pint tonight	.33	.13	.25	.24	future
pancake day tomorrow pancake day tomorrow xxxxx	.50	.50	1	.67	future

Note: Each rater judged each message using fractions of the day in the past or future (e.g., -3 = three days in the past; .33 = eight hours in the future). Ratings were averaged for the final message classifications.

Or Peer Review

Table 2
Correlations between temporal orientation and age, gender, and individual differences

	n	past		present		future	
		r	95% CI	r	95% CI	r	95% CI
Age	3240	.21	(.18, .25)	-.23	(-.26, -.20)	.16	(.12, .19)
Gender	3240	.05	(.02, .09)	-.13	(-.17, -.10)	.16	(.13, .20)
Openness	3240	-.01	(-.04, .02)	.09	(.06, .12)	-.14	(-.17, -.10)
Conscientiousness	3240	.04	(.00, .07)	-.11	(-.14, -.07)	.14	(.10, .17)
Extraversion	3240	-.04	(-.07, .00)	-.01	(-.04, .02)	.05	(.02, .09)
Agreeableness	3240	.03	(-.01, .06)	-.06	(-.09, -.02)	.07	(.03, .10)
Neuroticism	3240	-.01	(-.05, .02)	.02	(-.01, .05)	-.02	(-.05, .02)
Impulsiveness	744	-.01	(-.09, .07)	.07	(-.01, .15)	-.10	(-.18, -.02)
Satisfaction with Life	1369	.04	(-.02, .09)	-.08	(-.13, -.02)	.08	(.02, .14)
Depression symptoms	411	-.09	(-.20, .02)	.16	(.04, .29)	-.16	(-.29, -.03)

Note: Correlations with age, gender, and Big 5 personality were calculated with subset 1; impulsiveness, satisfaction with life, and depressive symptoms were calculated with subset 2. Correlations with impulsiveness, life satisfaction, and depressive symptoms were adjusted for age and gender. Bold indicates that the 95% confidence interval did not contain zero.

464x155mm (72 x 72 DPI)

Table 3*Personality items correlated with higher past, present, and future orientation*

High past orientation	High present orientation	High future orientation
Do not like art (.10)	Cut others to pieces (.16)	Complete tasks successfully (.15)
Would describe my experiences as somewhat dull	Can say things beautifully (.14)	Avoid philosophical discussions (.12)
Am easy to satisfy (.10)	Don't put my mind on the task at hand (.12)	Carry out my plans (.12)
Rarely lose my composure (.08)	Have frequent mood swings (.12)	Finish what I start (.11)
Don't like to draw attention to myself (.08)	Am hard to get to know (.10)	Make plans and stick to them (.10)
	Believe in the importance of art (.09)	Do things according to a plan (.10)
	Know how to captivate people (.09)	Respect others (.10)
	Do just enough work to get by (.09)	Am always prepared (.10)
	Suspect hidden motives in others (.09)	Follow through with my plans (.10)
	Believe that I am better than others (.09)	Do not like poetry (.08)
	Enjoy wild flights of fantasy (.09)	Am not interested in abstract ideas (.07)
	Find it difficult to get down to work (.09)	
	Make demands on others (.08)	
	Retreat from others (.08)	
	Carry the conversation to a higher level (.08)	
	Get back at others (.08)	
	Often feel blue (.07)	
	Mess things up (.07)	

Note: N = 636. Participants indicated how accurately each statement described them. Correlations (r) are in parentheses. All correlations are adjusted for age and gender, and only correlations with 95% confidence intervals that did not contain zero are listed.

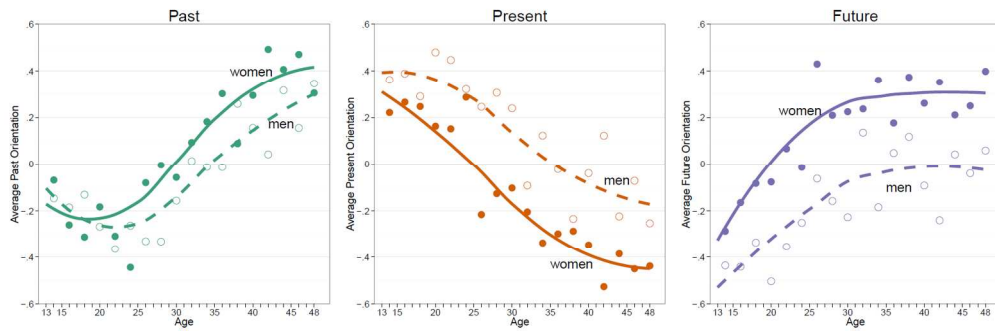


Figure 1. Average temporal orientations from ages 13 to 48. Lines are LOESS smoothers calculated across individuals, separately for women (solid lines) and men (dashed lines). Points indicate the average orientation within two-year age group (e.g., 13-14 year olds, 15-16 year olds, etc.) separately for women (shaded) and men (hollow). Each point represents 90 participants.

Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

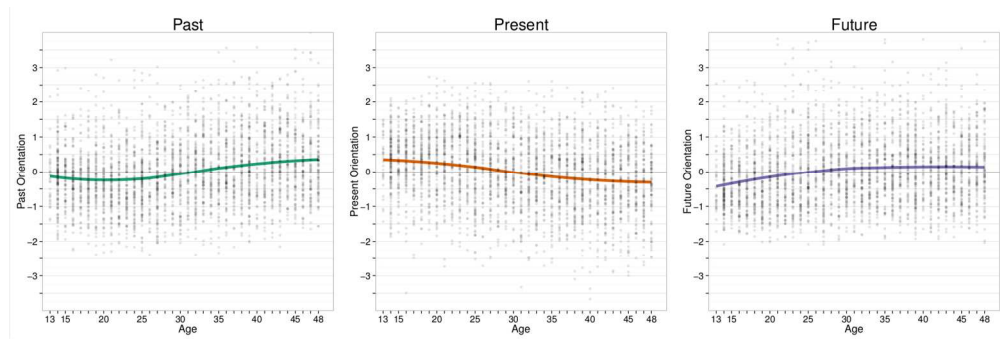


Figure A1. Average temporal orientations from ages 13 to 48. Lines are LOESS smoothers calculated across all individuals. Points indicate the average orientation of each participant.
602x199mm (72 x 72 DPI)

Peer Review

item content	Past orientation		Present orientation		Future orientation	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Complete tasks successfully	.046	.211	-.130	.000	.154	.000
Avoid philosophical discussions	.008	.829	-.086	.020	.124	.001
Carry out my plans	.063	.088	-.118	.001	.118	.001
Finish what I start	.040	.273	-.099	.007	.112	.002
Make plans and stick to them	.022	.543	-.083	.025	.105	.004
Do things according to a plan	.043	.243	-.096	.010	.104	.005
Respect others	.036	.332	-.088	.017	.099	.007
Am always prepared	.009	.816	-.069	.060	.098	.007
Follow through with my plans	.032	.381	-.085	.022	.098	.008
Do not like poetry	-.007	.852	-.049	.184	.082	.024
Am not interested in abstract ideas	-.007	.847	-.043	.242	.074	.044
Make friends easily	-.019	.604	-.034	.364	.071	.052
Do not mind being the centre of attention	-.018	.627	-.032	.388	.067	.066
Am not interested in theoretical discussions	-.011	.759	-.036	.332	.067	.068
Cheer people up	-.027	.458	-.023	.527	.064	.082
Rarely look for a deeper meaning in things	-.021	.566	-.026	.479	.062	.092
Tend to vote for conservative political candidates	.022	.553	-.054	.144	.061	.095
Am very pleased with myself	.006	.869	-.042	.255	.059	.109
Seldom feel blue	.041	.266	-.065	.080	.058	.113
Do not like art	.100	.006	-.104	.005	.058	.111
Get chores done right away	-.081	.028	.018	.619	.054	.145
Get stressed out easily	-.022	.546	-.020	.579	.053	.139
Talk to a lot of different people at parties	-.016	.656	-.023	.540	.052	.159
Have difficulty understanding abstract ideas	-.001	.986	-.033	.372	.051	.159
Am skilled in handling social situations	-.022	.550	-.018	.631	.050	.175
Make people feel at ease	-.024	.504	-.016	.663	.050	.175
Believe that others have good intentions	.047	.204	-.062	.093	.049	.186
Feel comfortable around people	-.009	.802	-.025	.500	.048	.191
Am exacting in my work	.025	.495	-.047	.195	.048	.187
Am easy to satisfy	.096	.009	-.094	.011	.047	.204
Am the life of the party	-.065	.076	.016	.675	.042	.248
Pay attention to details	-.009	.801	-.020	.585	.041	.269
Rarely get irritated	.018	.622	-.037	.319	.038	.294
Start conversations	-.018	.621	-.013	.735	.038	.303
Do not enjoy going to art museums	.068	.064	-.069	.060	.038	.299
Am concerned about others	-.013	.720	-.016	.672	.037	.306
Worry about things	-.031	.397	-.004	.912	.037	.300
Warm up quickly to others	-.011	.774	-.016	.659	.036	.328
Rarely lose my composure	.078	.033	-.073	.046	.034	.355
Am not easily frustrated	.011	.770	-.025	.491	.028	.440
Trust what people say	.059	.110	-.057	.126	.028	.451
Fear for the worst	-.038	.290	.009	.804	.025	.490
Believe that I am better than others	-.158	.000	.089	.015	.024	.510

1								
2		Treat all people equally	.047	.201	-.044	.229	.021	.571
3		Remain calm under pressure	-.025	.497	.004	.918	.019	.596
4	Believe that too much tax money goes to support artists		.003	.936	-.014	.698	.019	.602
5		Hold a grudge	.024	.518	-.026	.481	.016	.662
6		Feel comfortable with myself	-.020	.591	.004	.921	.014	.693
7		Sympathise with others feelings	.010	.792	-.016	.675	.014	.702
8		Find it difficult to approach others	-.058	.108	.032	.377	.009	.795
9		Panic easily	.006	.874	-.009	.796	.009	.810
10		Am out for my own personal gain	-.011	.761	.002	.949	.008	.834
11		Accept people as they are	.002	.954	-.006	.876	.007	.854
12		Am not easily bothered by things	.034	.348	-.017	.652	-.009	.799
13		Am relaxed most of the time	.049	.181	-.026	.483	-.010	.787
14		Make demands on others	-.117	.001	.085	.023	-.011	.756
15		Have a good word for everyone	.026	.474	-.008	.828	-.014	.696
16		Dont like to draw attention to myself	.077	.037	-.039	.292	-.018	.631
17		Get excited by new ideas	.040	.284	-.011	.759	-.023	.540
18		Have a vivid imagination	-.017	.654	.027	.472	-.024	.509
19		Need a push to get started	.014	.710	.008	.831	-.026	.477
20		Enjoy hearing new ideas	.037	.314	-.005	.889	-.030	.418
21	Would describe my experiences as somewhat dull		.100	.007	-.045	.221	-.031	.390
22		Have a sharp tongue	-.001	.981	.021	.563	-.032	.382
23		Seldom get mad	.027	.450	.005	.892	-.036	.327
24		Have little to say	-.012	.734	.033	.373	-.038	.298
25		Make a mess of things	-.041	.252	.054	.135	-.041	.248
26		Insult people	-.018	.616	.039	.289	-.041	.255
27		Dont see things through	-.002	.955	.030	.421	-.044	.232
28		Shirk my duties	.024	.504	.017	.631	-.051	.154
29		Contradict others	-.025	.499	.052	.160	-.055	.134
30		Feel threatened easily	.013	.718	.028	.442	-.056	.118
31	Tend to vote for liberal political candidates		-.041	.271	.066	.076	-.060	.100
32		Get back at others	-.054	.140	.077	.037	-.064	.081
33		Enjoy thinking about things	.068	.064	-.004	.923	-.064	.083
34		Am filled with doubts about things	-.005	.896	.045	.216	-.065	.073
35		Dislike myself	-.009	.813	.048	.188	-.066	.071
36		Dont talk a lot	.043	.234	.017	.649	-.070	.055
37		Keep others at a distance	.031	.404	.026	.486	-.071	.053
38		Have frequent mood swings	-.103	.004	.115	.001	-.073	.043
39		Avoid contact with others	.058	.116	.009	.805	-.073	.047
40		Mess things up	-.038	.284	.073	.043	-.074	.039
41		Am often down in the dumps	.044	.227	.019	.600	-.075	.041
42		Know how to captivate people	-.060	.105	.089	.016	-.077	.038
43		Keep in the background	.033	.363	.029	.442	-.078	.034
44		Leave things unfinished	-.009	.813	.057	.123	-.079	.031
45		Have a rich vocabulary	.053	.148	.018	.624	-.082	.025
46		Suspect hidden motives in others	-.053	.145	.089	.016	-.082	.024
47		Carry the conversation to a higher level	-.037	.318	.079	.033	-.084	.022
48		Often feel blue	-.028	.444	.074	.048	-.085	.021
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								

1							
2	Enjoy wild flights of fantasy	-.046	.210	.088	.017	-.089	.015
3	Believe in the importance of art	-.054	.139	.094	.011	-.089	.015
4	Waste my time	.030	.405	.038	.301	-.090	.014
5	Cut others to pieces	-.130	.000	.156	.000	-.108	.003
6	Dont put my mind on the task at hand	-.076	.038	.124	.001	-.114	.002
7	Find it difficult to get down to work	-.017	.633	.086	.018	-.115	.001
8	Do just enough work to get by	-.021	.551	.089	.014	-.115	.001
9	Am hard to get to know	-.029	.431	.100	.007	-.124	.001
10	Can say things beautifully	-.082	.026	.135	.000	-.125	.001
11	Retreat from others	.003	.945	.084	.023	-.132	.000
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							
39							
40							
41							
42							
43							
44							
45							
46							
47							
48							
49							
50							
51							
52							
53							
54							
55							
56							
57							
58							
59							
60							

For Peer Review