Political psycholinguistics: A comprehensive analysis of the language habits of liberal and conservative social media users

Article *in* Journal of Personality and Social Psychology · January 2020

DOI: 10.1037/pspp0000275

John T Jost
New York University

Richard Bonneau
New York University

Journal of Personality and Social Psychology

Political Psycholinguistics: A Comprehensive Analysis of the Language Habits of Liberal and Conservative Social Media Users

Joanna Sterling, John T. Jost, and Richard Bonneau

Online First Publication, January 9, 2020. http://dx.doi.org/10.1037/pspp0000275

CITATION

Sterling, J., Jost, J. T., & Bonneau, R. (2020, January 9). Political Psycholinguistics: A Comprehensive Analysis of the Language Habits of Liberal and Conservative Social Media Users. *Journal of Personality and Social Psychology*. Advance online publication. http://dx.doi.org/10.1037/pspp0000275

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Personality and Social Psychology:
Personality Processes and Individual Differences

ISSN: 0022-3514

http://dx.doi.org/10.1037/pspp0000275

Political Psycholinguistics: A Comprehensive Analysis of the Language Habits of Liberal and Conservative Social Media Users

Joanna Sterling
Princeton University

John T. Jost and Richard Bonneau
New York University

For nearly a century social scientists have sought to understand left–right ideological differences in values, motives, and thinking styles. Much progress has been made, but—as in other areas of research—this work has been criticized for relying on small and statistically unrepresentative samples and the use of reactive, self-report measures that lack ecological validity. In an effort to overcome these limitations, we employed automated text analytic methods to investigate the spontaneous, naturally occurring use of language in nearly 25,000 Twitter users. We derived 27 hypotheses from the literature on political psychology and tested them using 32 individual dictionaries. In 23 cases, we observed significant differences in the linguistic styles of liberals and conservatives. For instance, liberals used more language that conveyed benevolence, whereas conservatives used more language pertaining to threat, power, tradition, resistance to change, certainty, security, anger, anxiety, and negative emotion in general. In 17 cases, there were also significant effects of ideological extremity. For instance, moderates used more benevolent language, whereas extremists used more language pertaining to inhibition, tentativeness, affiliation, resistance to change, certainty, security, anger, anxiety, negative affect, swear words, and death-related language. These research methods, which are easily adaptable, open up new and unprecedented opportunities for conducting unobtrusive research in psycholinguistics and political psychology with large and diverse samples.

Keywords: political ideology, psycholinguistics, quantitative text analysis, social cognition, social media

> Linguistic habits are often important symptoms of unspoken sentiments.
>
> —Umberto Eco, 2001, *Five Moral Pieces*

With the concept of parapraxis—an error or slip in speech that betrays one's underlying motivational state—Sigmund Freud was among the first psychologists to regard language as a device for exploring the recesses of the human mind, but he was by no means the last. To name just a few, Hermann Rorschach (1921), Henry Murray (1943), and David McClelland (1979) developed pioneering methods for inferring mental states on the basis of verbal descriptions that people offered in response to inkblots, drawings, or hypothetical social episodes. These methods required content analyses—performed by highly trained coders—that were notoriously effortful and time-intensive (see also Sorrentino & Roney, 2000).

Contemporary psychologists are fortunate to have at their disposal far more efficient technologies for conducting automatic analyses of huge corpuses of linguistic texts. One important historical step was the creation of *The General Inquirer* (Stone, Dunphy, Smith, & Ogilvie, 1966), which was a computer program that classified language in terms of the expression of psychological needs for achievement, affiliation, and power, as specified by the motivational theory of McClelland, Atkinson, Clark, and Lowell (1953). One of the most popular software platforms for estimating psychological characteristics on the basis of spoken or written language is the Linguistic Inquiry and Word Count (LIWC) program (e.g., Pennebaker, Booth, Boyd, & Francis, 2015; Pennebaker & King, 1999; Pennebaker, Mehl, & Niederhoffer, 2003).

In the present research program, we employ LIWC and several other dictionary-based methods for the quantitative analysis of textual data in research on personality and social psychology. Dictionary-based methods make use of hybrid (or semisupervised) procedures that merge qualitative and quantitative methods. On the qualitative side, researchers identify concepts and catego-

Joanna Sterling, Department of Psychology, Princeton University; John T. Jost, Department of Psychology, New York University; Richard Bonneau, Center for Genomics and Systems Biology, New York University.

This article is based on a doctoral dissertation submitted to the Department of Psychology at New York University by Joanna Sterling under the supervision of John T. Jost. We thank the other members of the dissertation committee, namely Eric Knowles, Patrick Shrout, and Jay Van Bavel. This research was supported by the INSPIRE program of the National Science Foundation (Awards # SES-1248077 and # SES-1248077-001) as well as the Global Institute for Advanced Study (GIAS) and Research Investment Fund (RIF) at New York University. John T. Jost also gratefully acknowledges support from National Science Foundation Award #BCS-1627691. We thank Duncan Penfold-Brown, Jonathan Ronen, and Yvan Scher for their invaluable assistance with computer programming and data procurement. Study materials and code for data analysis have been made accessible on the Open Science Framework. The data provided have been deidentified per Twitter's terms of service: https://osf.io/4fsva/.

Correspondence concerning this article should be addressed to John T. Jost, Department of Psychology, New York University, Meyer Hall, 6 Washington Place, 6th Floor, New York, NY 10003. E-mail: john.jost@nyu.edu

ries of interest to them. Dictionaries are constructed by means of thesaurus-style language searches, in which a researcher or group of researchers identifies key terms or concepts that pertain to a shared underlying construct. Once the dictionaries are formed, the analysis is purely quantitative. A computer program automatically tallies word counts based on the constructed dictionaries with perfect reliability. The research goal is to capture salient psychological themes or dimensions in the thinking or communication styles of individuals or groups of people.

The benefits of using text analysis are numerous. For one thing, language is a form of behavior, the act of solidifying a piece of information in one's own memory or communicating it to another person or group. Thus, linguistic analysis enables researchers to study naturally occurring behavior rather than attitudes or beliefs in isolation. Importantly, the psycholinguistic cues embedded in one's communications are often not noticeable in a single instance; patterns emerge only in aggregation when processed by computers or trained researchers (Pennebaker & King, 1999; Pennebaker et al., 2003). For instance, studies employing the LIWC platform have turned up a number of interesting discoveries, such as the following: (a) people who are high (vs. low) in social status use more first person plural pronouns and fewer first person singular pronouns (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2014); and (b) people who are more (vs. less) feminine use more references to others and more positive feeling words, as well as fewer negations, articles, prepositions, swear words, long words, references to money, and numbers (Newman, Groom, Stone, & Pennebaker, 2006).

Liberal–Conservative Differences in Psychological Characteristics

For almost as long as Freud and his followers have been analyzing linguistic behavior, social scientists have sought to identify the social, cognitive, and motivational underpinnings of political ideology (e.g., Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; Eysenck, 1954; Jost, Glaser, Kruglanski, & Sulloway, 2003a, 2003b; Lane, 1962; McClosky, 1958; Rokeach, 1960; Tomkins, 1963; Wilson, 1973). Political ideology is often measured by asking participants to indicate their position on a continuous scale (ranging from very liberal [or leftist] to very conservative [or rightist]) or their evaluations of specific political parties, candidates, issues, or opinions. Applying these methods in the context of surveys, questionnaires, interviews, experiments, and observational studies, researchers have made considerable progress over the last 15 years or more in understanding psychological differences, including needs for certainty and security, between self-identified liberals and conservatives (e.g., Caprara & Vecchione, 2017; Carney, Jost, Gosling, & Potter, 2008; Gerber et al., 2011; Jost, 2017a, 2017b; Jost et al., 2003a, 2003b; Sibley, Osborne, & Duckitt, 2012). Nevertheless, there are clear limitations associated with the use of these research methods, which rely heavily upon self-report measures of political and psychological characteristics and are often based on fairly small sample sizes.

The advent of social media and the pervasive use of communication platforms such as Twitter and Facebook create new and unprecedented opportunities for studying the linguistic behavior of many thousands of citizens (see Jost, Barberá, et al., 2018, for a review). Linguistic analysis of publicly available social media data allows researchers to access much larger and more statistically representative samples of individuals and texts in a more cost-effective manner than is possible using traditional laboratory methods. Through the analysis of big data sources generated by social media usage, researchers have begun to explore the role of political ideology in the structure and function of online social networks (Barberá, Jost, Nagler, Tucker, & Bonneau, 2015; Boutyline & Willer, 2017), the expression of personal values (Jones et al., 2018; Neiman, Gonzalez, Wilkinson, Smith, & Hibbing, 2016a, 2016b), the use of moral and emotional language (Brady, Wills, Jost, Tucker, & Van Bavel, 2017; Robinson, Boyd, & Fetterman, 2014; Sterling & Jost, 2018; Sylwester & Purver, 2015; Wojcik, Hovasapian, Graham, Motyl, & Ditto, 2015), and support versus opposition to protest activity (Langer, Jost, Bonneau, Metzger, Noorbaloochi, et al., 2019; Smith, Gavin, & Sharp, 2015; Theocharis, Lowe, van Deth, & García-Albacete, 2015), to take just a few examples.

In the present research program we used dictionary-based methods to provide a comprehensive assessment—using large samples of texts generated by thousands of ordinary citizens—of 27 hypotheses derived from several different theories about the nature of psychological differences between liberals (or leftists) and conservatives (or rightists). We used dictionaries that have been frequently used in the psychological literature to measure political ideology but that have not been comprehensively evaluated elsewhere (e.g., Sterling & Jost, 2018). To our knowledge, this is the most ambitious and comprehensive investigation to date of political psycholinguistics, that is, the study of ideological divergence in the use of language (e.g., see Lakoff, 2004; Wetherell & Potter, 1992). Because of the number and breadth of theoretical perspectives under consideration, we have divided the article into five sections to address: (a) dispositional motives, (b) personal values, (c) motivated social cognition, (d) needs for uniqueness and conformity, and (e) emotional expression. Before turning to these specific theoretical domains, however, we first consider the role of ideological extremity, because previous research suggests that—in addition to qualitative psychological differences between liberals and conservatives—there are also differences between extremists and moderates (Jost et al., 2003a).

The Role of Ideological Extremity

As noted above, many studies conducted over the past 15 years have revealed that several key psychological characteristics covary with liberal-conservative (or left-right) orientation. At the same time, it has been suggested that ideological extremity (whether on the left or right) should be associated with certain psychological characteristics, including dogmatism and mental rigidity (e.g., Eysenck, 1954/1999; Greenberg & Jonas, 2003; van Prooijen & Krouwel, 2019; but see Sidanius, 1988, for a different view). Furthermore, according to terror management theory existential needs to increase self-esteem and decrease death anxiety should motivate extreme forms of worldview defense (Anson, Pyszczynski, Solomon, & Greenberg, 2009; Castano et al., 2011; Greenberg & Jonas, 2003). The idea here is that feelings of threat should be associated with ideological extremity in general rather than right-wing conservatism per se.

With the rise of big data, researchers are better situated than ever before to model complex, nonlinear effects including simultaneous influences of conservatism and extremity. Following Jost and colleagues (2003a), we hypothesized that—for some psychological variables, such as needs for certainty and security—there will be an asymmetrical quadratic effect such that epistemic and existential motives to reduce uncertainty and insecurity will be stronger on the right than the left and, at the same time, at ideologically extreme rather than moderate positions. For these variables, the overall pattern is expected to resemble a J, such that extreme liberals should be more focused on certainty and security than moderate liberals, moderate conservatives should be more focused on certainty and security than moderate liberals, and extreme conservatives should be more concerned with certainty and security than both moderate conservatives and extreme liberals (see Jost et al., 2003a; Sterling, Jost, & Pennycook, 2016). By unobtrusively assessing "real-world" communication behaviors in very large samples of social media users, we conduct the most comprehensive investigation of these curvilinear, J-shaped patterns to date.

Overview of the Present Research Program

A number of previously published studies, which we discuss in more detail below, reveal that there are meaningful connections between political ideology and language use (e.g., Brundidge, Reid, Choi, & Muddiman, 2014; Cichocka, Bilewicz, Jost, Marrouch, & Witkowska, 2016; Kemmelmeier, 2008; Neiman et al., 2016b; Okdie & Rempala, 2019; Robinson et al., 2014; Tetlock, 1983; Tetlock, Bernzweig, & Gallant, 1985). However, most of these studies focused on a relatively small number of linguistic categories and made use of relatively small samples (in some cases, necessarily so, given that they were focused on the rhetoric of political elites). In our judgment, a more comprehensive approach is necessary to more fully explore the topic of political psycholinguistics. In the present research program, we integrate findings from five major theoretical domains to estimate the magnitudes of ideological effects with respect to a wide range of psychological variables. We also estimate the extent to which ideological extremity is associated with each psychological variable over and above the effects of left-right ideology. In addition to synthesizing and advancing theoretical contributions to this area of research, we also make a methodological contribution by empirically evaluating several common assumptions employed in all linguistic research that relies on word count methods.

To advance these goals, we derived 27 language categories from the LIWC program (Pennebaker et al., 2015) and other sources (e.g., Harvard IV/General Inquirer, Stone et al., 1966). Each of these language categories has been linked in previous work to psychological characteristics associated with political ideology. In light of the steeply increasing popularity of dictionary-based methods in social and personality psychology, we chose to implement (and validate) methods currently in use rather than to generate our own sets of dictionaries.

LIWC is a dictionary-based software program frequently utilized in psychological research because of its extensive scale (it includes approximately 6,400 word components and 90 psychologically relevant word categories), reliability (Kahn, Tobin, Massey, & Anderson, 2007; Pennebaker & King, 1999; Pennebaker et al., 2003; Tausczik & Pennebaker, 2010), and transparency of method. Dictionaries used to measure other variables were adapted from those used by Neiman and colleagues (2016a). We investigated the degree to which these linguistic indicators of psychological characteristics were related to a given communicator's political ideology in several language samples. We retrieved approximately 3,200 tweets[1] sent by each of nearly 25,000 Twitter users.

For the sake of clarity and concision, we summarize our findings in five separate sections, each of which covers one of the major theoretical domains that we investigated. The sample, methods, and statistical analyses we employed in each of the five domains were identical. To minimize redundancy, we start by describing the shared methodological framework before delving into details of the specific theoretical domains investigated in the five substantive sections of the article.

Method

In the present research program, we observed naturally occurring differences in language use as ecologically valid and unobtrusive measures of the psychological underpinnings of political ideology in a large sample of ordinary social media users. First we assessed directional effects pertaining to liberal-conservative differences in language usage and then nondirectional effects pertaining to differences between moderates and extremists. Drawing on five domains of theory-driven research (dispositional motives, personal values, motivated social cognition, needs for uniqueness and conformity, and emotion), we distilled a total of 27 hypotheses to investigate using dictionary-based methods. In Table 1, we provide details pertaining to the original dictionary sources as well as sample words and a summary of our major findings.

Language Sample

We started by selecting a random sample of 50,000 Twitter users from Barberá's (2015) U.S. Election dataset, which contains more than 20 million Twitter accounts in total (https://github.com/pablobarbera/twitter_ideology/tree/master/2016-election). We used this method of selection because it enabled us to draw on the most complete list of Twitter users for whom we could estimate political ideology (our key explanatory variable, see details below). After excluding non-English language accounts, a final sample of 24,988 Twitter users remained; each user contributed between 1 and 3,244 Tweets ($M = 468.37$, $SD = 871.05$), resulting in a total sample of 11,703,650 tweets.[2] We cannot speak definitively to the demographic characteristics of the Twitter users included in our sample. We do know from a study conducted by Pew Research Center in 2016 that although Twitter users are far more statistically representative of the U.S. population than convenience samples typically used in psychological research, they are by no means perfectly representative of the population (Pew Research Institute, 2016). According to Pew, Twitter users make up about a quarter of adults who are online. They tend to be

[1] Despite limits publicized by Twitter, the maximum number of tweets researchers can retrieve from a particular user's public account is, in practice, slightly greater than 3,200.

[2] This data set was also used in research by Sterling and Jost (2018), which tested a different set of hypotheses derived from moral foundations theory.

Table 1
Language Categories, Hypotheses, Results, Citations, Toolkits Used, and Percentages of Total Words Included in Each Category

Language aspect	Hypothesis/used more often by?	Result/hypothesis supported?	Citation/previous research	Toolkit used	Pct. of total words
Dispositional motives					
Affiliation V1	Liberals	Null	**Fetterman, Boyd, & Robinson, 2015**	LIWC 2015	2.86%
Affiliation V2	Liberals	Disconfirmed	**Fetterman et al., 2015**	Harvard IV	3.49%
Power V1	Conservatives	**Confirmed**	**Fetterman et al., 2015**	LIWC 2015	2.36%
Power V2	Conservatives	**Confirmed**	**Fetterman et al., 2015**	Harvard IV	2.39%
Personal values					
Conformity	Conservatives	Null	**Neiman et al., 2016a, 2016b**	Neiman et al., 2016a, 2016b	.00%
Power	Conservatives	**Confirmed**	**Neiman et al., 2016a, 2016b**	Neiman et al., 2016a, 2016b	.07%
Security	Conservatives	**Confirmed**	**Neiman et al., 2016a, 2016b**	Neiman et al., 2016a, 2016b	.03%
Tradition	Conservatives	**Confirmed**	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.08%
Achievement V1	Conservatives	**Confirmed**	Sylwester & Purver, 2015	LIWC 2015	1.44%
Achievement V2	Conservatives	**Confirmed**	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.16%
Benevolence	Liberals	**Confirmed**	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.01%
Universalism	Liberals	Null	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.00%
Stimulation	Liberals	Null	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.01%
Self-direction	Liberals	Null	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.15%
Social	Liberals	Disconfirmed	*Novel prediction*	LIWC 2015	9.40%
Motivated social cognition					
Certainty V1	Conservatives	**Confirmed**	Sylwester & Purver, 2015	LIWC 2015	1.07%
Certainty V2	Conservatives	**Confirmed**	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.01%
Resistance to change	Conservatives	**Confirmed**	**Neiman et al., 2016a, 2016b**	Neiman et al., 2016a, 2016b	.11%
Inequality	Conservatives	Null	Neiman et al., 2016a, 2016b	Neiman et al., 2016a, 2016b	.11%
Past focus	Conservatives	**Confirmed**	**Robinson, Cassidy, Boyd, & Fetterman, 2015**	LIWC 2015	1.57%
Tentative	Liberals	Disconfirmed	Sylwester & Purver, 2015	LIWC 2015	1.24%
Future focus	Liberals	Disconfirmed	**Robinson et al., 2015**	LIWC 2015	.97%
Anxiety	Conservatives	Null	**Robinson, Boyd, & Fetterman, 2014**; Sylwester & Purver, 2015	LIWC 2015	.17%
Inhibition	Conservatives	**Confirmed**	*Novel prediction*	LIWC 2007	.47%
Threat	Conservatives	**Confirmed**	**Neiman et al., 2016a, 2016b**	Neiman et al., 2016a, 2016b	.07%
Risk focus	Conservatives	**Confirmed**	*Novel prediction*	LIWC 2015	.39%
Death	Conservatives	Null	Sylwester & Purver, 2015	LIWC 2015	.19%
Uniqueness and conformity motivations					
I	Liberals	Null	**Sylwester & Purver, 2015**	LIWC 2015	2.57%
We	Conservatives	**Confirmed**	**Sylwester & Purver, 2015**	LIWC 2015	.66%
Emotion					
Positive emotion	Liberals	Null	**Sylwester & Purver, 2015**; **Wojcik, Hovasapian, Graham, Motyl, & Ditto, 2015**	LIWC 2015	4.47%
Negative emotion	Conservatives	**Confirmed**	Sylwester & Purver, 2015; **Wojcik et al., 2015**	LIWC 2015	1.49%
Anger	Liberals	Disconfirmed	**Robinson, et al., 2014**	LIWC 2015	.53%
Anxiety	Conservatives	Null	**Robinson, et al., 2014**; Sylwester & Purver, 2015	LIWC 2015	.17%

Note. Results that supported the hypothesis in the present study (third column) and in previous research (fourth column) are identified in boldface. The last column refers to the total percent of words from each category out of the sum total of all words used.

younger, more educated, and of a higher socioeconomic status in comparison with the general population.

To ensure that our estimates were not dependent upon language patterns exhibited by particular types of users, we conducted a series of 500 nested bootstraps. In each bootstrap, we first sampled 80% of the Twitter users with replacement. Within the selected 80%, we then sampled with replacement 500 of each user's tweets (or, if they sent fewer than 500 tweets, we retained their total number) to accommodate the matrix limitations on most statistical software. This also helped ensure that any effects we obtained were not contingent upon specific periods of time.[3] All effect size confidence intervals are based on the results of these bootstraps. This research was considered exempt from federal oversight at 45 CFR 46 101(b) [2] by the Institutional Review Board of New York University (Protocol #12–9058).

Before coding the tweets, we preprocessed the data to remove URLs, hashtags, and punctuation. All tweets were coded using the LIWC 2015 program. Dictionaries that were not intrinsic to LIWC were manually converted to a LIWC format prior to conducting analyses. LIWC outputs the percentage of words that make up each category of interest out of the total number of words used in the document (in our case, a tweet). Twitter's character constraints at the time (140 characters) made raw percentages especially unstable. Therefore, we recoded the dependent measures so that they were dichotomous indicators of whether a given tweet did (1) or did not (0) contain at least one word from the relevant category.

Measures

To estimate each Twitter user's political ideology, we leveraged the online social networks in which individuals are embedded using quantitative methods developed and validated by Barberá (2015), who (a) identified approximately 300 elite Twitter accounts that span the ideological spectrum and (b) collected all tweets by ordinary citizens who follow those accounts. Rather than assigning ideological weights to elites through traditional means, such as expert coding, we used a statistical model (specifically, a Bayesian Spatial Following model) to estimate the ideological placements of elite and ordinary users based on patterns of *following* behavior. The method operates according to the notion that an individual's ideological position can be predicted by the average ideological position of the elites he or she follows (see also Barberá et al., 2015).

This method of imputing ideology relies upon an assumption of political homophily: Twitter users generally prefer to follow political elites who hold ideological positions that are more (vs. less) similar to their own (as in spatial voting models; Enelow & Hinich, 1984). It estimates the positions of both individual users and elites based on the clustering that occurs in their follower networks. Individuals who belong to the same clusters are predicted to hold similar ideological positions. Ideological estimates obtained in this way were validated by Barberá (2015) for both elite and mass samples through expert ratings, state-level ideological estimates, individual users' party registrations, and campaign contributions, as well as self-identification on Twitter. The average ideological position in our sample was extremely close to the scale midpoint of zero ($M = -.01$, $SD = 1.14$).

Statistical Analyses

We first conducted a series of multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users (using Stata, Version 13) on each of the bootstrapped samples. We regressed the probability that a tweet contained at least one word from each of the relevant dictionaries on the Twitter user's mean-centered ideological position. We also adjusted for the log-transformed word count of each tweet, because longer tweets are more likely to contain all types of language. The data set was organized at the level of the individual tweet, with the regression fitting the probability of a given word type being included in a single tweet. Tweets were nested within Twitter user, with a minimum of 1 and a maximum of 500 tweets composed by a single Twitter user ($M = 416.64$, $SD = 144.88$). This was done to minimize the likelihood that any model would be dominated by the word use patterns of a small number of hyper-active users. Reported effect sizes for political ideology are based on the median model from 500 bootstraps (where each bootstrap is a random draw from each user as described above). Confidence bounds consist of the smallest and largest effect sizes for ideology. We first built models with linear and fixed effects for all word types. After the construction of linear models, we fit similar models with an additional quadratic term that accounted for the symmetrical effect of ideological extremity. This model was then run on all bootstrapped samples used in the first set of analyses.

We also conducted nonparametric analyses for each word type (to evaluate the null, linear, and quadratic models described above). Parametric statistics, like the linear and quadratic regressions we conducted, impose model assumptions on how the data deviate from predicted values. Given these assumptions, the researcher determines whether each model performs better than a null model. Model comparison can lead to problems in which some functional forms selected (i.e., the models) are inappropriate; here we seek to detect trends with methods that do not make strong assumptions about the functional forms that generate the effect. Parametric statistics can detect significant evidence for a model that accounts for some amount of variance in the data, but over the range of predictors, certain regions may not be explaining variance in the dependent variable accurately. For all of these reasons, we modeled expected deviations of ratios of word-types from mean ratios (word type enrichments in bins of ideologically sorted users) with a nonparametric method. These model-free methods are useful for both detecting and visualizing trends and also aid in evaluating the quality of parametric regression models discussed above. We plot predictions from these enrichment analyses alongside the regression models to evaluate which segments of the data are being misrepresented (e.g., see Figure 3). In this way, we can also use enrichment analyses to determine whether the patterns in our data set generalize to other samples drawn from the same population.

Enrichment analyses proceed on the basis of the null hypothesis that some variable of interest—here, the use of a particular linguistic category as represented by frequencies of words from a word-type dictionary—is equally distributed throughout all observations within a data set. The null (or baseline) ratio is the proportion of observations within the whole dataset that contains words reflecting the variables of interest. To begin with, the

[3] Each sample of 500 tweets made up less than 1/6 of the total number of tweets we could collect from a particular individual. If an effect were limited to a particular point in time, there would be high variance in the point estimate (depending on whether the relevant point in time was included in the sample or not). In this way, bootstrapping allows us to account for time-specific effects by including their variation in our general error estimates.

dataset is rank-sorted by ideology and divided into equally sized bins. For our dataset this produces bins of observations with very low ideology estimates (extremely liberal), bins with very high ideology estimates (extremely conservative), and bins arranged throughout the spaces between. The analysis then proceeds to compare the proportion of observations within each bin containing the variable of interest with the proportion of observations in the whole dataset that contains the same variable. As in a chi-square analysis, we then test whether the concentration of the variable in each bin is statistically likely or unlikely given what we know about its concentration in the whole dataset. This enables us to determine whether certain levels of the construct on which the data are sorted are associated with more or less frequent occurrences of the variables of interest.

Thus, we first sorted the dataset based on the political ideology of each Twitter user (using R, Version 3.3.2; see https://osf.io/4fsva/) and then split the data set into 50 equally sized groups, each of which contained 240,000 tweets. We tested whether these groups of tweets, which were arranged along the ideological spectrum, included significantly more or less instances of certain linguistic categories than we would expect on average. Because each bin was tested individually, no particular model was imposed on the dataset. Significance was determined using a multiple hypothesis corrected alpha (adjusted $\alpha = .0000213$). After the nonparametric enrichment analyses were completed, we imposed the predicted lines from our two regression analyses (i.e., linear [dotted lines] and quadratic [dashed lines] models) on top of the data averages (solid lines) to assess model fit for each language category (see Figure 3 and Appendix B for analysis details).

All reported effects are in the form of odd ratios, which confer intuitive measures of effect size. The odds ratio can be interpreted as the increase in the odds of a tweet containing a particular type of language for each unit increase in ideology. Because ideology was coded so that higher numbers indicated greater political conservatism, numbers larger than one reveal that conservatives used more words from a given category than liberals, whereas numbers less than one reveal that liberals used more words from that category than conservatives.

In addition to ideological differences in the proportions of tweets expressing each psychological construct, we were also interested in potential differences in the specific words used by each ideological group to express each construct. To compose lists of the most frequently used words, we divided the sample of Twitter users into three groups: liberals, moderates, and conservatives. Moderates were defined as having ideology estimates within one standard deviation ($SD = 1.065$) around zero (an estimate implying that a given Twitter user is neither strongly liberal nor strongly conservative; 13,165 Twitter users). Liberals were defined as having ideology estimates below this cutoff (estimates < -0.532; 7,153 Twitter users in total). Conservatives were defined as having ideology estimates above this cutoff (estimates > 0.532; 4,670 Twitter users in total). We then took all of the tweets from liberal and conservative Twitter users and separately derived counts for each word within each dictionary (using the Quanteda package in R, Version 3.3.2, https://osf.io/4fsva/). We recorded the 10 most frequently used words for liberals and conservatives for each dictionary (see Table 2).

Results

Dispositional Motives: Affiliation and Power

According to a long and distinguished tradition of scientific research on human motivation, Atkinson (1958), McClelland (1987), and Winter (1991) proposed that there are three major dispositional needs or motives—defined as "recurrent concerns about a goal state that drives, orients, and selects behavior" (McClelland et al., 1953, p. 183)—that may operate nonconsciously and yet are reflected in the linguistic behavior of individuals. These are motives for achievement, affiliation, and power. Achievement motivation is characterized by a preoccupation with excellence and the attainment of success. Affiliation motivation is focused on the creation and maintenance of close, warm interpersonal relationships. Power motivation involves a desire for influence and control over others (Duncan & Peterson, 2010; Winter, 1991).

Theory and research suggest that there is a connection between at least two of the three dispositional motives and political ideology. Utilizing manual means of content coding of public speeches, Suedfeld and colleagues (1990) observed that liberal politicians in Canada used more imagery associated with affiliation motives than did conservative politicians. This is consistent with liberal-socialist ideals, which emphasize social security and communal welfare, as opposed to conservative-capitalist ideals, which emphasize self-reliance and economic competition. The ideological difference in the use of language pertaining to affiliation motivation was replicated by Fetterman, Boyd, and Robinson (2015), who used dictionary-based methods to analyze language used by ordinary citizens (in political chat rooms) and political elites (on news websites and in official political addresses). These researchers also observed that conservatives used more language associated with power than did liberals. This difference is consistent with other evidence indicating that conservatives exhibit higher levels of personal and social dominance (Pratto, Sidanius, Stallworth, & Malle, 1994; Sidanius & Pratto, 1993) and stronger desires to preserve existing power structures (Altemeyer, 1981; Wilson & Sibley, 2013), in comparison with liberals. We therefore hypothesized that liberals would use more affiliation language than conservatives (H1), whereas conservatives would use more power language than liberals (H2).

In some cases, the same theoretical constructs—such as affiliation, power, and achievement—have been operationalized by different researchers of language in different ways (see Appendix A for complete dictionary word lists). In two pairs of dictionaries (i.e., affiliation and power dictionaries from LIWC and Harvard IV), the operationalizations were produced by different groups of researchers who relied on similar dictionary creation algorithms to achieve reasonably comprehensive representations of the constructs of interest. Differences between these dictionaries are minor and result from slightly different methodological decisions made throughout the word vetting process. The three overlapping dictionaries produced by Neiman and colleagues (2016a, 2016b) were developed with more limited goals in mind. Rather than attempting to measure all aspects of a given theoretical construct, these researchers were focused on capturing a single element of that construct as it related to a specific psychological theory. For example, the achievement and power dictionaries were constructed

Table 2
Most Frequently Used Words in Each Language Category

Category	Liberals	Conservatives
Dispositional motives		
Affiliation V1	we, love, our, us, game, help, tweet, team, twitter, **friends**	we, our, love, us, game, help, team, tweet, **family**, twitter
Affiliation V2	like, get, love, our, go, us, make, back, way, thank	like, get, our, love, us, go, back, make, way, thank
Power V1	up, best, over, help, down, win, big, god, **police**, president	up, god, over, best, win, down, help, big, president, **law**
Power V2	have, make, back, right, still, way, take, please, **show**, **say**	have, back, make, right, **god**, way, still, take, **win**, please
Personal values		
Conformity	submission, go along, conform	submission, go along, conform
Power	power, order, control, rule, position, status, charge, capital, manage, authority	power, control, order, rule, position, charge, status, capital, authority, manage
Security	protect, safety, defense, threat, strength, defend, guard, shelter, secure, shield	protect, defense, safety, threat, strength, defend, guard, secure, shelter, shield
Tradition	family, continue, faith, religion, foundation, ceremony, custom, **routine**, belief, tradition	family, faith, continue, religion, foundation, custom, belief, tradition, ceremony, **relations**
Achievement V1	best, first, work, better, win, team, top, trying, try, **working**	best, first, work, win, better, team, top, **leader**, trying, try
Achievement V2	win, job, goal, success, victory, achievement, succeed, finishing, prosperity, **triumph**	win, job, goal, success, victory, achievement, succeed, finishing, prosperity, **accomplishment**
Benevolence	improve, benefit, care for, assist, enhance, nurture, lend a hand	improve, benefit, assist, care for, enhance, nurture, lend a hand
Universalism	understanding, tolerant, open-mindedness, universalism	understanding, tolerant, open-mindedness, universalism
Stimulation	challenge, excitement, enthusiasm, thrill, stimulation	challenge, excitement, enthusiasm, thrill, stimulation
Self-direction	think, mind, create, choose, consider, decide, awareness, strength, purpose, **innovation**	think, mind, choose, create, strength, consider, decide, purpose, **seek**, awareness
Social	rt, you, your, we, love, our, they, people, who, he	rt, you, your, we, he, they, our, who, people, love
Motivated social cognition		
Certainty V1	all, never, ever, every, always, everyone, sure, must, **everything**, true	all, never, ever, every, always, sure, everyone, must, **nothing**, true
Certainty V2	doubt, confusion, disorder, uncertainty, underlying, hesitation, ambiguity, indecision, vagueness, improbability	doubt, confusion, disorder, uncertainty, underlying, hesitation, indecision, ambiguity, vagueness, improbability
Resistance to change	keep, hold, protect, continue, original, normal, foundation, native, prevent, defeat	keep, hold, protect, continue, foundation, original, normal, defeat, prevent, native
Inequality	class, power, law, group, system, lead, order, rule, **privilege**, advantage	law, power, class, lead, group, order, system, rule, advantage, **authority**
Past focus	was, been, got, had, did, were, made, didn't, done, said	was, been, got, did, had, were, made, said, didn't, done
Tentative	if, or, some, most, may, any, hope, someone, something, **pretty**	if, or, some, most, may, any, hope, someone, something, **anyone**
Future focus	will, going, tonight, then, may, hope, tomorrow, I'll, coming, **gonna**	will, going, tonight, then, may, hope, tomorrow, coming, I'll, **won't**
Anxiety	terror, threat, fear, shame, risk, stress, avoid, **struggle**, worry, doubt	terror, threat, fear, risk, doubt, shame, worry, avoid, **afraid**, stress
Inhibition	stop, wait, keep, safe, hold, forget, **waiting**, save, protect, control	keep, stop, wait, safe, hold, control, protect, forget, save, **secure**
Threat	fear, hurt, loss, risk, avoid, upset, warning, threat, failure, terror	fear, hurt, loss, warning, threat, terror, risk, failure, avoid, upset
Risk focus	stop, bad, wrong, problem, fail, protect, **worst**, safe, lose, secure	stop, bad, problem, wrong, fail, protect, secure, lose, **threat**, safe
Death	kill, death, dead, war, die, murder, died, alive, **wars**, **dying**	kill, war, death, dead, murder, die, died, alive, **memorial**, **suicide**
Uniqueness and conformity motivations		
I	I, my, me, I'm, I've, I'll, I'd, I'm, myself, mine	I, my, me, I'm, I've, I'll, I'm, I'd, myself, mine
We	we, our, us, were, lets, well, we've, let's, wed, ourselves	we, our, us, lets, were, let's, well, we've, ourselves, wed
Emotion		
Positive emotion	love, good, great, thanks, happy, best, **lol**, thank, please, well	good, love, great, thanks, best, happy, thank, please, **win**, well
Negative emotion	kill, bad, shit, miss, hate, fight, **fuck**, wrong, problem, **sorry**	kill, bad, fight, miss, hate, **attack**, shit, problem, **terror**, wrong
Anger	kill, shit, hate, fight, fuck, damn, **fuckin**, attack, hell, war	kill, fight, hate, attack, shit, war, hell, damn, **murder**, fuck
Anxiety	terror, threat, fear, shame, risk, stress, avoid, **struggle**, worry, doubt	terror, threat, fear, risk, doubt, shame, worry, avoid, **afraid**, stress

Note. Words are listed in the order of frequency, so the most frequently used words are listed first. Liberal and conservative tweets were distinguished at the ordinary citizen level by taking one standard deviation around the mean. Anyone who was above this cutoff was coded as conservative (4,670 Twitter users), and anyone who fell below this cutoff was coded as liberal (7,153 Twitter users). Words that were differentially used by liberals and conservatives are shown in boldface.

to measure these value orientations as described by Schwartz's (1992) theory of basic values, which we describe below. Thus, the Neiman dictionaries take into account the specific aspects of each underlying construct that are directly related to more than one relevant theory and are therefore discussed in more than one section of the article.[4]

Results for Dispositional Motives

Liberal–conservative political orientation. To examine hypotheses concerning the motivations of liberals and conservatives, we used a mix of dictionaries taken from LIWC and the Harvard IV (Fetterman et al., 2015; see Figure 1). Consistent with H2, conservatives used more language on Twitter referencing power than liberals (V1: OR = 1.069, 99% CI [1.057, 1.082]; V2: OR = 1.066, 99% CI [1.053, 1.079]). However, they were also more likely to reference affiliation, but only with respect to one of the

[4] As a rule, the dictionaries constructed by Neiman and colleagues were focused on psychological constructs as they pertain to specific theories. The dictionaries included in LIWC and Harvard IV, on the other hand, were constructed to measure more general definitions of those psychological constructs.

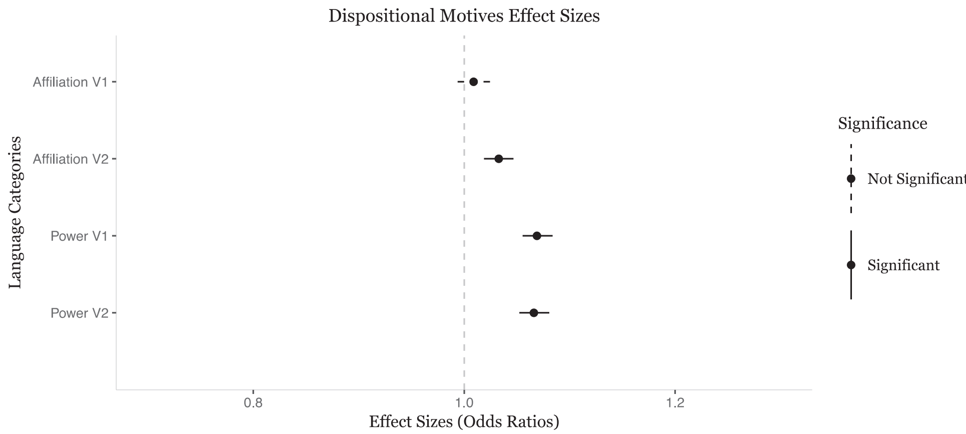

Figure 1. Odds ratio estimates of the effect of political ideology on the dispositional motives language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

two operationalizations of this construct (V2: OR = 1.033, 99% CI [1.020, 1.046]). Thus, H1 was not supported.

The word-level analyses helped to elucidate these results with respect to affiliation. While some frequently used words appeared in both categories (e.g., “we,” “love,” “our,” “us”), there were some notable differences. LIWC’s version of the dictionary included unique terms that referred to social media (“tweet,” “twitter”) and sports or games (“game,” “team”). The Harvard dictionary, by contrast, includes unique words that seem action-oriented (“get,” “make,” “go”), apparently tapping into approach motivation in a way that the LIWC dictionary does not. The inclusion of approach-oriented words may help to explain why we find ideological differences in the Harvard affiliation dictionary but not in LIWC.

The most frequently used words in the Harvard version of the power dictionary differed between liberals and conservatives. Specifically, words like “show” and “say” appeared in the top word list for liberals but not conservatives, whereas words such as “god” and “win” appeared in the list for conservatives but not liberals. A post hoc interpretation is that whereas liberals focus on the means of communication (which may or may not be directly linked to power in every case), conservatives emphasize established authorities and power-related outcomes (see Table 2).

The role of ideological extremity. To weigh the evidence in support of the J-curve hypothesis, which combines linear and quadratic effects of political ideology, we present the results from multilevel mixed-effects logistic regression models and compare how the two regression-based prediction models (one that includes the quadratic term and one that does not) fit the observed patterns of the data in each case using enrichment analyses.

We found that ideological extremity was associated with the use of words from three of the four dispositional motive dictionaries (see Figure 2). Having a more extreme ideological position was associated with sending more tweets referencing power (V1: OR = 1.026, 99% CI [1.013, 1.038]; V2: OR = 1.029, 99% CI [1.020, 1.038]), and (in one of two cases) affiliation (V1: OR = 0.997, 99% CI [0.986, 1.007]; V2: OR = 1.016, 99% CI [1.004, 1.029]). Integrating results from the two regression models into the nonparametric analyses allows us to better evaluate evidence bearing on the J-curve hypothesis. In Figure 3 we have plotted the ratio of the observed proportion of power language for each of the 50 bins over the observed proportion of power language throughout the dataset as a whole. Because the bins are arranged according to the ideology of the Twitter users (along the *x* axis), these points can be utilized as a smoothed scatterplot representing patterns within the data. To determine how well the regression analyses are performing, we fit predicted lines from each model onto these plots; thus, linear effects are depicted as dotted lines, and quadratic effects are depicted as dashed lines. We also depict a moving average of the enrichment ratios as solid lines. Comparing the moving average to the two relationships specified by the parametric models confirms support for the J-curve hypothesis with respect to power and (in one case) affiliation (see Figure 4).

Personal Values: The Schwartz Taxonomy

Personal values, which are closely related to dispositional motives in the tradition of Atkinson (1958), McClelland (1987), and Winter (1991), are defined as the criteria by which individuals select and evaluate behaviors, events, objects, and ideas (Schwartz, 1992). They are centered on specific goals or motivational concerns that were presumably adaptive in an evolutionary sense (Schwartz & Bilsky, 1987, 1990). According to Schwartz (1992), there are 10 basic value dimensions that are widely shared across cultures, namely: universalism, benevolence, conformity, tradition, security, power, achievement, hedonism, stimulation, and self-direction. In terms of self-report measures completed by ordinary citizens, endorsement of these values is clearly linked to political orientation. Specifically, liberals and leftists tend to prioritize universalism, benevolence, stimulation, and self-direction, whereas conservatives and rightists tend to prioritize conformity, tradition, security, and power (Caprara, Schwartz, Capanna, Vec-

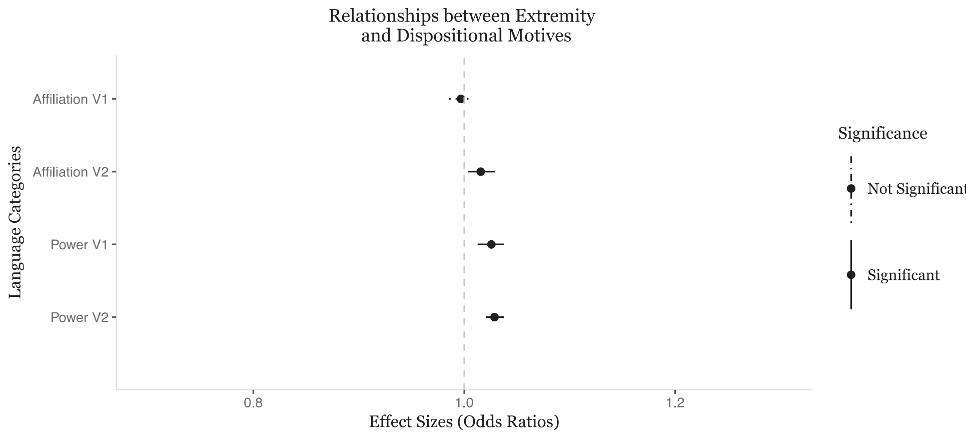

Figure 2. Odds ratio estimates of the quadratic effect of political ideology on the dispositional motives language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

chione, & Barbaranelli, 2006; Devos, Spini, & Schwartz, 2002; Jost, Basevich, Dickson, & Noorbaloochi, 2016; Piurko, Schwartz, & Davidov, 2011; Schwartz, Caprara, & Vecchione, 2010; Vecchione, Caprara, Dentale, & Schwartz, 2013).

Consistent with these results, two studies of social media usage suggested that Democratic elites were more likely to use words related to benevolence, universalism, and social security, whereas Republican elites were more likely to use words related to power, conformity, tradition, and national security (Jones et al., 2018; Neiman et al., 2016a, 2016b). Thus, we hypothesized that conservatives would use more language expressing power (H2), conformity (H3), security (H4), tradition (H5), and achievement (H6) than liberals, and that liberals would use more language expressing benevolence (H7), universalism (H8), stimulation (H9), and self-direction (H10) than conservatives. In line with their values of benevolence and universalism, we predicted that liberals will also use language that refers to others or social words (H11) than conservatives.

Results for Personal Values

Liberal–conservative political orientation. To assess the degree to which liberals and conservatives differed in their emphasis on specific personal values, we utilized dictionaries developed by Neiman and colleagues (2016a, 2016b). Results are summarized in Figure 5. Consistent with Schwartz's (1992) theory of values, conservatives were more likely than liberals to send tweets that referenced power (H2; OR = 1.064, 99% CI [1.038, 1.096]), security (H4; OR = 1.152, 99% CI [1.097, 1.199]), tradition (H5; OR = 1.090, 99% CI [1.061, 1.126]), and achievement (H6; V1: OR = 1.030, 99% CI [1.017, 1.045]; V2: OR = 1.105, 99% CI [1.074, 1.135])–as well as social words (H11; OR = 1.040, 99% CI [1.022, 1.063]). Liberals were more likely than conservatives to send tweets that referenced benevolence (H7; OR = 0.917, 99% CI [0.852, 0.995]). There were no ideological differences in the use of the four remaining categories of values. However, it is important to point out that some of these categories included very

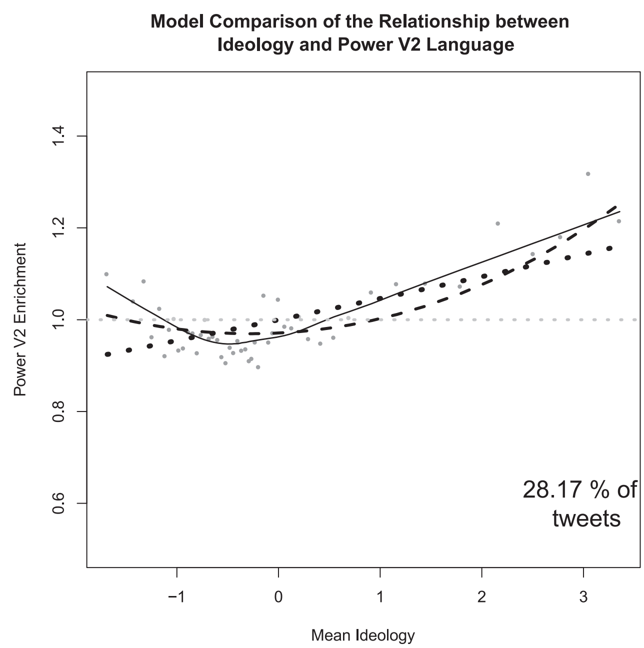

Figure 3. Model comparisons for power language. Tweets were ordered by political ideology and then the dataset was split into 50 equal segments (each segment contains 240,000 tweets). Then, enrichment analyses were conducted separately for each segment to test whether tweets within that segment were significantly more or less likely to contain this specific type of language than we would expect on average. Significance was determined using a multiple hypothesis corrected alpha (α = .0037). Dark gray dots represent groups of tweets that significantly differed from the average proportion. Light gray dots represent groups of tweets did not significantly differ from the average proportion. Solid lines are lowess lines fit to the results of the enrichment analyses (taking into account 66% of the data at a time). The dotted line represents the expected values implied by the linear regression model. The dashed line represents the expected values implied by the quadratic regression model (see Appendix B for details).

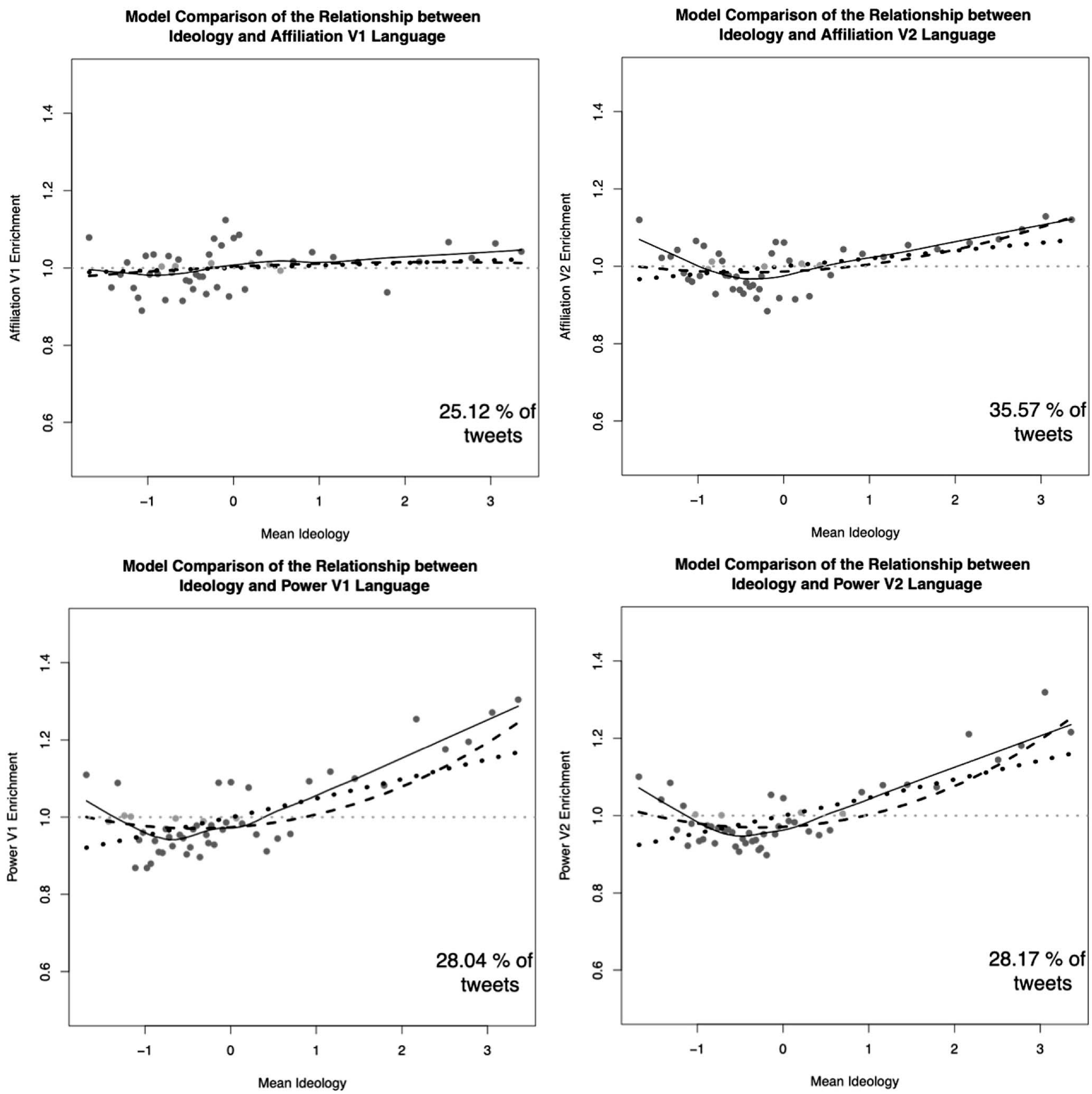

Figure 4. Model comparisons for dispositional motives variables. Tweets were ordered by political ideology and then the dataset was split into 50 equal segments (each segment contains 240,000 tweets). Then, enrichment analyses were conducted separately for each segment to test whether tweets within that segment were significantly more or less likely to contain this specific type of language than we would expect on average. Significance was determined using a multiple hypothesis corrected alpha ($\alpha = .0037$). Dark gray dots represent groups of tweets that significantly differed from the average proportion. Light gray dots represent groups of tweets did not significantly differ from the average proportion. Solid lines are lowess lines fit to the results of the enrichment analyses (taking into account 66% of the data at a time). The dotted line represents the expected values implied by the linear regression model. The dashed line represents the expected values implied by the quadratic regression model (see Appendix B for details).

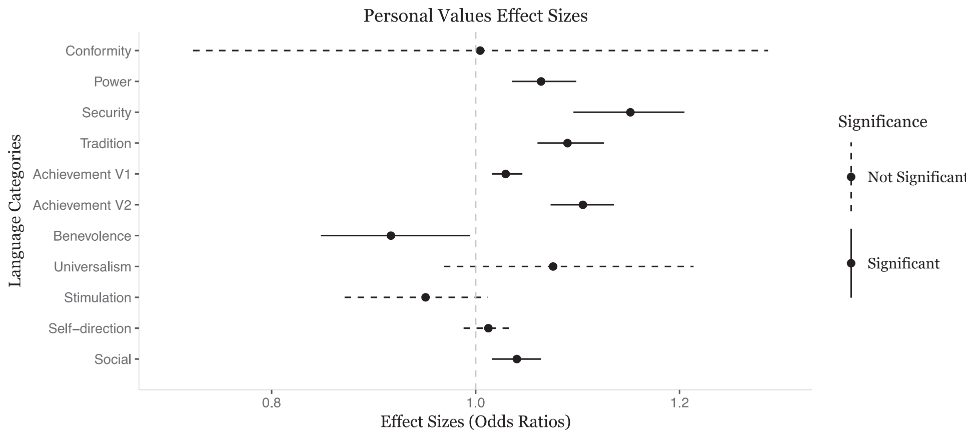

Figure 5. Odds ratio estimates of the effect of political ideology on the personal values language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

few words (i.e., less than 10). Presumably, this limits conclusions we can draw from these analyses within the domain of personal values.

The role of ideological extremity. We observed significant quadratic relationships with respect to four of the 11 personal values dictionaries (see Figure 6). Extremists were more likely than moderates to send messages containing language pertaining to security (OR = 1.061, 99% CI [1.030, 1.098]) and self-direction (OR = 1.028, 99% CI [1.010, 1.050]), whereas moderates were more likely to send messages containing benevolence words (OR = 0.937, 99% CI [0.879, 0.997]). There was no relationship between extremity per se and the use of the other seven language categories. We obtained support for the J-curve hypothesis in the cases of security and (to a marginal extent) social language (see Figure 7). In addition, we found that liberals and moderates used more benevolence language than did conservatives and extremists.

Theory of Ideology as Motivated Social Cognition

According to the theory of political conservatism as motivated social cognition, epistemic needs to achieve certainty, order, and structure and existential needs to reduce fear, threat, and insecurity promote an affinity for the maintenance of tradition and hierarchy, both of which are enshrined in conservative ideology (Jost et al., 2003b). Consistent with this formulation, individuals who score higher on subjective and objective measures of uncertainty avoidance, intolerance of ambiguity, cognitive/perceptual rigidity, dogmatism, and personal needs for order, structure, and cognitive closure tend to be more politically conservative, whereas individuals who score higher on subjective and objective measures of integrative complexity, need for cognition, and cognitive reflection tend to be more liberal (Jost, 2017a; Jost, Sterling, & Stern, 2018;

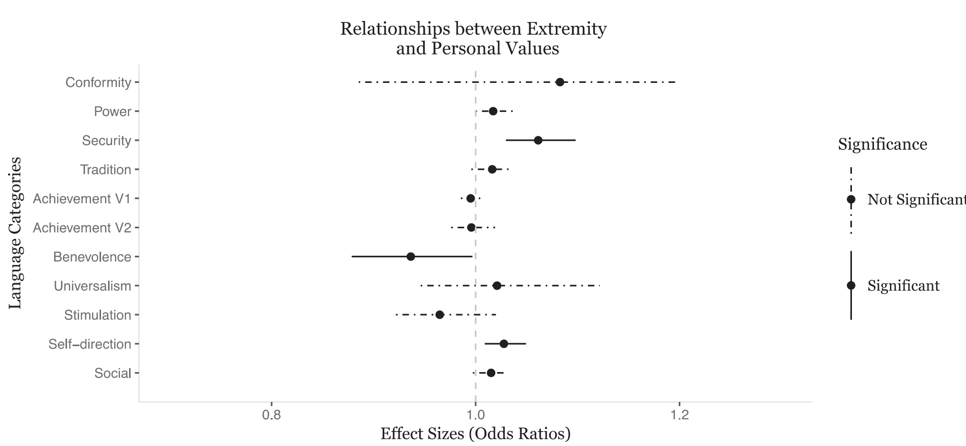

Figure 6. Odds ratio estimates of the quadratic effect of political ideology on the personal values language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

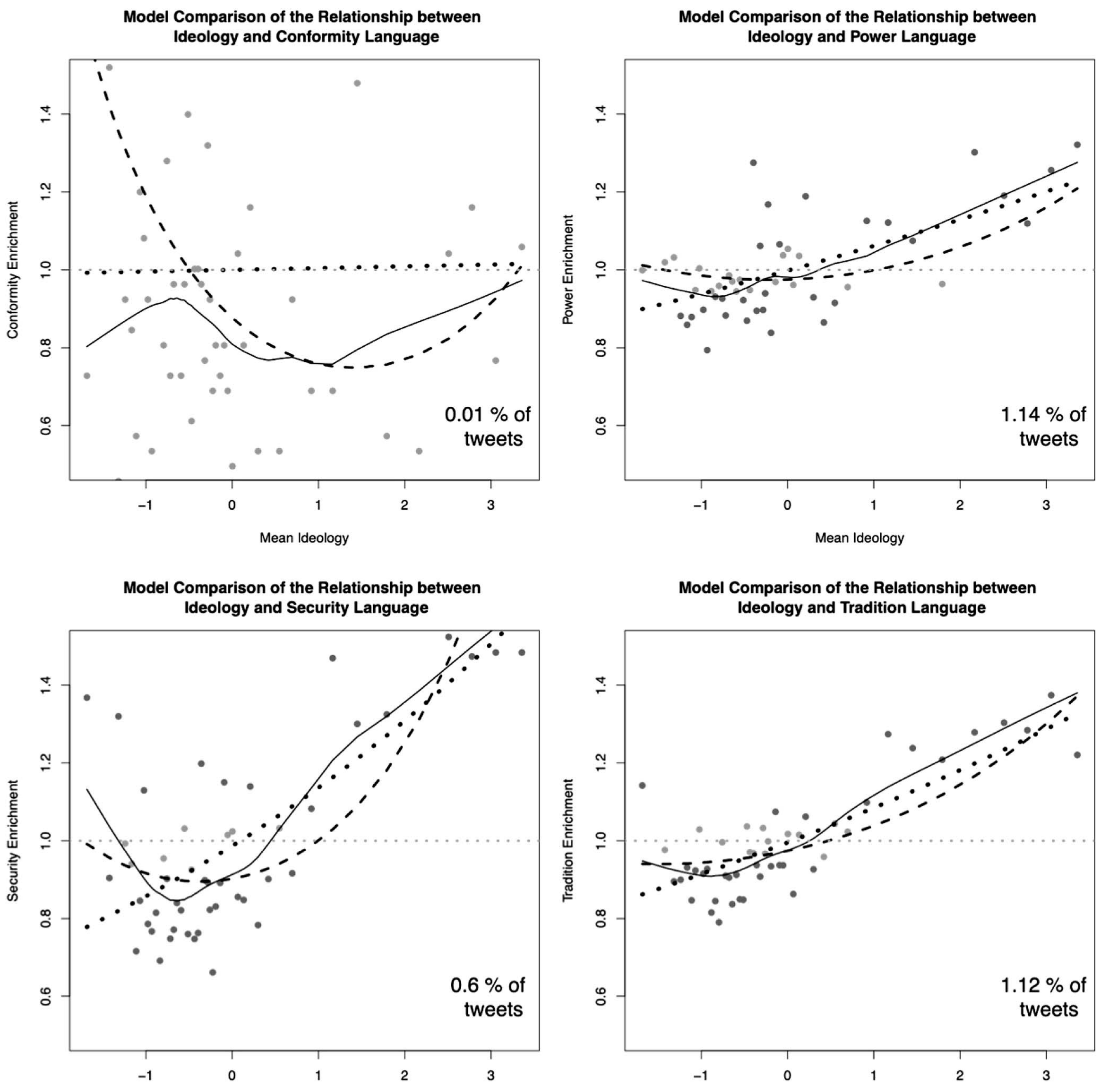

Figure 7. Model comparisons for personal values variables. Tweets were ordered by political ideology, and then the dataset was split into 50 equal segments (each segment contains 240,000 tweets). Then, enrichment analyses were conducted separately for each segment to test whether tweets within that segment were significantly more or less likely to contain this specific type of language than we would expect on average. Significance was determined using a multiple hypothesis corrected alpha (α = .0037). Dark gray dots represent groups of tweets that significantly differed from the average proportion. Light gray dots represent groups of tweets did not significantly differ from the average proportion. Solid lines are lowess lines fit to the results of the enrichment analyses (taking into account 66% of the data at a time). The dotted line represents the expected values implied by the linear regression model. The dashed line represents the expected values implied by the quadratic regression model (see Appendix B for details). (*Figure continues on next page.*)

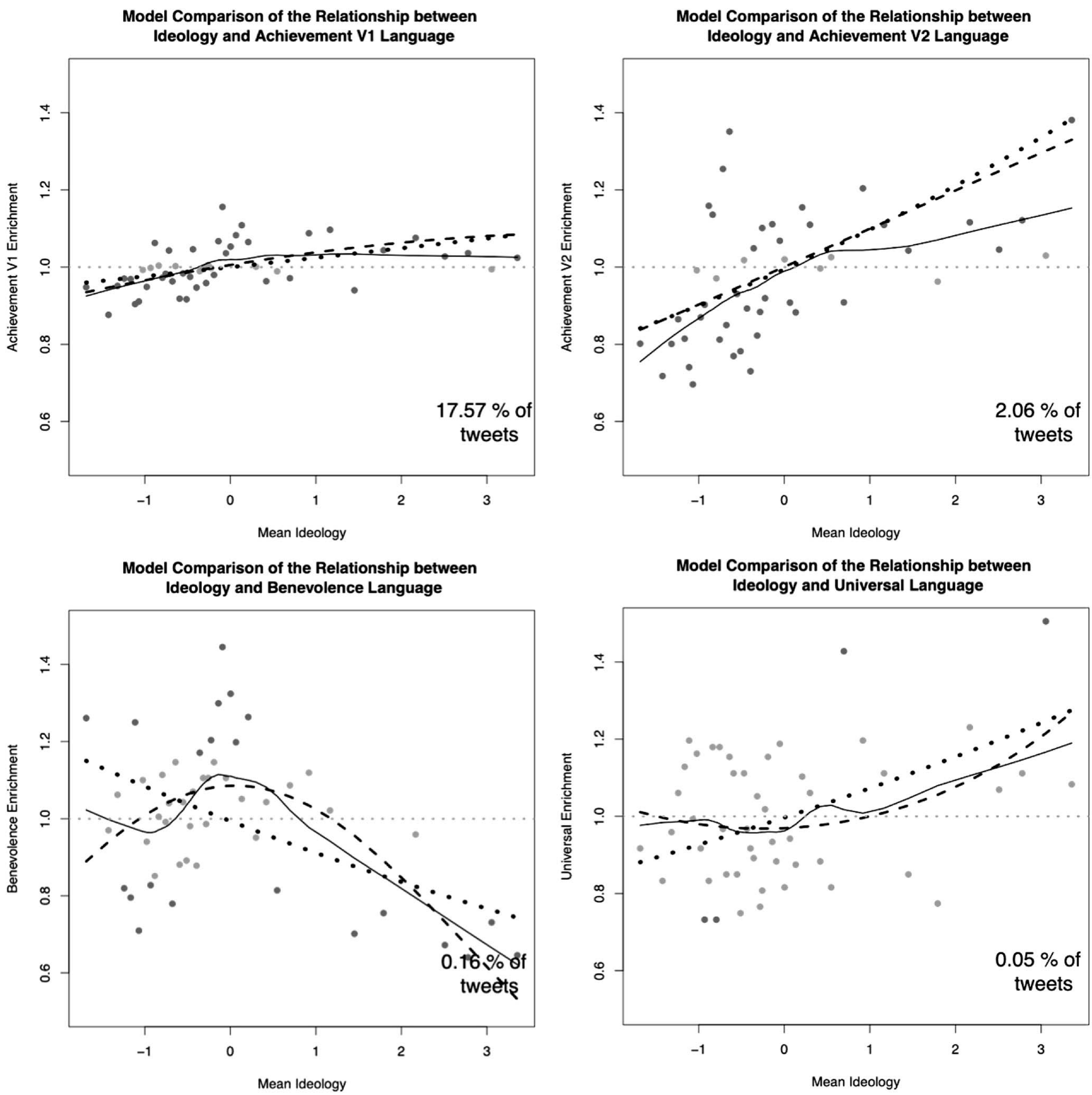

Figure 7. (continued) (*Figure continues on next page.*)

Rokeach, 1960). In addition, subjective perceptions of threat and exposure to objectively threatening circumstances, such as terrorist attacks, are positively associated with preferences for conservative leaders, opinions, and policies (and negatively associated with preferences for liberal leaders, opinions, and policies; Jost, 2017a; Jost, Stern, Rule, & Sterling, 2017; Wilson, 1973).

In terms of language use, several studies show that speeches and opinions given by liberals exhibit more integrative complexity (manually coded in terms of the number of arguments raised and the extent to which those arguments were reconciled) than those given by conservatives (Brundidge et al., 2014; Kemmelmeier, 2008; Tetlock, 1983; Tetlock et al., 1985). An analysis of State of the Union addresses revealed that Republican presidents used a higher proportion of nouns, which elicit clearer and more stable and definite perceptions of reality than other parts of speech such as verbs and adjectives, than did Democratic presidents (Cichocka et al., 2016). There is also some evidence that conservatives use more anxiety (Robinson et al., 2014) and threat-related language (Neiman et al., 2016b) as well as more resistance to change (Okdie & Rempala, 2019) and references to the past (Robinson, Cassidy,

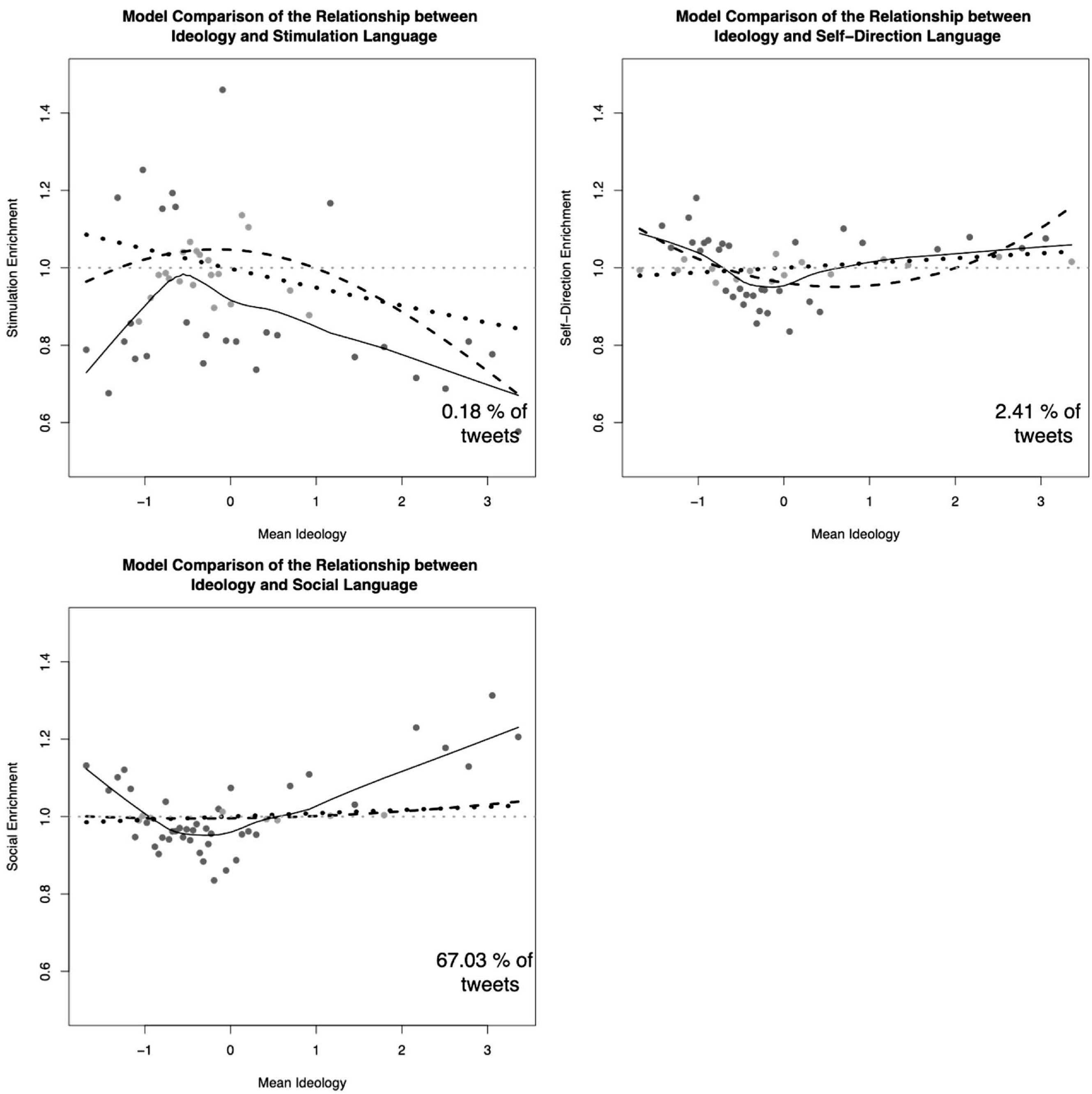

Figure 7. (continued)

Boyd, & Fetterman, 2015). For all of these reasons, we hypothesized that conservatives would use more language indicating certainty (H12), resistance to change (H13), acceptance of inequality (H14), and a focus on the past (H15), whereas liberals would use more language indicating tentativeness (H16) and a focus on the future (H17). In addition, we hypothesized that conservatives would use more language conveying anxiety (H18), inhibition (H19), threat sensitivity (H20), risk focus (H21), and death salience (H22).

Results for the Theory of Ideology as Motivated Social Cognition

Liberal–conservative political orientation. We obtained support for hypotheses derived from Jost et al.'s (2003b) theory of ideology as motivated social cognition in seven of 12 cases (see Figure 8). Conservatives were more likely than liberals to send messages that contained language reflecting certainty (H12; V1: OR = 1.031, 99% CI [1.015, 1.045]; V2: OR = 1.085, 99% CI

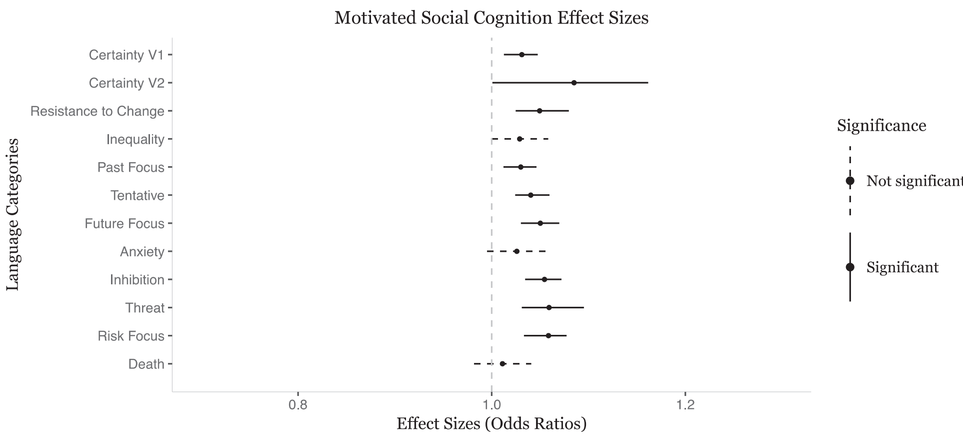

Figure 8. Odds ratio estimates of the effect of political ideology on the motivated social cognition language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

[1.001, 1.157]), resistance to change (H13; OR = 1.050, 99% CI [1.027, 1.075]), references to the past (H15; OR = 1.030, 99% CI [1.016, 1.046]), inhibition (H19; OR = 1.055, 99% CI [1.038, 1.070]), threat (H21; OR = 1.059, 99% CI [1.033, 1.089]), and risk focus (H21; OR = 1.059, 99% CI [1.035, 1.077]). Unexpectedly, conservatives were also more likely than liberals to use tentative language (H16; OR = 1.040, 99% CI [1.026, 1.057]) and to refer to the future (H17; OR = 1.050, 99% CI [1.035, 1.066]). There were no ideological differences in the use of the other three language categories. The top words in each category were very similar for liberals and conservatives (see Table 2).

The role of ideological extremity. We observed significant quadratic variation for 10 of the 12 motivated social cognition dictionaries (see Figure 9). Extremists were more likely than moderates to send messages containing language pertaining to anxiety (OR = 1.044, 99% CI [1.027, 1.060]), certainty (but only significantly for V1: OR = 1.033, 99% CI [1.018, 1.079]; V2: OR = 1.029, 99% CI [1.001, 1.157]), death (OR = 1.055, 99% CI [1.038, 1.070]), future (OR = 1.030, 99% CI [0.973, 1.079]), inhibition (OR = 1.046, 99% CI [1.033, 1.057]), past (OR = 1.047, 99% CI [1.035, 1.060]), resistance to change (OR = 1.024, 99% CI [1.004, 1.041]), risk (OR = 1.041, 99% CI [1.028, 1.054]), and tentativeness (OR = 1.043, 99% CI [1.032, 1.057]).

We obtained partial support (or better) for the J-curve hypothesis with respect to eight of the 12 dictionaries, namely: anxiety, certainty, inhibition, past, resistance to change, risk, and tentativeness (see Figure 10). For each of these categories, conservatives used more language of that type than liberals, but extreme liberals also used these linguistic categories more than moderate liberals. Threat-related language was the only one that followed a strictly linear pattern. However, death-related language followed a purely quadratic or symmetrical pattern; extremists mentioned death more than moderates.[5]

Needs for Uniqueness and Conformity

It has been suggested that, in addition to epistemic and existential needs to reduce uncertainty and threat, there are relational needs to attain a sense of social belongingness and shared reality with like-minded others that promote an affinity for conservative ideology (Hennes, Nam, Stern, & Jost, 2012; Jost, Ledgerwood, & Hardin, 2008; Jost, van der Linden, Panagopoulos, & Hardin, 2018). Consistent with this notion, conservatives place a higher value than liberals on conformity, loyalty, and adherence to convention (e.g., Jost et al., 2016). They also score higher than liberals on subjective measures of the need to share reality among like-minded others (Hennes et al., 2012) and perceive more within-group consensus, even when it comes to judgments that have no political relevance (Stern, West, Jost, & Rule, 2014). Liberals, on the other hand, have been found to exhibit an “illusion of uniqueness,” perceiving their own judgments to be more distinctive than they actually are (Rabinowitz, Latella, Stern, & Jost, 2016; Stern, West, & Schmitt, 2014).

With respect to the use of language, Sylwester and Purver (2015) assessed desires for uniqueness and conformity in terms of pronoun use. They argued that greater use of first person singular pronouns (“I” words) would reflect a stronger desire for uniqueness, whereas greater use of first person plural pronouns (“we” words) would reflect perceptions of within-group similarity and a stronger desire for conformity. As predicted, these researchers observed that supporters of the Democratic Party used more first-person singular pronouns in their social media messages, whereas supporters of the Republican Party used more first person plural pronouns. Building on this work, we hypothesized that liberals would use more first person singular pronouns (H23), whereas conservatives would use more first person plural pronouns (H24) than liberals.

[5] Enrichment analyses suggested that neither the linear nor quadratic relationships best characterized the relationship between ideology and future-related language. Although there was a significant quadratic effect of ideology on future language in the regression models, upon further examination we saw that there appeared to be a positive relationship between ideology and future language for conservatives but little to no relationship for liberals.

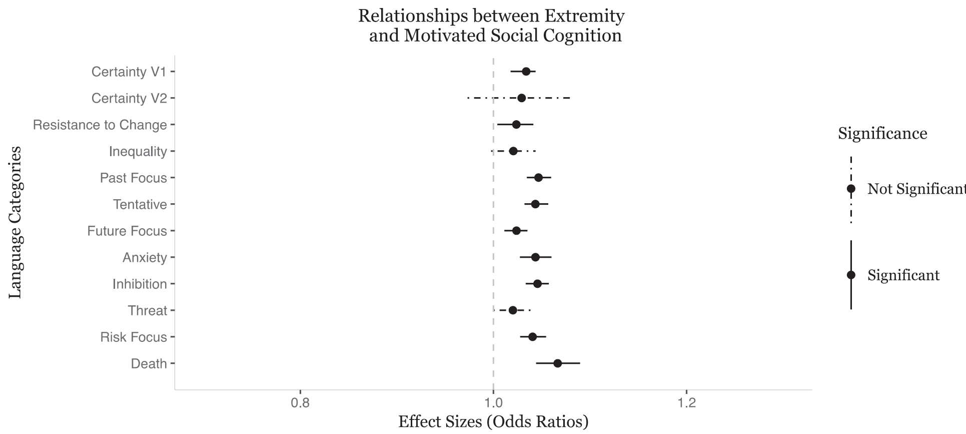

Figure 9. Odds ratio estimates of the quadratic effect of political ideology on the motivated social cognition language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

Results for Uniqueness and Conformity

Liberal–conservative political orientation. We utilized two categories from LIWC to examine the expression of needs pertaining to uniqueness and conformity (see Figure 11). Consistent with Stern et al.'s (2014) hypothesis concerning preferences for conformity and ingroup similarity (H24), conservatives used more first person plural pronouns ("we;" OR = 1.045, 99% CI [1.028, 1.066]) than liberals. There were no differences in the use of "I" (H23).

The role of ideological extremity. Ideological extremity predicted the use of language pertaining to uniqueness and conformity (see Figure 12). That is, extremists were more likely than moderates to use both singular (OR = 1.035, 99% CI [1.015, 1.058]) and plural (OR = 1.041, 99% CI [1.025, 1.059]) personal pronouns. We observed a J-curve pattern with respect to the use of first person plural pronouns (see Figure 13). With respect to the use of "I" words, we observed that increased liberalism was associated with the use of more first person singular pronouns; there was no apparent relationship for conservatives.

Emotional Expression

Insofar as conservative ideology offers explanations and justifications for social and economic inequality that exonerate the overarching social system, Napier and Jost (2008) hypothesized (and found) that conservatives and rightists in North America and Europe would be less troubled than liberals and leftists by the high degree of inequality in society. Specifically, they found that conservatives reported more life satisfaction and greater subjective well-being than liberals, and that this effect was partially attributable to the belief that inequality in society is fair and just. The ideological gap in self-reported happiness has been replicated many times (Bixter, 2015; Burton, Plaks, & Peterson, 2015; Butz, Kieslich, & Bless, 2017; Choma, Busseri, & Sadava, 2009; Cichocka & Jost, 2014; Okulicz-Kozaryn, Holmes, & Avery, 2014; Onraet, Van Assche, Roets, Haesevoets, & Van Hiel, 2017; Schlenker, Chambers, & Le, 2012). Nevertheless, when Wojcik et al. (2015) investigated the language used by liberal and conservative citizens and legislators, they observed that liberals were more likely than conservatives to use positive emotion words. Thus, Wojcik and colleagues suggested that subjective measures of happiness were distorted by the fact that conservatives tend to score higher on measures of self-deception and impression management. Although the use of positive emotion words is not necessarily a direct measure of subjective or objective well-being, we considered the hypotheses that liberals would use more positive emotion words (H25), whereas conservatives use more negative emotion words (H26).

It has also been theorized that liberals are more strongly motivated than conservatives by approach-related emotions such as moral outrage, anger at injustice, and excitement as well as a promotion-oriented focus on the attainment of positive outcomes, whereas conservatives are more strongly motivated than liberals by avoidance-related emotions such as anxiety, feeling calm, and a prevention-oriented focus on risk management (Janoff-Bulman, Sheikh, & Baldacci, 2008; Pattershall & Eidelman, 2008). Consistent with these ideas, Robinson and colleagues (2014) observed that liberals used more language related to anger than conservatives did, and conservatives used more language related to anxiety than liberals did. At the same time, public opinion research suggests that rightists may be more driven by anger (and fear), in comparison with leftists, so we considered this possibility as well (Jost, 2019; Vasilopoulos, Marcus, Valentino, & Foucault, 2019). Thus, we investigated the hypothesis proposed by Robinson et al. (2014) that liberals would use more anger-related language than conservatives (H27) and (as noted above) that conservatives would use more anxiety-related language than liberals (H18).

Results for Emotional Expression

Liberal–conservative political orientation. To assess differences in emotional expression, we used four LIWC emotion categories (see Figure 14). As hypothesized, conservatives sent more messages containing negative emotion than liberals (H26;

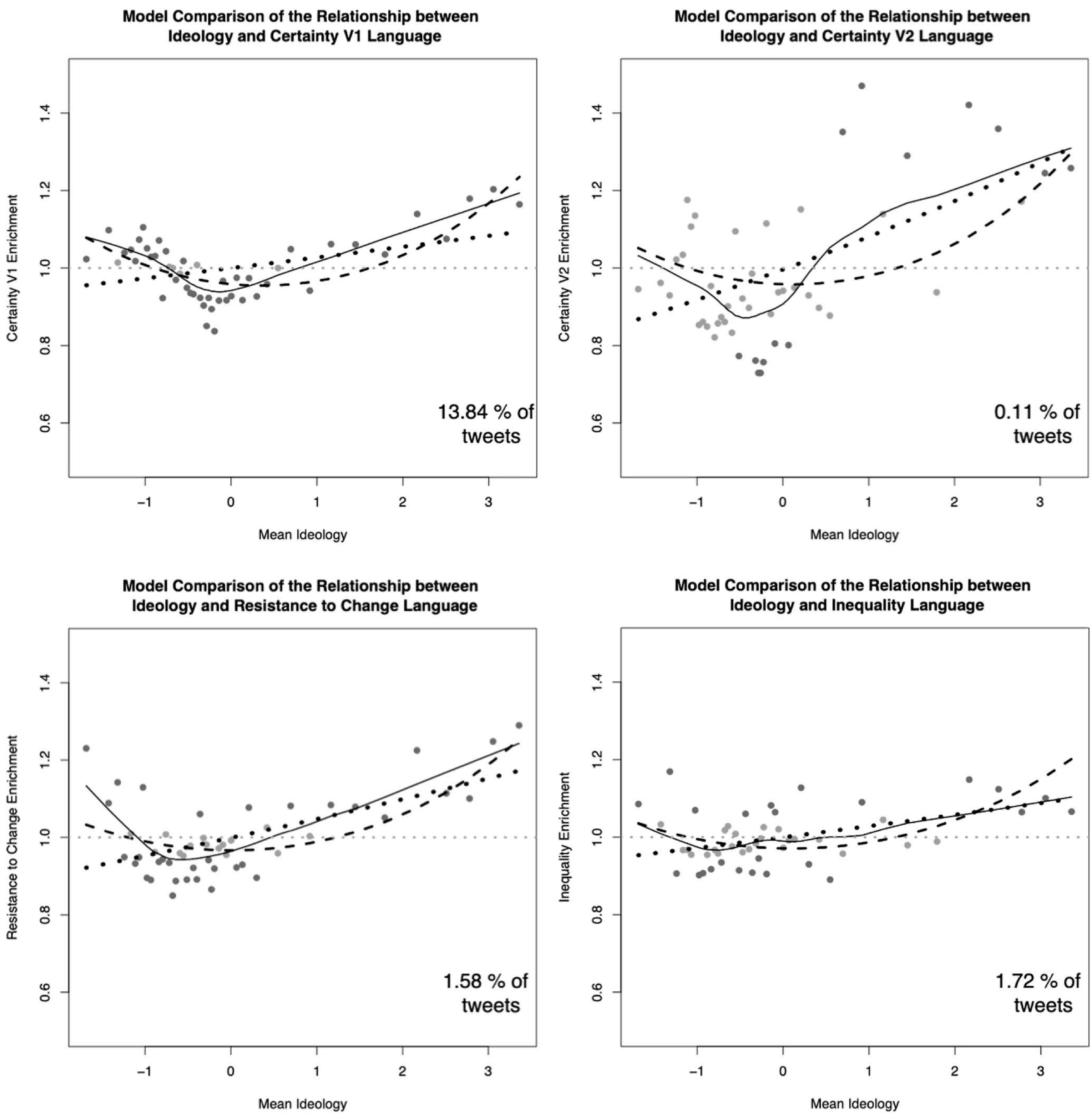

Figure 10. Model comparisons for motivated social cognition variables. Tweets were ordered by political ideology and then the dataset was split into 50 equal segments (each segment contains 240,000 tweets). Then, enrichment analyses were conducted separately for each segment to test whether tweets within that segment were significantly more or less likely to contain this specific type of language than we would expect on average. Significance was determined using a multiple hypothesis corrected alpha ($\alpha = .0037$). Dark gray dots represent groups of tweets that significantly differed from the average proportion. Light gray dots represent groups of tweets did not significantly differ from the average proportion. Solid lines are lowess lines fit to the results of the enrichment analyses (taking into account 66% of the data at a time). The dotted line represents the expected values implied by the linear regression model. The dashed line represents the expected values implied by the quadratic regression model (see Appendix B for details). (*Figure continues on next page.*)

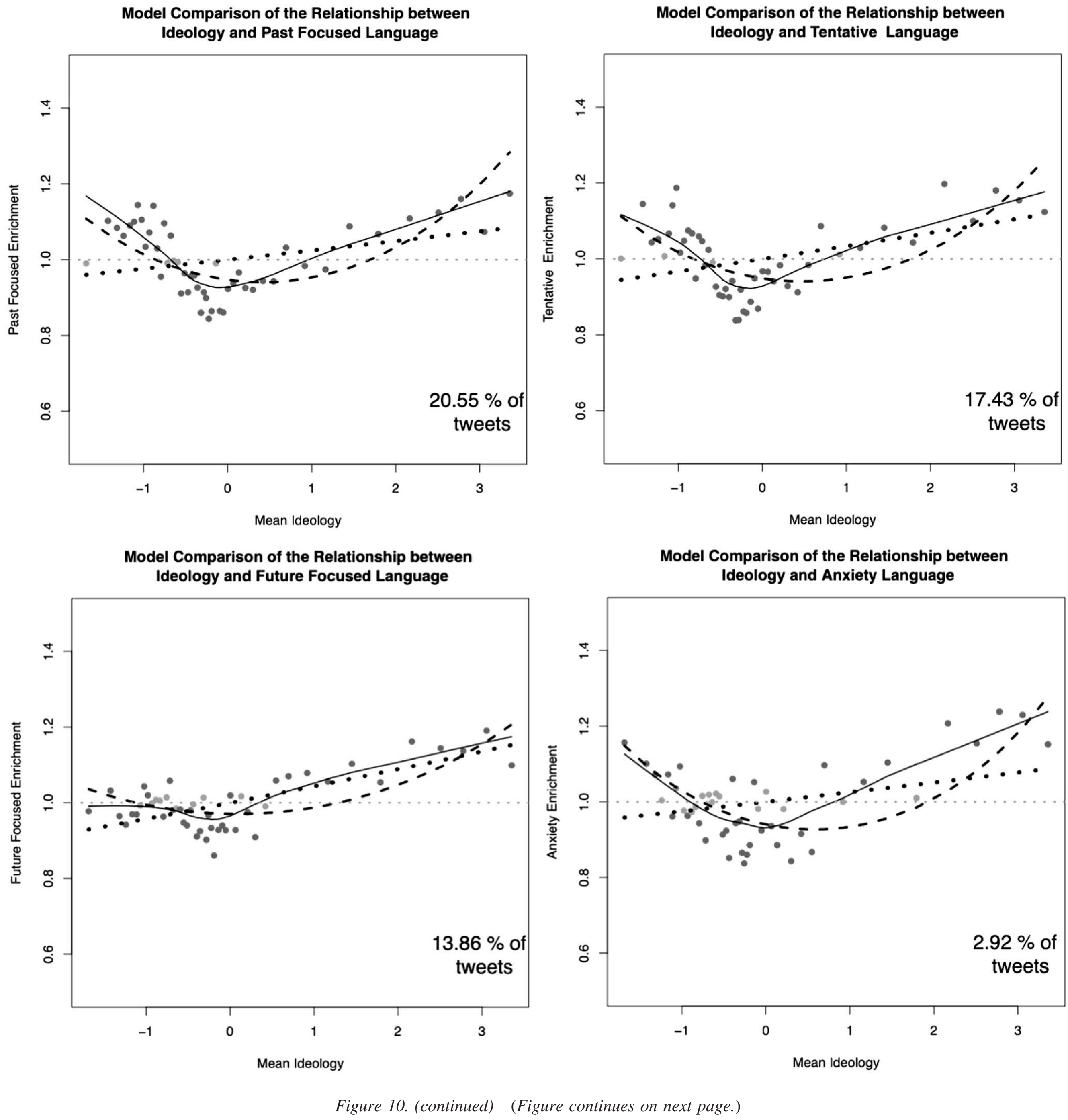

Figure 10. (continued) (*Figure continues on next page.*)

OR = 1.040, 99% CI [1.019, 1.058]). They also sent more tweets with anger-related language, OR = 1.041, 99% CI [1.019, 1.065], consistent with public opinion data suggesting that rightists tend to express more anger than leftists (Jost, 2019; Vasilopoulos et al., 2019). No ideological differences were observed in the use of positive emotion (H25) or anxiety language (H18; as reported above), contrary to expectations. Consistent with the theory of ideology as motivated social cognition, conservatives' top negative emotion words included threat sensitivity ("attack," "terror"), whereas the language of liberals was less likely to include overtly threatening content (e.g., "fuck," "sorry").

The role of ideological extremity. Both extremists and conservatives were more likely than moderates and liberals to use words related to anger (OR = 1.095, 99% CI [1.075, 1.113]), anxiety, and negative emotion in general (OR = 1.066, 99% CI [1.051, 1.075]; see Figures 15 and 16). There were no effects of extremity on the use of words related to positive affect.

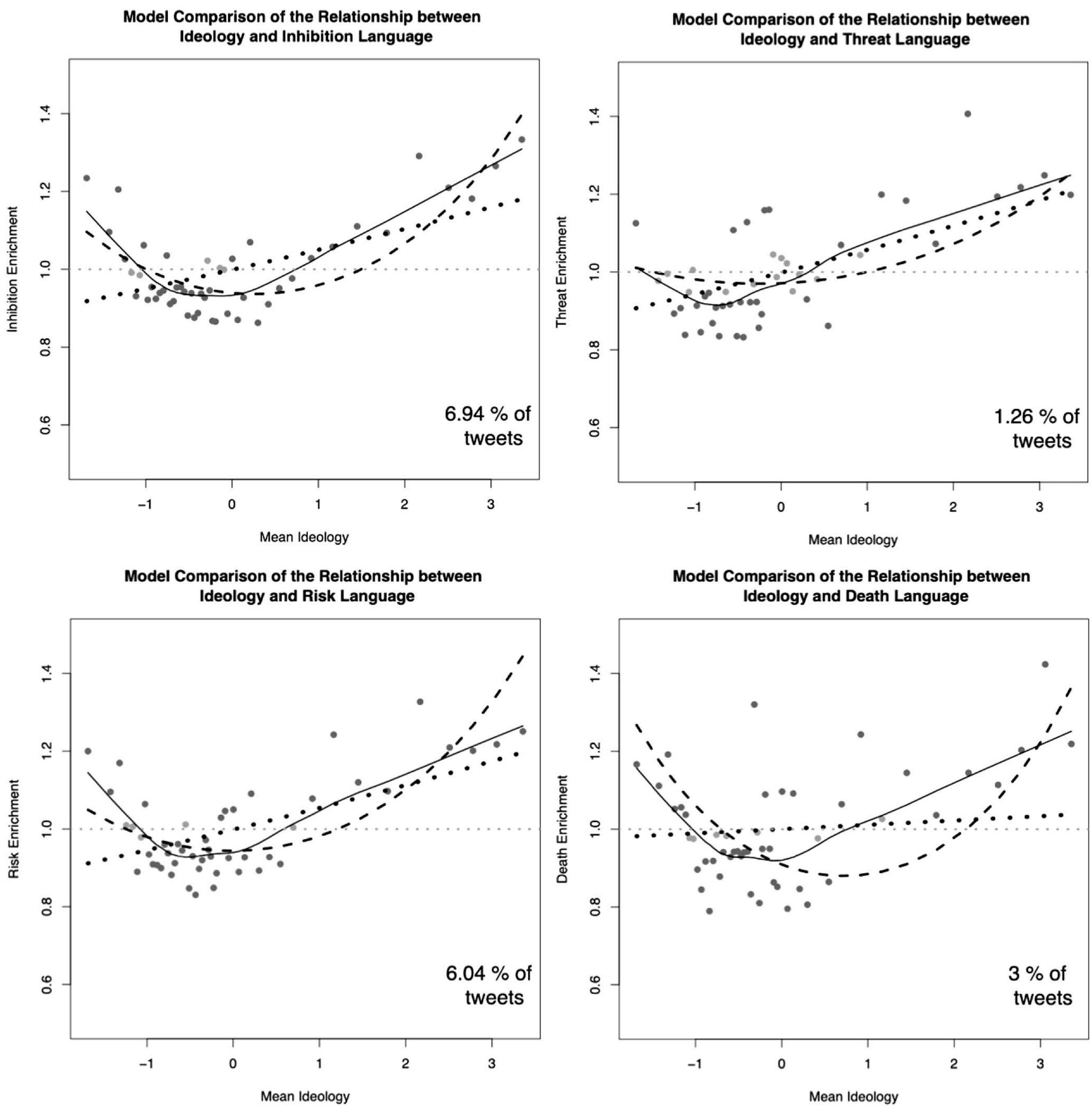

Figure 10. (continued)

Discussion

In this research program we sought to develop a theory-driven, empirically rigorous and comprehensive understanding of the connections between political ideology and the use of language by employing dictionary-based tools for analyzing naturally occurring language as an ecologically valid, unobtrusive measure of needs, values, and motives linked to liberalism and conservatism (see Webb, Campbell, Schwartz, & Sechrest, 1969). The incorporation of naturally occurring language as a complement to traditional experimental and survey-based measures of basic values and psychological motives advances our knowledge of these relationships in a number of exciting ways. For one thing, we were able to harvest data from a much larger and more representative sample than is typically the case in political psychology. In doing so, we bypassed concerns about the generalizability of results based exclusively on university students and instead focused on the population of social media users in the United States (Henry, 2008;

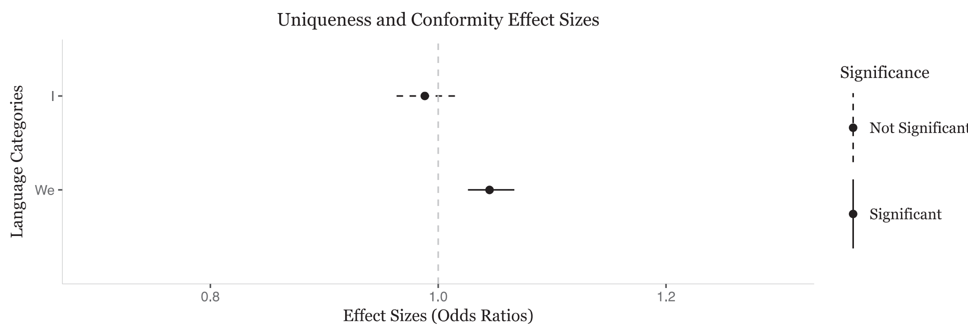

Figure 11. Odds ratio estimates of the effect of political ideology on the uniqueness and conformity language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

Sears, 1986). Importantly, we also utilized nonreactive measurements that remove the influence of the researcher during data collection and thus minimize demand characteristics and experimenter bias in the construction of studies, thereby alleviating concerns about alleged ideological biases of researchers.

Assessing naturally occurring language (rather than closed-ended responses) may also provide more sensitive measures of psychological characteristics. In the current research program, we analyzed spontaneous forms of discourse to explore complex relationships between psychological and political variables and to identify both similarities and differences in the language used by liberals and conservatives. Specifically, we concatenated predictions from 50 years of research on linguistic differences between liberals and conservatives and used dictionary-based methods to assess the robustness of predictions across a sample of approximately 25,000 Twitter users. On the basis of this data set, we comprehensively evaluated the evidence bearing on a large number of hypotheses (27 of them) pertaining to linear and quadratic effects of political orientation on the use of language reflecting psychological needs, values, interests, and concerns.

In light of the fact that using naturally occurring language to measure psychological tendencies is fairly novel, it may be useful to compare the relationships we observed utilizing linguistic measures to those observed in traditional research on motive and value assessments. Overall, we observed 23 significant differences in the linguistic styles of liberals and conservatives (out of 32 tests). Eighteen of these 23 differences were consistent with evidence based on more traditional studies of ideological differences in values and motivations. More specifically, we observed that liberals tended to use language that conveyed benevolence, whereas conservatives used more language pertaining to power, certainty, resistance to change, tradition, threat, risk, security, anger, anxiety, and negative emotion in general.

Somewhat surprisingly, we observed that conservatives were more likely than liberals to use both tentative and future-focused language—two categories that convey subjective uncertainty. One possible explanation for these results is in terms of dimensional thinking (Chung & Pennebaker, 2008), which suggests that having a high need for certainty may lead people to use more words that indicate their present level (or degree) of certainty rather than the expression of certainty regardless of the situation. Experimental research designed to manipulate epistemic motivation and measure subsequent linguistic behavior would be necessary to fully explore this possibility.

In addition to obtaining a great deal of support for the notion that liberals and conservatives communicate differently from one another, we also found some evidence that ideological extremists communicate differently than moderates (or centrists). In 17 of 32 cases we observed results consistent with the J-curve pattern described by Jost et al. (2003a). That is, both conservatism and ideological extremity were associated with the use of language pertaining to certainty, resistance to change, tentativeness, anxiety, inhibition, security, anger, negative affect, and affiliation.[6] On the other hand, liberalism and moderation were associated with the use of benevolent language. Conservatism—but not ideological extremity—was associated with the use of language related to threat, power, and tradition. Extremity in general was associated with the increased use of death-related language. Although previous studies provided only sparse empirical evidence pertaining to linear versus quadratic effects, the results we obtained adhere to the logic of the theory of political ideology as motivated social cognition. In particular, these findings support the propositions that (a) all ideologies communicate some (subjective) degree of safety and security, and yet (b) some ideologies do so more thoroughly or explicitly than others (e.g., see Hennes et al., 2012).

For the remaining categories, neither a linear nor quadratic relationship accounted adequately for the data patterns. In eight cases, extremity within the group of liberals or conservatives was associated with language use. We detected no meaningful relationships between political ideology and the use of words pertaining to inequality, conformity, stimulation, universalism, feeling, or positive emotion words. It is important to note, however, that with the exception of the latter two categories, these null results were observed for some of the least frequently used categories. Low base rates might make it especially difficult to specify the nature of the relationship even in very large data sets.

Taking all of these findings in conjunction, there are a few noteworthy patterns overall. It is clear that in the majority of cases

[6] With respect to affiliation language, it is reasonable to suppose—given other research in political psychology—that conservatives would be especially likely to use words that pertain to the family and other close in-groups (see also Table 2).

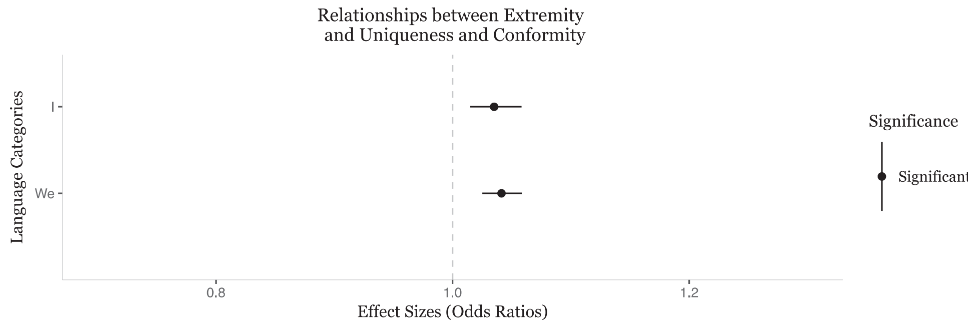

Figure 12. Odds ratio estimates of the quadratic effect of political ideology on the uniqueness and conformity language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

there are robust effects associated with left-right ideology and ideological extremity. If one were to ignore the role of extremity and focus on left-right ideology in isolation, the models would misestimate the psychological profiles of extreme liberals. The J-curve pattern (as we have been referring to it in this article) was especially evident with respect to epistemic and existential motivations. As hypothesized on the basis of the theory of political ideology as motivated social cognition, we find that conservatives do indeed possess stronger epistemic and existential needs to reduce uncertainty and threat (as evidenced by their language use). At the same time, we also see that extreme liberals use language pertaining to threat and certainty more often than moderate liberals. Death-related language was more prevalent at both ideological extremes.

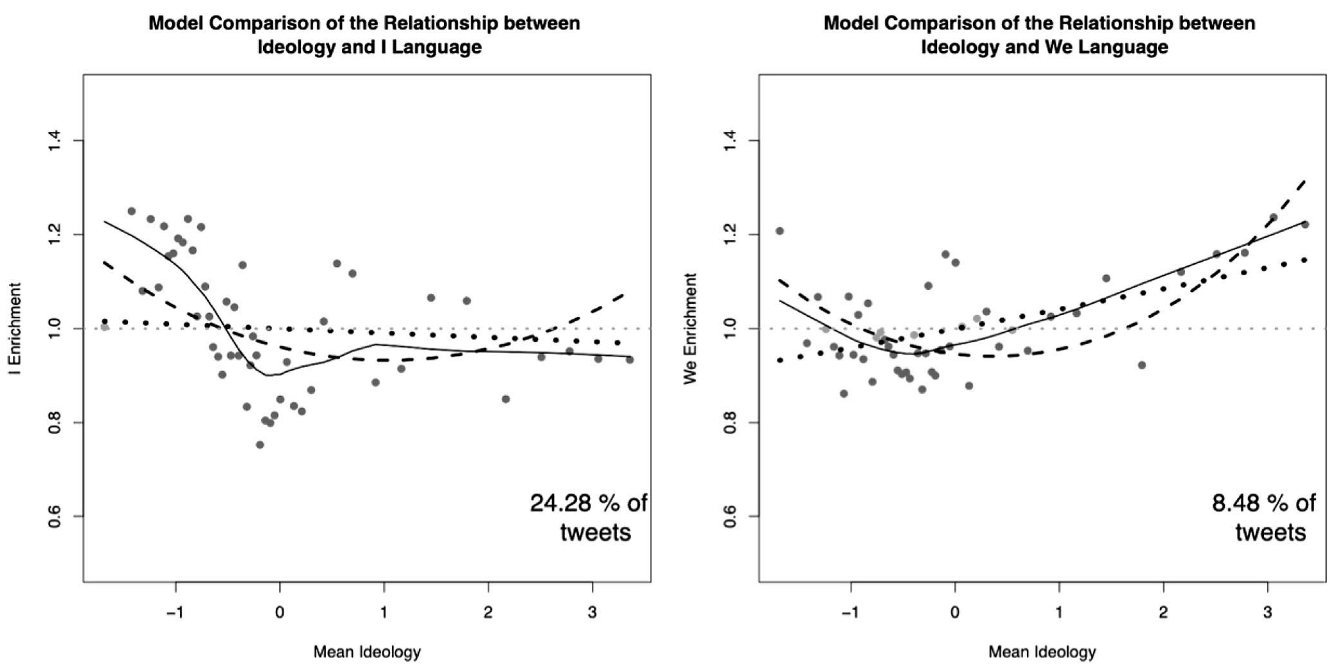

Figure 13. Model comparisons for uniqueness and conformity variables. Tweets were ordered by political ideology and then the dataset was split into 50 equal segments (each segment contains 240,000 tweets). Then, enrichment analyses were conducted separately for each segment to test whether tweets within that segment were significantly more or less likely to contain this specific type of language than we would expect on average. Significance was determined using a multiple hypothesis corrected alpha ($\alpha = .0037$). Dark gray dots represent groups of tweets that significantly differed from the average proportion. Light gray dots represent groups of tweets did not significantly differ from the average proportion. Solid lines are lowess lines fit to the results of the enrichment analyses (taking into account 66% of the data at a time). The dotted line represents the expected values implied by the linear regression model. The dashed line represents the expected values implied by the quadratic regression model (see Appendix B for details).

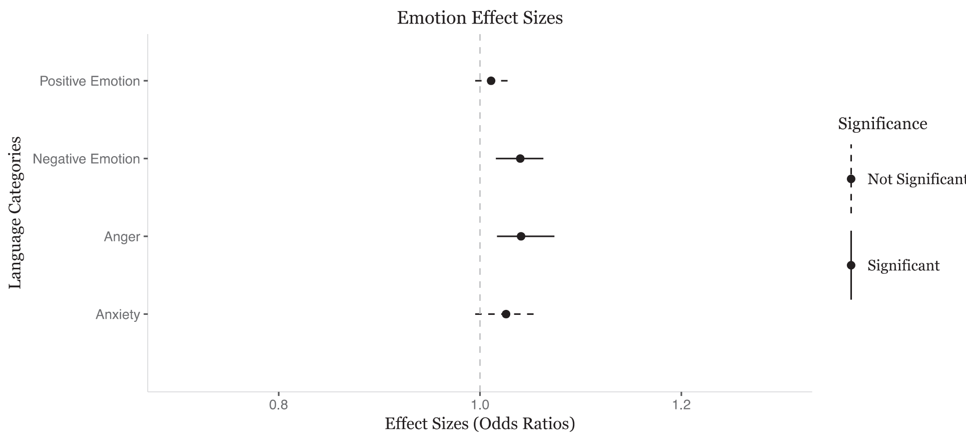

Figure 14. Odds ratio estimates of the effect of political ideology on the emotion language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

There were other categories of language that were associated more specifically with liberalism or conservatism (rather than ideological extremity). These categories may be linked to stances taken by the two major political parties in the United States. Conservatives were especially likely to use words related to power, tradition, and threat. To some degree, these themes correspond to the Republican emphasis on strong military forces, support for capitalism, and religious moral codes. Liberals, on the other hand, were especially likely to use words pertaining to benevolence, perhaps reflecting the Democratic emphasis on equality and social welfare provisions. These findings too, are consistent with the theory of political ideology as motivated social cognition (Jost et al., 2003a, 2003b).

Nevertheless, there were some hypotheses that were not supported, although they had been supported in previous research. For instance, we obtained no consistent evidence of a special link between anxiety and conservatism, and contrary to predictions, we found that conservatives used language referring to the future more than liberals. We cannot say definitively why these results did not accord with previous findings, but there are a few factors that may be useful to consider. One potential explanation is that we are measuring publicly observable behavior rather than privately reported attitudes. It seems likely that people would express psychological tendencies in their spontaneous online behavior that they would not necessarily recognize or report in the context of a research study. A less interesting explanation has to do with characteristics of the dictionaries we used. Shorter dictionaries, such as those developed by Neiman and colleagues (2016a, 2016b), may have been less reliable or valid than longer dictionaries such as LIWC and Harvard; we observed that the smaller dictionaries were more likely to produce null effects than the larger dictionaries.

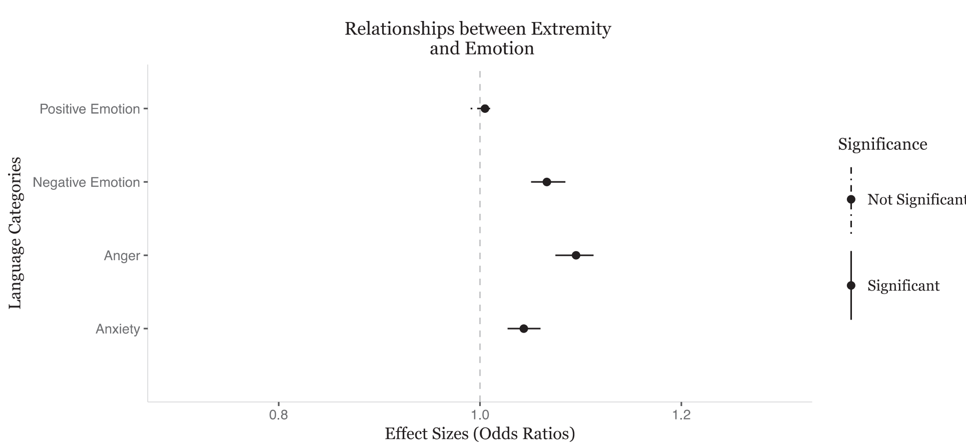

Figure 15. Odds ratio estimates of the quadratic effect of political ideology on the emotion language categories from a sample of approximately 25,000 ordinary citizen Twitter users. Effects computed using multilevel mixed-effects logistic regression analyses with random intercepts for Twitter users. Points depict median effect sizes from 500 bootstraps, sampled with replacement. Lines depict 99.6% confidence intervals. Solid lines differ meaningfully from zero. Dotted lines do not differ meaningfully from zero.

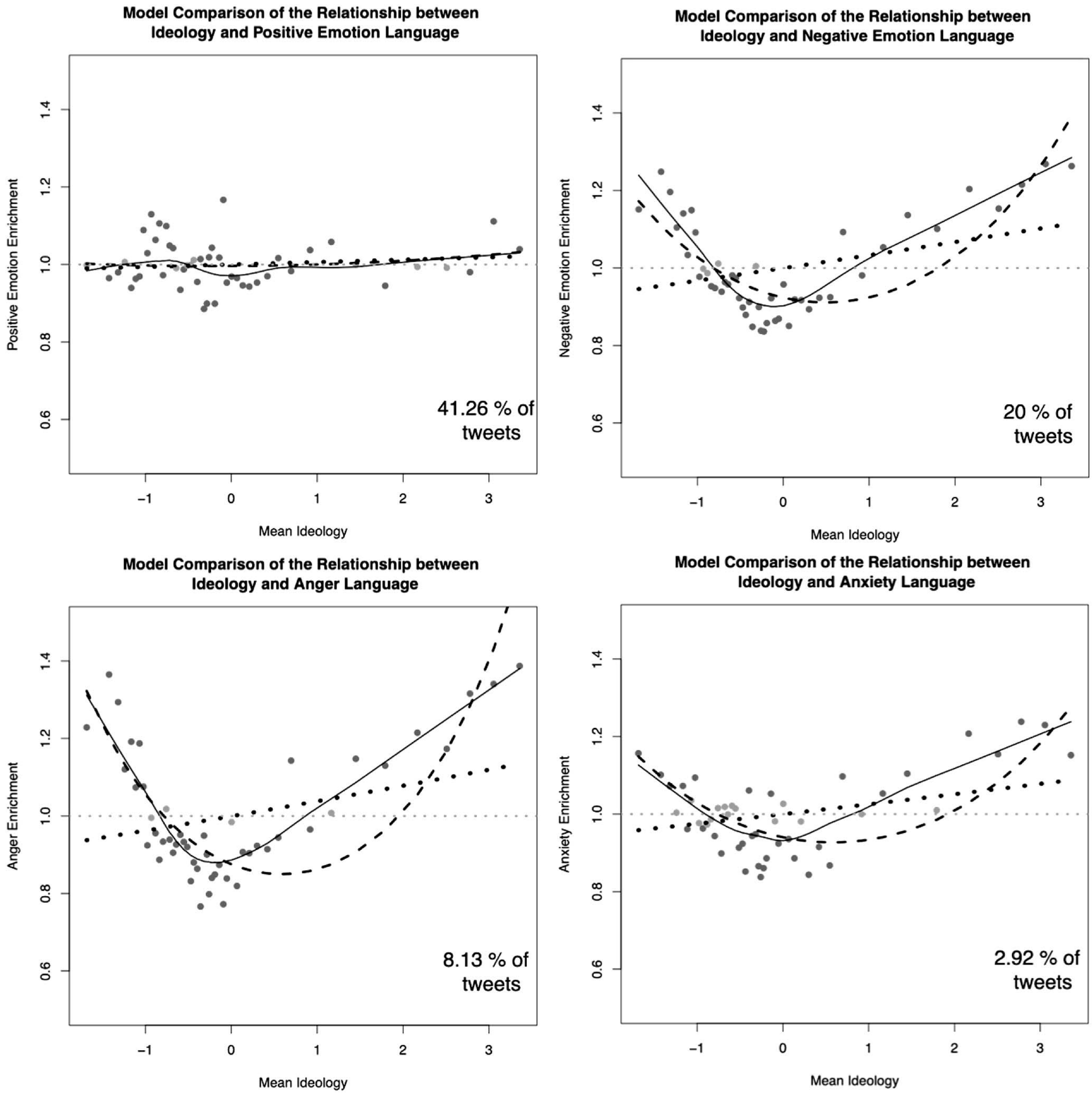

Figure 16. Model comparisons for emotion variables. Tweets were ordered by political ideology, and then the dataset was split into 50 equal segments (each segment contains 240,000 tweets). Then, enrichment analyses were conducted separately for each segment to test whether tweets within that segment were significantly more or less likely to contain this specific type of language than we would expect on average. Significance was determined using a multiple hypothesis corrected alpha (α = .0037). Dark gray dots represent groups of tweets that significantly differed from the average proportion. Light gray dots represent groups of tweets did not significantly differ from the average proportion. Solid lines are lowess lines fit to the results of the enrichment analyses (taking into account 66% of the data at a time). The dotted line represents the expected values implied by the linear regression model. The dashed line represents the expected values implied by the quadratic regression model (see Appendix B for details).

There are two additional findings of note. The first pertains to the issue of construct validity, namely the question of whether dictionaries—many of which contain a great number and variety of words—tap into similar psychological constructs for liberals and conservatives. Although the goal of dictionary-based methods is to measure the same underlying construct, it is conceivable that liberals and conservatives are using very different vocabularies even within the same semantic space. In such a situation the claim that one is consistently measuring the same construct may be problematic. Fortunately, when we analyzed the most frequently used dictionary words for liberals and conservatives, we observed a remarkable degree of similarity in the structure of psychological constructs for the overwhelming majority of dictionaries (see Table 2). In two other cases, the language differences were revealing. With respect to negative affect, conservatives often used words referring to threat, such as "terror" and "attack," but liberals did not. With respect to power-related language, conservatives tended to use words like "god" and "win," whereas liberals were more likely to use communication-related words such as "show" and "say."

There are a few key limitations associated with our textual approach to assessing the motivational and value-based underpinnings of political ideology. First, there are some inconsistencies in the ways that researchers have created their dictionaries, especially when it comes to unidimensional versus multidimensional conceptions. Whereas the LIWC and Harvard versions of the power dictionary include words that indicate high and low power, most of the dictionaries developed by Neiman and colleagues (2016a, 2016b) contain only one level (i.e., words that indicate having power). Given the possible role of dimensional thinking, as described above, we recommend that future research prioritize bipolar (rather than unipolar) dictionary constructions. Of course, there are other limitations as well. Using a specific word does not mean that a Twitter user endorses the concept. In addition, many words have multiple meanings, but in most cases only one meaning fits the conceptual category. Twitter messages are also very brief, and it is difficult, if not impossible to capture contextual aspects of language use with the methods we have employed.

Second, although our sample was more statistically representative of the population than in most psychological research, there are clearly limitations associated with our sampling frame. The fact that all of our participants were social media users means that the sample was younger, more male, and more educated than the rest of the population (Pew Research Institute, 2016). Furthermore, all participants followed at least three political elite accounts; this enabled us to estimate their ideological position accurately. Although the list of elites we utilized was reasonably comprehensive and contained news outlets and think tanks as well as politicians, this procedure most likely produced a sample that was more politically involved and sophisticated, in comparison with the general public. It is also unfortunate, as noted above, that we were not able to adjust for specific demographic characteristics of individual users in our sample. This means that certain factors—such as gender or race/ethnicity—that covary with political orientation may contribute to the ideological differences we have observed. On this point, it is worth noting that—in a study of members of the U.S. Congress, for whom information about sex and ethnicity is available, we have replicated some (but not all) of the ideological differences observed here (Sterling & Jost, 2019; see also Brady, Wills, Burkart, Jost, & Van Bavel, in press; Sterling & Jost, 2018).

Third, we obtained consistent support for many of the hypotheses specified in previous work, but it is also important to recognize that most of the effects sizes were relatively small. There are a number of reasons to expect small effect sizes in this context of research on psycholinguistics. As noted in the introduction, dictionary-based methods are imprecise when it comes to measuring linguistic behavior because they ignore features of the social context and focus on stable patterns over time. Dictionary-based methods function rather well when it comes to analyzing very large samples of text, but they still capture a good deal of noise along with the psychological signal. It is also important to keep in mind that we harvested all messages sent by the Twitter users in our sample regardless of their topic or structure. Because of the informal and diverse nature of communication on Twitter, the language sample we analyzed included an extremely wide range of topics as well as slang, typos, misspellings, and abbreviations. All of these linguistic features are likely to have contributed statistical noise. Finally, the follower-based method that we have used to estimate users' ideological positions proceeds from an assumption of homophily. Although this method of ideological estimation has been thoroughly validated in previous research (see Barberá, 2015), it almost surely mischaracterizers some users' ideological positions.

Conclusion

In summary, this research program contributes to the development of methodological tools that allow social and behavioral scientists to reach beyond the traditional confines of the research laboratory to study naturally occurring communication behavior. We have also refined measures of psychological characteristics of liberals and conservatives in a manner that does not rely upon self-report measures or awareness of one's own psychological characteristics. Furthermore, we used these unobtrusive methods in an attempt to resolve a number of theoretical controversies in the political psychology literature. Specifically, we evaluated the evidence bearing on the question of ideological symmetries and asymmetries (Jost, 2017a), taking into account the role of ideological extremity as well as liberalism-conservatism (Jost et al., 2003a). Importantly, we observe both directional and nondirectional effects of political ideology, suggesting that the correspondence between psychological characteristics and ideological preferences is affected by conservatism and extremity in combination.

In addition to the theoretical, methodological, and empirical contributions that the automated psychological analysis of text offers the field of psychology, there are a number of practical implications of this work. One of the most exciting is the potential to harness these psychological expressions on highly representative samples and hard-to-reach populations in real time. Based on archival and experimental studies, we know that exogenous threats (such as terrorist attacks) can produce ideological shifts (Jost et al., 2017). As the precision of linguistic tools continue to develop, it is possible to imagine detecting macrolevel changes in public opinion based on subtle changes in language usage. Indeed, one can envision—some years hence, assuming that further empirical progress is made—the prospect of reverse-engineering political ideology. That is, it may one day be possible to unobtrusively estimate fairly precise ideological positions—of individuals and even of entire populations—on the basis of dynamic patterns of linguistic habits for, as Umberto Eco pointed out, very often these habits are "symptoms of unspoken sentiments."

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York, NY: Harper.

Altemeyer, R. A. (1981). *Right-wing authoritarianism*. Winnipeg, Canada: University of Manitoba Press.

Anson, J., Pyszczynski, T., Solomon, S., & Greenberg, J. (2009). Political ideology in post-9/11 America: A terror management perspective on maintenance and change of the status quo. In J. T. Jost, A. C. Kay, & H. Thórisdóttir (Eds.), *Social and psychological bases of ideology and system justification* (pp. 210–240). New York, NY: Oxford University.

Atkinson, J. W. (1958). Towards experimental analysis of human motivation in terms of motives, expectancies, and incentives. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society* (pp. 288–305). Princeton, NJ: Van Nostrand.

Barberá, P. (2015). Birds of the same feather tweet together. Bayesian ideal point estimation using twitter data. *Political Analysis, 23,* 76–91. http://dx.doi.org/10.1093/pan/mpu011

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science, 26,* 1531–1542. http://dx.doi.org/10.1177/0956797615594620

Bixter, M. T. (2015). Happiness, political orientation, and religiosity. *Personality and Individual Differences, 72,* 7–11. http://dx.doi.org/10.1016/j.paid.2014.08.010

Boutyline, A., & Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology, 38,* 551–569. http://dx.doi.org/10.1111/pops.12337

Brady, W. J., Wills, J. A., Burkart, D., Jost, J. T., & Van Bavel, J. J. (in press). An ideological asymmetry in the diffusion of moralized content on social media among political leaders. *Journal of Experimental Psychology: General*. Advance online publication. http://dx.doi.org/10.1037/xge0000532

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences of the United States of America, 114,* 7313–7318. http://dx.doi.org/10.1073/pnas.1618923114

Brundidge, J., Reid, S. A., Choi, S., & Muddiman, A. (2014). The "deliberative digital divide:" Opinion leadership and integrative complexity in the U.S. political blogosphere. *Political Psychology, 35,* 741–755. http://dx.doi.org/10.1111/pops.12201

Burton, C. M., Plaks, J. E., & Peterson, J. B. (2015). Why do conservatives report being happier than liberals? The contribution of neuroticism. *Journal of Social and Political Psychology, 3,* 89–102. http://dx.doi.org/10.5964/jspp.v3i1.117

Butz, S., Kieslich, P. J., & Bless, H. (2017). Why are conservatives happier than liberals? Comparing different explanations based on system justification, multiple group membership, and positive adjustment. *European Journal of Social Psychology, 47,* 362–372. http://dx.doi.org/10.1002/ejsp.2283

Caprara, G. V., Schwartz, S., Capanna, C., Vecchione, M., & Barbaranelli, C. (2006). Personality and politics: Values, traits, and political choice. *Political Psychology, 27,* 1–28. http://dx.doi.org/10.1111/j.1467-9221.2006.00447.x

Caprara, G. V., & Vecchione, M. (2017). *Personalizing politics and realizing democracy*. United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199982868.001.0001

Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology, 29,* 807–840. http://dx.doi.org/10.1111/j.1467-9221.2008.00668.x

Castano, E., Leidner, B., Bonacossa, A., Nikkah, J., Perrulli, R., Spencer, B., & Humphrey, N. (2011). Ideology, fear of death, and death anxiety. *Political Psychology, 32,* 601–621. http://dx.doi.org/10.1111/j.1467-9221.2011.00822.x

Choma, B. L., Busseri, M. A., & Sadava, S. W. (2009). Liberal and conservative political ideologies: Different routes to happiness? *Journal of Research in Personality, 43,* 502–505. http://dx.doi.org/10.1016/j.jrp.2008.12.016

Chung, C. K., & Pennebaker, J. W. (2008). Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality, 42,* 96–132. http://dx.doi.org/10.1016/j.jrp.2007.04.006

Cichocka, A., Bilewicz, M., Jost, J. T., Marrouch, N., & Witkowska, M. (2016). On the grammar of politics—Or why conservatives prefer nouns. *Political Psychology, 37,* 799–815. http://dx.doi.org/10.1111/pops.12327

Cichocka, A., & Jost, J. T. (2014). Stripped of illusions? Exploring system justification processes in capitalist and post-Communist societies. *International Journal of Psychology, 49,* 6–29. http://dx.doi.org/10.1002/ijop.12011

Devos, T., Spini, D., & Schwartz, S. H. (2002). Conflicts among human values and trust in institutions. *British Journal of Social Psychology, 41,* 481–494. http://dx.doi.org/10.1348/014466602321149849

Duncan, L. E., & Peterson, B. E. (2010). Gender and motivation for achievement, affiliation–intimacy, and power. In J. Chrisler & D. McCreary (Eds.), *Handbook of gender research in psychology* (pp. 41–62). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4419-1467-5_3

Enelow, J. M., & Hinich, M. J. (1984). *The spatial theory of voting: An Introduction*. Cambridge, UK: Cambridge University Press.

Eysenck, H. (1954). *The psychology of politics*. New Brunswick, NJ: Transaction.

Fetterman, A. K., Boyd, R. L., & Robinson, M. D. (2015). Power versus affiliation in political ideology: Robust linguistic evidence for distinct motivation-related signatures. *Personality and Social Psychology Bulletin, 41,* 1195–1206. http://dx.doi.org/10.1177/0146167215591960

Gerber, A. S., Huber, G. A., Doherty, D., Dowling, C. M., Raso, C., & Ha, S. E. (2011). Personality traits and participation in political processes. *The Journal of Politics, 73,* 692–706. http://dx.doi.org/10.1017/S0022381611000399

Greenberg, J., & Jonas, E. (2003). Psychological motives and political orientation—The left, the right, and the rigid: Comment on Jost et al. (2003). *Psychological Bulletin, 129,* 376–382. http://dx.doi.org/10.1037/0033-2909.129.3.376

Hennes, E. P., Nam, H. H., Stern, C., & Jost, J. T. (2012). Not all ideologies are created equal: Epistemic, existential, and relational needs predict system-justifying attitudes. *Social Cognition, 30,* 669–688. http://dx.doi.org/10.1521/soco.2012.30.6.669

Henry, P. J. (2008). College sophomores in the laboratory redux: Influences of a narrow data base on social psychology's view of the nature of prejudice. *Psychological Inquiry, 19,* 49–71. http://dx.doi.org/10.1080/10478400802049936

Janoff-Bulman, R., Sheikh, S., & Baldacci, K. (2008). Mapping moral motives: Approach, avoidance, and political orientation. *Journal of Experimental Social Psychology, 44,* 1091–1099. http://dx.doi.org/10.1016/j.jesp.2007.11.003

Jones, K. L., Noorbaloochi, S., Jost, J. T., Bonneau, R., Nagler, J., & Tucker, J. A. (2018). Liberal and conservative values: What we can learn from Congressional tweets. *Political Psychology, 39,* 423–443. http://dx.doi.org/10.1111/pops.12415

Jost, J. T. (2017a). Ideological asymmetries and the essence of political psychology. *Political Psychology, 38,* 167–208. http://dx.doi.org/10.1111/pops.12407

Jost, J. T. (2017b). The marketplace of ideology: "Elective affinities" in political psychology and their implications for consumer behavior. *Journal of Consumer Psychology, 27,* 502–520. http://dx.doi.org/10.1016/j.jcps.2017.07.003

Jost, J. T. (2019). Anger and authoritarianism mediate the effects of fear on support for the far right: What Vasilopoulos et al. (2019) really found. *Political Psychology, 40,* 705–711. http://dx.doi.org/10.1111/pops.12567

Jost, J. T., Barberá, P., Bonneau, R., Langer, M., Metzger, M. M., Nagler, J., . . . Tucker, J. A. (2018). How social media facilitates political protest: Informational and motivational pathways from social media to political protest. *Advances in Political Psychology, 39,* 85–118. http://dx.doi.org/10.1111/pops.12478

Jost, J. T., Basevich, E., Dickson, E. S., & Noorbaloochi, S. (2016). The place of values in a world of politics: Personality, motivation, and ideology. In T. Brosch & D. Sander (Eds.), *Handbook of value: Perspectives from economics, neuroscience, philosophy, psychology, and sociology* (pp. 351–374). New York, NY: Oxford University Press.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003a). Exceptions that prove the rule–Using a theory of motivated social cognition to account for ideological incongruities and political anomalies: Reply to Greenberg and Jonas (2003). *Psychological Bulletin, 129,* 383–393.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003b). Political conservatism as motivated social cognition. *Psychological Bulletin, 129,* 339–375. http://dx.doi.org/10.1037/0033-2909.129.3.339

Jost, J. T., Ledgerwood, A., & Hardin, C. D. (2008). Shared reality, system justification, and the relational basis of ideological beliefs. *Social and Personality Psychology Compass, 2,* 171–186. http://dx.doi.org/10.1111/j.1751-9004.2007.00056.x

Jost, J. T., Sterling, J., & Stern, C. (2018). Getting closure on conservatism, or the politics of epistemic and existential motivation. In C. Kopetz & A. Fishbach (Eds.), *The motivation-cognition interface; From the lab to the real world: A Festschrift in honor of Arie W. Kruglanski* (pp. 56–87). New York, NY: Routledge.

Jost, J. T., Stern, C., Rule, N. O., & Sterling, J. (2017). The politics of fear: Is there an ideological asymmetry in existential motivation? *Social Cognition, 35,* 324–353. http://dx.doi.org/10.1521/soco.2017.35.4.324

Jost, J. T., van der Linden, S., Panagopoulos, C., & Hardin, C. D. (2018). Ideological asymmetries in conformity, desire for shared reality, and the spread of misinformation. *Current Opinion in Psychology, 23,* 77–83. http://dx.doi.org/10.1016/j.copsyc.2018.01.003

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology, 33,* 125–143. http://dx.doi.org/10.1177/0261927X13502654

Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. *The American Journal of Psychology, 120,* 263–286.

Kemmelmeier, M. (2008). Is there a relationship between political orientation and cognitive ability? A test of three hypotheses in two studies. *Personality and Individual Differences, 45,* 767–772. http://dx.doi.org/10.1016/j.paid.2008.08.003

Lakoff, G. (2004). *Don't think of an elephant*. White River Junction, VT: Chelsea Green.

Lane, R. E. (1962). *Political ideology: Why the American common man believes what he does*. New York, NY: The Free Press.

Langer, M., Jost, J. T., Bonneau, R., Metzger, M., Noorbaloochi, S., & Penfold-Brown, D. (2019). Digital dissent: An analysis of the motivational contents of Tweets from an Occupy Wall Street demonstration. *Motivation Science, 5,* 14–34. http://dx.doi.org/10.1037/mot0000084

McClelland, D. C. (1979). Inhibited power motivation and high blood pressure in men. *Journal of Abnormal Psychology, 88,* 182–190. http://dx.doi.org/10.1037/0021-843X.88.2.182

McClelland, D. C. (1987). Biological aspects of human motivation. In F. Halisch & J. Kuhl (Eds.), *Motivation, intention and volition* (pp. 11–19). New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-70967-8_2

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive*. New York, NY: Appleton-Century-Crofts. http://dx.doi.org/10.1037/11144-000

McClosky, H. (1958). Conservatism and personality. *The American Political Science Review, 52,* 27–45. http://dx.doi.org/10.2307/1953011

Murray, H. A. (1943). *Contemporary psychopathology: A source book* (Vol. 2644). Cambridge, MA: Harvard University Press.

Napier, J. L., & Jost, J. T. (2008). Why are conservatives happier than liberals? *Psychological Science, 19,* 565–572. http://dx.doi.org/10.1111/j.1467-9280.2008.02124.x

Neiman, J. L., Gonzalez, F. J., Wilkinson, K., Smith, K. B., & Hibbing, J. R. (2016a). Speaking different languages or reading from the same script? Word usage of Democratic and Republican politicians. *Political Communication, 33,* 212–240. http://dx.doi.org/10.1080/10584609.2014.969465

Neiman, J. L., Gonzalez, F. J., Wilkinson, K., Smith, K. B., & Hibbing, J. R. (2016b). Corrigendum: Speaking different languages or reading from the same script? Word usage of Democratic and Republican politicians. *Political Communication, 33,* 346–349. http://dx.doi.org/10.1080/10584609.2016.1161349

Newman, M. L., Groom, C. L., Stone, L. D., & Pennebaker, J. W. (2006). *Sex differences in language use: An analysis of text samples from 70 studies*. Unpublished manuscript.

Okdie, B. M., & Rempala, D. M. (2019). Brief textual indicators of political orientation. *Journal of Language and Social Psychology, 38,* 106–125. http://dx.doi.org/10.1177/0261927X18762973

Okulicz-Kozaryn, A., Holmes, O., & Avery, D. R. (2014). The subjective well-being political paradox: Happy welfare states and unhappy liberals. *Journal of Applied Psychology, 99,* 1300–1308. http://dx.doi.org/10.1037/a0037654

Onraet, E., Van Assche, J., Roets, A., Haesevoets, T., & Van Hiel, A. (2017). The happiness gap between conservatives and liberals depends on country-level threat: A worldwide multilevel study. *Social Psychological and Personality Science, 8,* 11–19. http://dx.doi.org/10.1177/1948550616662125

Pattershall, J., & Eidelman, S. (2008). *Prevention, promotion, and politics: Evidence that regulatory focus affects political conservatism*. Presented at the annual meeting of the Society for Personality and Social Psychology (SPSP).

Pennebaker, J. W., Booth, R. J., Boyd, R. L., & Francis, M. E. (2015). *Linguistic inquiry and word count: LIWC2015 operator's manual*. Austin, TX: Pennebaker Conglomerates.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77,* 1296–1312. http://dx.doi.org/10.1037/0022-3514.77.6.1296

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language. use: Our words, our selves. *Annual Review of Psychology, 54,* 547–577. http://dx.doi.org/10.1146/annurev.psych.54.101601.145041

Pew Research Institute. (2016). *Social media updates 2016*. Washington, DC: Author. Retrieved from https://www.pewinternet.org/2016/11/11/social-media-update-2016/

Piurko, Y., Schwartz, S. H., & Davidov, E. (2011). Basic personal values and the meaning of left-right political orientations in 20 countries. *Political Psychology, 32,* 537–561. http://dx.doi.org/10.1111/j.1467-9221.2011.00828.x

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67,* 741–763. http://dx.doi.org/10.1037/0022-3514.67.4.741

Rabinowitz, M., Latella, L., Stern, C., & Jost, J. T. (2016). Beliefs about childhood vaccination in the United States: Political ideology, false consensus, and the illusion of uniqueness. *PLoS ONE, 11,* e0158382. http://dx.doi.org/10.1371/journal.pone.0158382

Robinson, M. D., Boyd, R. L., & Fetterman, A. K. (2014). An emotional signature of political ideology: Evidence from two linguistic content-coding studies. *Personality and Individual Differences, 71,* 98–102. http://dx.doi.org/10.1016/j.paid.2014.07.039

Robinson, M. D., Cassidy, D. M., Boyd, R. L., & Fetterman, A. K. (2015). The politics of time: Conservatives differentially reference the past and liberals differentially reference the future. *Journal of Applied Social Psychology, 45,* 391–399. http://dx.doi.org/10.1111/jasp.12306

Rokeach, M. (1960). *The open and closed mind*. Oxford, UK: Basic Books.

Rorschach, H. (1921). *Psychodiagnostics* (translated by P. Lemkau). Bern, Switzerland: Ernst Bircher.

Schlenker, B. R., Chambers, J. R., & Le, B. M. (2012). Conservatives are happier than liberals, but why? Political ideology, personality, and life satisfaction. *Journal of Research in Personality, 46,* 127–146. http://dx.doi.org/10.1016/j.jrp.2011.12.009

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). New York, NY: Academic Press. http://dx.doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology, 53,* 550–562. http://dx.doi.org/10.1037/0022-3514.53.3.550

Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology, 58,* 878–891. http://dx.doi.org/10.1037/0022-3514.58.5.878

Schwartz, S. H., Caprara, G. V., & Vecchione, M. (2010). Basic personal values, core political values, and voting: A longitudinal analysis. *Political Psychology, 31,* 421–452. http://dx.doi.org/10.1111/j.1467-9221.2010.00764.x

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51,* 515–530. http://dx.doi.org/10.1037/0022-3514.51.3.515

Sibley, C. G., Osborne, D., & Duckitt, J. (2012). Personality and political orientation: Meta-analysis and test of a Threat-Constraint Model. *Journal of Research in Personality, 46,* 664–677. http://dx.doi.org/10.1016/j.jrp.2012.08.002

Sidanius, J. (1988). Political sophistication and political deviance: A structural equation examination of context theory. *Journal of Personality and Social Psychology, 55,* 37–51. http://dx.doi.org/10.1037/0022-3514.55.1.37

Sidanius, J., & Pratto, F. (1993). The dynamics of social dominance and the inevitability of oppression. In P. Sniderman & P. E. Tetlock (Eds.), *Prejudice, politics, and race in America today* (pp. 173–211). California: Stanford University Press.

Smith, L. G., Gavin, J., & Sharp, E. (2015). Social identity formation during the emergence of the occupy movement. *European Journal of Social Psychology, 45,* 818–832. http://dx.doi.org/10.1002/ejsp.2150

Sorrentino, R. M., & Roney, C. J. R. (2000). *The uncertain mind: Individual differences in facing the unknown*. Philadelphia, PA: Psychology Press/Taylor & Francis.

Sterling, J., & Jost, J. T. (2018). Moral discourse in the Twitterverse: Effects of ideology and political sophistication on language use among U.S. citizens and members of Congress. *Journal of Language and Politics, 17,* 195–221. http://dx.doi.org/10.1075/jlp.17034.ste

Sterling, J., & Jost, J. T. (2019). *A psycholinguistic analysis of social media communication from U.S. Congressional accounts: Effects of ideology, national threat, and platform*. Unpublished manuscript.

Sterling, J., Jost, J. T., & Pennycook, G. (2016). Are neoliberals more susceptible to bullshit? *Judgment and Decision Making, 11,* 352–360.

Stern, C., West, T. V., Jost, J. T., & Rule, N. O. (2014). "Ditto heads": Do conservatives perceive greater consensus within their ranks than liberals? *Personality and Social Psychology Bulletin, 40,* 1162–1177. http://dx.doi.org/10.1177/0146167214537834

Stern, C., West, T. V., & Schmitt, P. G. (2014). The liberal illusion of uniqueness. *Psychological Science, 25,* 137–144. http://dx.doi.org/10.1177/0956797613500796

Stone, P. J., Dunphy, D. C., Smith, M. S., & Ogilvie, D. M. (1966). *The general inquirer: A computer approach to content analysis: Studies in psychology*. Cambridge, MA: MIT Press.

Suedfeld, P., Bluck, S., Ballard, E. J., & Baker-Brown, G. (1990). Canadian federal elections: Motive profiles and integrative complexity in political speeches and popular media. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 22,* 26–36. http://dx.doi.org/10.1037/h0078936

Sylwester, K., & Purver, M. (2015). Twitter language use reflects psychological differences between democrats and republicans. *PLoS ONE, 10,* e0137422. http://dx.doi.org/10.1371/journal.pone.0137422

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29,* 24–54. http://dx.doi.org/10.1177/0261927X09351676

Tetlock, P. E. (1983). Cognitive style and political ideology. *Journal of Personality and Social Psychology, 45,* 118–126. http://dx.doi.org/10.1037/0022-3514.45.1.118

Tetlock, P. E., Bernzweig, J., & Gallant, J. L. (1985). Supreme Court decision making: Cognitive style as a predictor of ideological consistency of voting. *Journal of Personality and Social Psychology, 48,* 1227–1239. http://dx.doi.org/10.1037/0022-3514.48.5.1227

Theocharis, Y., Lowe, W., Van Deth, J. W., & García-Albacete, G. (2015). Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication, & Society, 18,* 202–220. http://dx.doi.org/10.1080/1369118X.2014.948035

Tomkins, S. S. (1963). Left and right: A basic dimension of ideology and personality. In R. W. White (Ed.), *The study of lives* (pp. 388–411). Chicago, IL: Atherton. http://dx.doi.org/10.1037/12238-017

van Prooijen, J. W., & Krouwel, A. P. M. (2019). Psychological features of extreme political ideologies. *Current Directions in Psychological Science, 28,* 159–163. http://dx.doi.org/10.1177/0963721418817755

Vasilopoulos, P., Marcus, G. E., Valentino, N. A., & Foucault, M. (2019). Fear, anger, and voting for the far right: Evidence from the November 13, 2015 Paris terror attacks. *Political Psychology, 40,* 679–704. http://dx.doi.org/10.1111/pops.12513

Vecchione, M., Caprara, G., Dentale, F., & Schwartz, S. H. (2013). Voting and values: Reciprocal effects over time. *Political Psychology, 34,* 465–485. http://dx.doi.org/10.1111/pops.12011

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1969). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago, IL: Rand McNally.

Wetherell, M., & Potter, J. (1992). *Mapping the language of racism: Discourse and the legitimation of exploitation*. New York, NY: Columbia University.

Wilson, G. D. (Ed.). (1973). *The psychology of conservatism*. London, UK: Academic Press.

Wilson, M. S., & Sibley, C. G. (2013). Social dominance orientation and right-wing authoritarianism: Additive and interactive effects on political conservatism. *Political Psychology, 34,* 277–284. http://dx.doi.org/10.1111/j.1467-9221.2012.00929.x

Winter, D. G. (1991). A motivational model of leadership: Predicting long-term management success from TAT measures of power motivation and responsibility. *The Leadership Quarterly, 2,* 67–80. http://dx.doi.org/10.1016/1048-9843(91)90023-U

Wojcik, S. P., Hovasapian, A., Graham, J., Motyl, M., & Ditto, P. H. (2015). Conservatives report, but liberals display, greater happiness. *Science, 347,* 1243–1246. http://dx.doi.org/10.1126/science.1260817

Appendix A

Some Operationalization Differences Among Dictionaries Measuring the Same Psychological Construct

Table A1
Operationalization Differences Among Dictionaries Measuring the Same Construct

Construct	Operationalization	Dictionary contents (First twenty terms only)
Achievement	LIWC 2015	abilit*, able, accomplish*, ace, achievable, achieve*, achievi*, acquir*, acquisition*, actualiz*, adequa*, advanc*, advantag*, ahead, ambition, ambitions, ambitious, ambitiously, ambitiousness, attain, etc.
	Neiman et al., 2016a, 2016b	accomplishment, achievement, attainment, competence, completion, finishing, fulfillment, goal, job, prevail, proficiency, prosperity, succeed, success, triumph, victory, win
Affiliation	LIWC 2015	accompan*, accomplice*, affil*, alliance*, allies, ally, amigo*, associate, associates, associating, association, associations, bae, banter*, belong*, bestfriend*, bf, bff*, bfs, boyfriend*, etc.
	Harvard IV	abide, absorption, accede, acceptance, accompany, acquaintance, adherence, adherent, admire, adopt, advice, advise, advocacy, affair, affection, affectionate, affiliate, affiliation, affix, agree, etc.
Certainty	LIWC 2015	absolute, absolutely, accura*, all, altogether, always, apparent, assur*, blatant*, certain*, clear, clearly, commit, commitment*, commits, committed, committing, complete, completed, completely, etc.
	Neiman et al., 2016a, 2016b	ambiguity, confusion, disorder, doubt, hesitation, improbability, indecision, putoff, uncertainty, underlying, vagueness
Power	LIWC 2015	a-list*, above, acclaimed, administr*, age, allow*, amateur*, ambition, ambitions, ambitious, apolog*, apprentic*, approv*, armies, army, asham*, assault*, assertive, attack*, attendant, etc.
	Harvard IV	abolish, accomplish, accomplishment, accord, achievement, adjudication, administer, administration, administrative, administrator, admit, admonish, adviser, advisor, advocate, affirm, afford, agency, alliance, allot, etc.
	Neiman et al., 2016a, 2016b	authority, be in charge of, be in command of, capital, charge, control, dominance, exploit, governance, have power over, jurisdiction, manage, order, position, power, punish, rank, regulation, rule, stature, status, supremacy

Table A2
Table of Correlations of Scores Derived From Different Dictionaries Used to Estimate the Same Psychological Construct

Category	Operationalization	1	2
Achievement	1. LIWC 2015		
	2. Neiman et al., 2016a, 2016b	.246	
Affiliation	1. LIWC 2015		
	2. Harvard IV	.352	
Certainty	1. LIWC 2015		
	2. Neiman et al., 2016a, 2016b	.004	
Power	1. LIWC 2015		
	2. Harvard IV	.300	
	3. Neiman et al., 2016a, 2016b	.156	.161

(*Appendices continue*)

Appendix B

Details of the Enrichment Analysis

Table B1
Components of the Enrichment Equation: An Illustration

Variable	Tweets with relevant words	Tweets without relevant words	Total
Within bin	*a*	*b*	
Outside bin	*c*	*d*	
Total			*N*

Note. Multiple hypothesis corrected alpha: $\alpha_{ew} = \frac{\alpha_{pc}}{b*l} = \frac{.05}{50*27} =$ .0037.

Enrichment equation (hypergeometric distribution):

$$p(a \mid b, \quad c, \quad d) = \frac{\binom{a+b}{a}\binom{c+d}{c}}{\binom{N}{a+c}}$$

Binning Procedure

The dataset was formatted such that each row was a tweet. To construct the bins, we first ordered each tweet based on the political ideology estimates of the twitter users. We then split the sorted dataset into 50 equal sized bins (each bin containing 240,000 tweets and at minimum 75 users). Enrichment analyses were conducted on each bin individually.

In the initial binning of ideologically ranked tweets there are some Twitter users who have tweets that are divided between two different bins. To ensure that this was not biasing our results or over weighting active users, we sorted the dataset by ideology again and created bins with as close to 240,000 tweets as possible without dividing users among multiple bins. Specifically, we added users consecutively based on their ideology rankings to a bin until the number of tweets in the bin exceeded 240,000. When adding a user's collection of tweets to a bin exceed the bin's limit, we sorted that user into a new bin and continued adding tweets grouped by user until that bin's tweet count exceeded 240,000. Conducting the analyses in this way did not alter the enrichment analysis in any substantive manner (see alternative binning code here: https://github.com/jlsterling/Dissertation_code).

Linear expected values computation:

$$h(P_{lin}) = \beta_0 + \beta_1 I_b + \beta_2 ln(\overline{WC_b})$$

$$h(P_{lin}) = -\frac{1}{(1 + \exp(P_{lin})}$$

Converting to odds ratio:

$$Odds_{lin} = \frac{h(P_{lin})_b}{E(h(P_{lin})_b)}$$

Quadratic expected values computation:

$$h(P_{quad}) = \beta_0 + \beta_1 I_b + \beta_2 I_b^2 + \beta_3 ln(\overline{WC_b})$$

$$h(P_{quad}) = -\frac{1}{(1 + \exp(P_{quad})}$$

Converting to odds ratio:

$$Odds_{lin} = \frac{h(P_{quad})_b}{E(h(P_{quad})_b)}$$

Table B2
Definitions

b	Number of bins
l	Number of language predictions
P_{lin}	Expected proportion from linear model of tweets containing relevant words within bin
I	Average ideology estimate per bin
WC	Average word count per bin
P_{quad}	Expected proportion of tweets from quadratic model containing relevant words within bin

(*Appendices continue*)

Appendix C

Table C1
Counts of Tweets Containing at Least One Dictionary Word by User Ideology

Language aspect	Liberals	Moderates	Conservatives
Dispositional motives			
Affiliation V1	1,133,867	1,172,680	633,617
Affiliation V2	1,629,348	1,618,233	915,169
Power V1	1,243,373	1,277,973	760,917
Power V2	1,270,093	1,269,250	757,121
Personal values			
Conformity	553	479	233
Power	49,015	52,734	31,698
Security	25,207	25,392	20,104
Tradition	46,684	50,211	33,625
Achievement V1	783,513	827,888	445,422
Achievement V2	86,711	99,433	55,295
Benevolence	7,068	8,102	3,266
Universalism	2,266	2,224	1,403
Stimulation	9,386	8,008	3,661
Self-direction	114,501	105,516	62,470
Social	3,063,800	3,061,074	1,720,382
Motivated social cognition			
Certainty V1	648,108	604,908	366,572
Certainty V2	4,601	4,529	3,234
Resistance to change	70,958	70,534	43,557
Inequality	77,286	79,106	44,386
Past focus	975,942	893,637	535,812
Tentative	825,066	751,997	463,248
Future focus	629,462	617,938	374,232
Anxiety	135,329	126,282	79,757
Inhibition	316,085	303,038	193,333
Threat	53,978	56,839	36,170
Risk focus	271,692	264,279	171,375
Death	134,403	132,613	84,633
Uniqueness and conformity motivations			
I	1,198,933	1,054,603	587,902
We	382,063	384,325	226,211
Emotion			
Positive emotion	1,903,774	1,892,755	1,032,032
Negative emotion	943,842	852,596	544,484
Anger	387,479	334,032	230,107
Anxiety	135,329	126,282	79,757
Total number of tweets	4,576,614	4,647,421	2,479,615

Received August 14, 2018
Revision received October 13, 2019
Accepted October 24, 2019 ■