

# Linguistic Markers of Grandiose Narcissism: A LIWC Analysis of 15 Samples

In press at:

*Journal of Language and Social Psychology*

The final published version may differ slightly.

Nicholas S. Holtzman<sup>1</sup>, Allison M. Tackman<sup>2</sup>, Angela L. Carey<sup>2</sup>,  
Melanie S. Brucks<sup>3</sup>, Albrecht C. P. Kufner<sup>4</sup>, Fenne große Deters<sup>5</sup>, Mitja  
D. Back<sup>4</sup>, M. Brent Donnellan<sup>6</sup>, James W. Pennebaker<sup>7</sup>, Ryne  
A. Sherman<sup>8</sup>, and Matthias R. Mehl<sup>2</sup>

## Abstract

Narcissism is unrelated to using first-person singular pronouns. Whether narcissism is linked to other language use remains unclear. We aimed to identify linguistic markers of narcissism. We applied the Linguistic Inquiry and Word Count to texts ( $k = 15$ ;  $N = 4,941$ ). The strongest positive correlates were: using words related to sports, second-person pronouns, and swear words. The strongest negative correlates were: using anxiety/fear words, tentative words, and words related to sensory/perceptual processes. Effects were small (each  $|r| < .10$ ).

## Keywords

Language, LIWC, narcissism, personality, text analysis

---

<sup>1</sup>Georgia Southern University

<sup>2</sup>University of Arizona

<sup>3</sup>Columbia University

<sup>4</sup>University of Münster

<sup>5</sup>University of Potsdam

<sup>6</sup>Michigan State University

<sup>7</sup>University of Texas at Austin

<sup>8</sup>Hogan Assessment Systems

## Corresponding Author:

Nicholas S. Holtzman, Department of Psychology, P.O. Box 8041; Georgia Southern University; Statesboro, GA 30460-8041

Email: [nick.holtzman@gmail.com](mailto:nick.holtzman@gmail.com)

Our words often betray who we are. Personality psychology has identified associations between personality traits—such as openness, conscientiousness, extraversion, agreeableness, and neuroticism—and word use (Fast & Funder, 2008; Ireland & Mehl, 2014; Schwartz et al., 2013; Yarkoni, 2010). Relatively little is known, however, about the degree to which narcissism reveals itself through word use. With a few choice exceptions (Bogolyubova, Panicheva, Tikhonov, Ivanov, & Ledovaya, 2018; Preotiuc-Pietro, Carpenter, Giorgi, & Ungar, 2016; Sumner, Byers, Boochever, & Park, 2012), studies on narcissism and language have been conducted on small samples (e.g., Holtzman, Vazire, & Mehl, 2010). Indeed, even some of the larger samples might still be considered underpowered (Preotiuc-Pietro, et al., 2016), given the modest effect sizes for psychological language markers. Accordingly, the goal of this paper was to use a collection of datasets, which we used previously (Carey et al., 2015), to provide a clearer answer as to whether there are any word patterns that can help identify narcissistic individuals.

We use the Linguistic Inquiry and Word Count (LIWC) program for our analyses, as it is currently one of the most commonly used and best validated approaches to studying word use in social/personality psychology. This program was developed by James Pennebaker and colleagues (Pennebaker, Francis, & Booth, 2001) and offers an efficient way to code text files on a number of dimensions. We chose to use the LIWC 2001 dictionary which counts how many times words from 87 categories have appeared in a text file. In the end, each text file is assigned a set of 87 scores, each of which indicate the percent of occurrence relative to the total number of words in the file. Using this technology, we aim to answer the following research question.

### **Research Question: Are there Robust LIWC Correlates of Narcissism?**

The key question is whether we can identify a robust set of LIWC correlates of grandiose narcissism. This is a controversial question, as a recent study of Twitter data (final  $N = 491$ )

suggests that there is not a robust set of LIWC correlates of narcissism (Preotiuc-Pietro, et al., 2016); the prediction equation led to a multiple  $R$  of .15 in their study (using the set of LIWC variables as predictors). The study provided a first glimpse at the LIWC profile of narcissism. The study was limited in sample size and generalizability (Twitter users only). Thus, the Preotiuc-Pietro and colleagues (2016) study provides a solid starting point for answering this research question, while it has some limitations that our study seeks to address.

We take an explicitly exploratory approach based on a large amount of text data derived from various communication contexts. We correlate narcissism with each of the LIWC 2001 categories, excluding the “I” category, which we studied in our previous publication on narcissism (Carey et al., 2015), and excluding the punctuation categories because punctuation can be arbitrary in verbatim transcripts of spoken language; we also excluded emoticons because they do not appear in (transcribed) spoken language. Thus, after the exclusions had been accounted for, we ran correlations with 72 LIWC categories. Our exploratory approach has the additional advantage of minimizing HARKing—hypothesizing after the results are known (Kerr, 1998).

There is reason to believe that narcissism might have detectable LIWC correlates, as there is evidence that theoretically and empirically related traits including extraversion and disagreeableness have robust LIWC correlates (Yarkoni, 2010). However, one might expect these correlations to be modest in magnitude for at least two reasons—one statistical and the other substantive. The statistical reason one might not expect large correlations is because effect sizes tend to be modest in studies that are adequately powered (Ioannidis, 2005); such effects discovered by large collaborative groups are presumably less influenced by publication bias than small studies—studies that often produce larger effects, likely due to publication bias. The substantive reason one might not expect large correlations is because narcissistic people can be versatile and

amorphous (i.e., quick to change) in their social presentation, potentially leading to an unstable language profile. In our study, we do not expect to find many correlations for which  $|r| > .10$ . Finally, in the supplemental materials, we also show results by gender.

## **Overview of Aims and Scope**

In sum, our objective is to test for significant LIWC correlates of narcissism. Ultimately, our study was motivated by the goal to find a stable LIWC-based linguistic profile for grandiose narcissism that could then be integrated into, and interpreted on the background of, the existing LIWC research (Ireland & Mehl, 2014; Tausczik & Pennebaker, 2010).

## *Method*

The data that we used to conduct the meta-analysis in this paper are posted on the Open Science Framework (<https://osf.io/caf8n/files/>). The raw data are the same as what we used in Carey and colleagues (2015). Methodological differences include: First, in the Carey study, the data were analyzed at the level of the individual. In this study, we analyzed the data at the level of the aggregate sample statistic and then we meta-analytically synthesized the results. Prior research has shown that analysis at the individual level and at the sample level produce similar results (Steinberg et al., 1997). Second, in the Carey study, the data were analyzed with context as a moderator; here, we refrain from doing so to not further exacerbate the multiple comparison problem.

*Samples and Procedures.* Information about the samples is summarized in Table 1. Sample sizes were determined by convenience and were based on logistical constraints. These samples are fully described in Carey and colleagues (2015), and therefore they are not described here. Minor differences in sample sizes were evident, and this most certainly occurred because of the massive data management effort that this project entailed.

**Table 1.** Sample Descriptions and Narcissism POMP Scores.

Sample	Task	Lang.	Measure	$\alpha$	Total Sample			Females			Males		
					<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
1	1.1 Video-taped self-descriptions	G	NPI 15	.75	101	25.48	19.33	78	25.30	19.29	23	26.09	19.88
2	2.1 Describe yourself to group members	G	NPI 40	.78	68	35.96	14.70	34	32.35	11.84	34	39.56	16.50
3	3.1 Introduce yourself to group members	G	NPI 40	.82	340	36.42	15.76	233	35.06	15.09	107	39.37	16.83
4	4.1 Write down specific attributes you have 4.2 Write down goals 4.3 Video-taped self-descriptions	G	NPI 40	.84	73	39.90	17.03	52	38.22	16.93	21	44.05	16.98
5	5.1 Introduce yourself into camera	G	NPI 40	.82	133	36.20	16.09	77	34.94	15.64	53	38.63	16.87
6	6.1 "About me" on German social media profile	G	NPI 15	.84	54	35.75	19.20	44	35.69	19.05	10	36.00	20.89
7	7.1 Direct replication of Raskin & Shaw (1988)	E	NPI 40	.83	242	40.65	16.63	199	40.17	16.52	43	42.88	17.19
8	8.1 Self-defining memory essay 8.2 High point in life essay 8.3 Low point in life essay 8.4 Describe yourself essay	E	NPI 40	.82	312	40.89	16.23	208	40.02	15.44	103	42.88	17.54
9	9.1 Important memory or life episode essay 9.2 Important moment in romantic life essay 9.3 Person you care most about essay 9.4 Recent problem essay 9.5 Like most about yourself essay 9.6 Like least about yourself essay	E	NPI 40	.84	453	42.59	17.12	326	41.79	16.89	124	44.74	17.71
10	10.1 Facebook status updates	G	NPI 15	.70	74	32.61	19.18	57	30.39	17.97	17	40.06	21.71
11	11.1 Facebook status updates	E	NPI 40	.83	132	45.94	17.53	84	42.00	15.39	48	52.83	19.01
12	12.1 Stream of consciousness task 12.2 Thematic apperception test	E	NPI 16	.70	751	34.55	19.45	472	31.97	18.22	279	38.91	20.68
13	13.1 Stream of consciousness task	E	NPI 16	.65	921	31.20	17.96	562	28.89	17.40	359	34.81	18.24
14	14.1 Stream of consciousness task 14.2 Thematic apperception test 14.3 "Who am I?" writing task	E	Dirty 12	.74	1,210	67.54	14.62	751	66.08	14.95		70.20	13.54
15	15.1 Facebook status updates	E	NPI 40	.77	77	39.22	14.4	60	37.54	14.28	17	45.15	13.59

Note. Lang. = Language; G = German; E = English; NPI = Narcissistic Personality Inventory; Dirty 12 = Dirty Dozen;  $\alpha$  = Cronbach's alpha. Narcissism Scores are in POMP units, reflecting percentage values relative to the maximum possible score (Cohen, Cohen, Aiken, & West, 1999); this facilitates comparisons across samples.

*Measures.* Important information about the study measures is also summarized in Table 1. The 40-item Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988) was administered in the original English version in Sample 7, Sample 8, Sample 9, Sample 11, and Sample 15. The 16-item NPI (Ames, Rose, & Anderson, 2006) was administered in Sample 12 and Sample 13 and it uses select items from the NPI-40. The German version of the 40-item NPI (Schütz, Marcus, & Sellin, 2004) was administered in Sample 2, Sample 3, Sample 4, and Sample 5. The German 15-item NPI version (Schütz, et al., 2004) was administered in Sample 1, Sample 6, and Sample 10. Finally, the Dirty Dozen (Jonason & Webster, 2010), a newer 12-item measure of the Dark Triad, was administered in Sample 14.

We chose to use LIWC 2001 to analyze the data for all 15 data sets. All videotaped introductions and interactions were transcribed and saved as text documents. All written documents were saved as text documents and, where sample size permitted, were manually cleaned and spell checked. The resulting text documents from English samples were submitted to the English version of LIWC 2001 (Pennebaker, et al., 2001). The resulting text documents from the German samples were submitted to the psychometrically validated German LIWC 2001 dictionary (Wolf et al., 2008). For the purpose of this study, we focused on how narcissism is related to 72 LIWC categories. All LIWC measures are in a percentage-based metric.

*Data Analytic Strategy.* First, we aimed to arrive at one LIWC category score (in raw form) per participant. Thus, in the datasets in which there were multiple language tasks, we ran LIWC on each task, and then we averaged the LIWC scores across the tasks. Next, we correlated the LIWC scores with narcissism using weighted Pearson's correlations (Bills & Li, 2005); we weighted each participant by the number of words provided. A weighting method was selected because some participants provided much more linguistic information (i.e., more words) than others. The

advantage of this method is that it differentially weights participants who have provided more information (in some cases, hundreds more words), and thus it appropriately down-weights cases that are based on less text data.

Next, we aimed to aggregate the correlations across the 15 data sets. To do so, we used the *metaphor* package for meta-analysis in R (Viechtbauer, 2010); we used the fixed effects model because we aimed to determine the effect sizes for this set of 15 studies, rather than make inferences about broader sets of studies in the population (Viechtbauer, 2010). We ran this process three times: for the total sample, for females, and for males. Given the large number of effects that we report, we focus on the effect sizes, rather than modeling the variability among the effects. We computed 95% confidence intervals for each point estimate. To enhance readability (i.e., to avoid displaying the numerous LIWC effects), we focus on the top 15 (most positive) and bottom 15 (most negative) LIWC correlates of narcissism.

*Descriptive Statistics.* The means and standard deviations for the narcissism measures are shown in Table 1, and the means and standard deviations for the 72 LIWC word categories, by sample, are listed in the Supplemental Material – LIWC Descriptives (AKA Supple~2.XLS at <https://osf.io/caf8n/files/>). Note that some of the cells are empty because of differences between the English and German dictionaries.

## **Results**

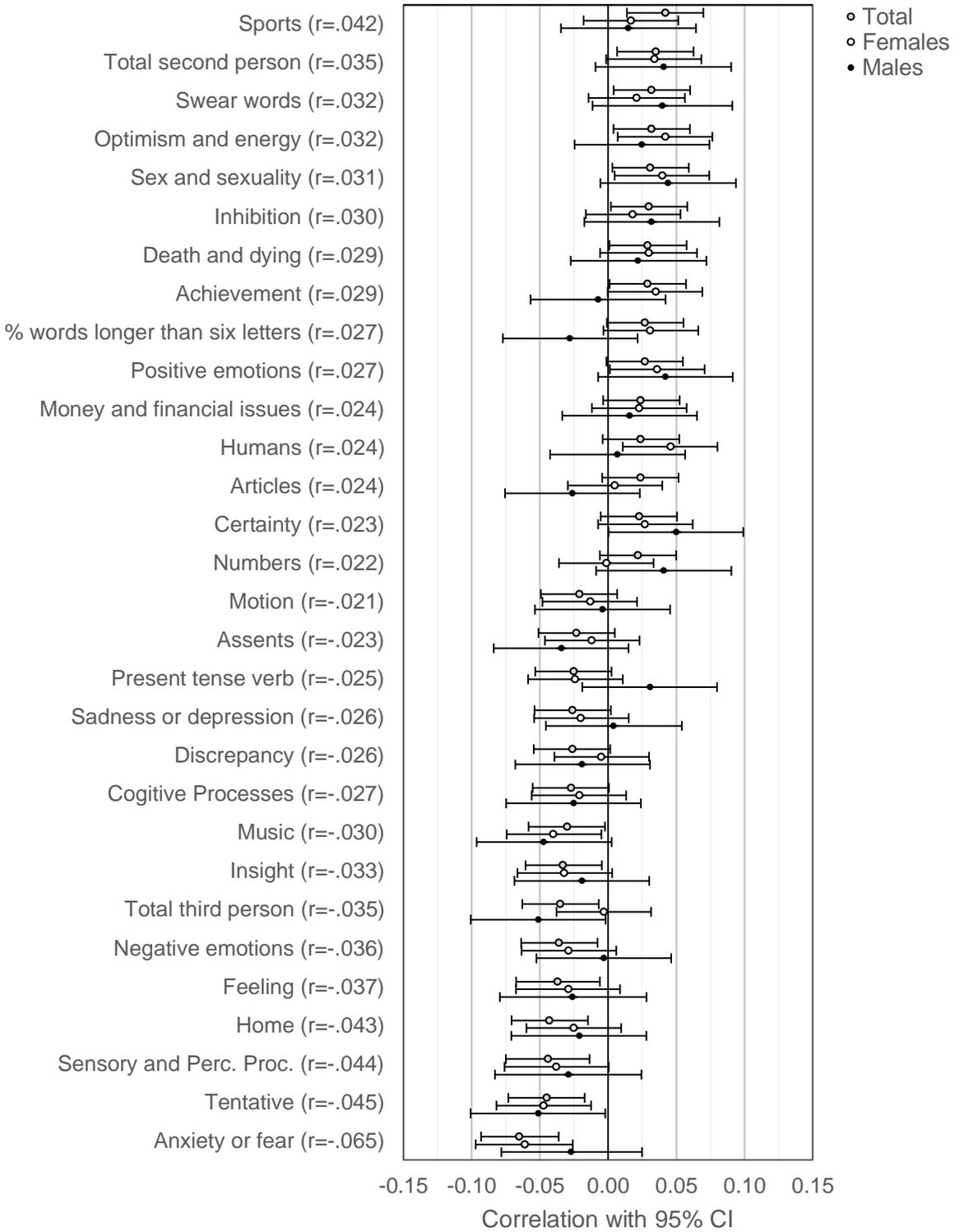
*Testing the Overall Effect.* To estimate the degree to which narcissism is associated with relative word frequencies in the LIWC dictionary in general, we conducted a randomization test following the method prescribed by Sherman and Funder (2009). First, we calculated the weighted correlations between narcissism and 72 LIWC categories, weighted by word count, for each of the 15 datasets. (The word count category was not weighted by word count as it did not make sense to

weight that variable by itself.) Second, we meta-analytically combined the weighted correlations across the studies, weighting by sample size, to get 72 correlations between narcissism and LIWC. Third, we computed the observed average absolute association between narcissism and LIWC ( $\overline{|r|} = .0194$ ). This value serves as a gauge for the overall association between narcissism and LIWC.

To determine if this value is larger than one would expect under a random model of no association between narcissism and LIWC, we did the following 100,000 times: (1) within each data set we randomly re-assigned narcissism scores to participants (i.e., shuffled the narcissism scores); (2) we repeated the steps just previously described to calculate an average absolute association between narcissism and LIWC. We saved the values of each of these 100,000 randomizations as a sampling distribution for our observed value. For only 3 of the 100,000 randomizations did the randomizations yield a value greater than or equal to the observed absolute effect size value. Thus, the probability of the observed relationship between narcissism and LIWC word frequencies is estimated to be 3 in 100,000 ( $p = .00003$ ). This constitutes evidence that overall, narcissism has real (i.e., beyond chance) relationships to word frequencies (i.e., across all tested LIWC variables).

*LIWC Profile of Narcissism.* Next, a meta-analysis of the 15 studies was conducted on the weighted correlations. (See the Supplemental Material – Weighted Correlations by Sample; AKA Supple~3.XLS at <https://osf.io/caf8n/files/>). The top 15 (most positive) and bottom 15 (most negative) LIWC correlates of narcissism are displayed in Figure 1; we display the effects for the total sample as well as for females only and for males only. The full sets of correlations for all LIWC variables for the total sample, for females only, and for males only, are available in the Supplemental Material – Figure of Overall Narcissism Effects (AKA Supple~4.XLS at <https://osf.io/caf8n/files/>). In the total sample (i.e., ignoring gender), 17 of the 72 effects were

statistically significant at the  $p < .05$  level; seven were statistically significant at the more stringent  $p < .01$  level; five were significant at the even more stringent  $p < .005$  level (Benjamin et al., 2018). Narcissism was most positively correlated with using sports-related words ( $r = .042, p = .003, 95\%$  CI [.014; .070]), second person pronouns such as “you” ( $r = .035, p = .02, 95\%$  CI [.007; .063]), and swear words ( $r = .032, p = .025, 95\%$  CI [.004; .060]). Narcissism was most negatively related to using “anxiety or fear” words ( $r = -.065, p < .001, 95\%$  CI [-.093; -.036]), tentative words ( $r = -.045, p = .002, 95\%$  CI [-.073; -.017]), and senses words ( $r = -.044, p = .005, 95\%$  CI [-.075; -.014]).



**Figure 1.** Correlation coefficients with 95% confidence intervals for the top 15 (most positive) and bottom 15 (most negative) LIWC correlates of narcissism. Parenthetical  $r$ s are for Total  $r$ s.

When applying a full Bonferroni correction for the entire study,  $(.05/[3*72]) = .0002$ , the only statistically significant finding that emerged was the negative relationship between narcissism and the LIWC category “anxiety or fear” ( $p < .0001$ ); none of the other effects, including the gender-specific effects (which are available in a file called Supplemental Material – Gender Differences; AKA Supple~1.XLS at <https://osf.io/caf8n/files/>) were significant at this maximally stringent level, although this stringent level is quite likely too conservative (e.g., García, 2004). Readers interested in a narrative about the gender associations are directed to the online Word document called Supplemental Material – Gender Based Profiles of Correlations, which is stored at the *JLSP* site.

### *Discussion*

Our exploratory analysis revealed several correlations based on the LIWC categories, and our randomization test revealed evidence that overall LIWC categories are related to narcissism. These effects tended to be small. Our Discussion focuses on four significant correlations in the total sample that could be viewed as particularly theoretically and empirically interesting: the positive association between narcissism and sexual word use, the positive association between narcissism and swear word use, the negative association between narcissism and tentative word use, and the negative association between narcissism and anxiety- or fear-related words.

First, consistent with empirical research (e.g., Holtzman, et al., 2010; Sumner, et al., 2012) and theory (Buss & Shackelford, 1997), narcissism was associated with using more sexual words. The idea is that narcissistic people will use sexual language to create a sexualized environment, perhaps as a means to signaling their own sexual availability or to prime sexual concepts in the minds of sexually-available others.

Second, narcissism was associated with using more swear words (cf. Feldman, Lian, Kosinski, & Stillwell, 2017). One explanation of this effect is that narcissistic people tend to be disagreeable (Paulhus, 1998), and one linguistic marker of disagreeableness is swearing (Kern et al., 2014, Figure 1; Mehl, Gosling, & Pennebaker, 2006). An alternative way to think about this is that it may be part of the general tendency for narcissistic people to seek attention, even if it is for socially unacceptable behavior. In line with this idea, Oscar Wilde once wrote that “there is only one thing in the world worse than being talked about, and that is not being talked about” (Wilde, 1890/2014, pp. 4-5). Narcissistic people tend to do what it takes to be talked about.

Third, narcissism was associated with using fewer tentative words. Narcissistic people tend to be dominant (Cheng, Tracy, & Henrich, 2010) and exploitative (Raskin & Terry, 1988); in general, they are self-assured and confident. It is a rare grandiose narcissistic individual who is submissive and tentative about things. So, perhaps it is unsurprising that narcissistic people are less tentative in their language (i.e., rarely using words like “maybe”, “perhaps”, and “guess”). This is consistent with previous work on narcissism and tentative word use (Sumner, et al., 2012).

Fourth, we also found that narcissism was negatively related to using words about anxiety and fear (e.g., “afraid”, “distracted”, “horror”). This was the most robust effect in the analysis. One interpretation is that narcissistic people—at least grandiose narcissistic people—actually experience little fear and anxiety. Another way to interpret this finding is that narcissistic people may experience some anxiety and fear, but these emotions may not register consciously and so they do not get expressed in language; this latter interpretation is consistent with the perspective that grandiose narcissism masks underlying vulnerabilities (Cheng, Tracy, & Miller, 2013). While our data cannot resolve this ongoing debate about whether narcissistic people really are vulnerable “deep down” (Campbell, Bosson, Goheen, Lakey, & Kernis, 2007), the data do point to this

interesting notion: Narcissistic people appear not to verbalize indicators of vulnerability such as anxiety and fear.

All in all, grandiose narcissism manifests in a linguistic profile that appears certain, confident, and perhaps low in vulnerability. Relatedly, previous research has highlighted the link between narcissism and invulnerability (Aalsma, Lapsley, & Flannery, 2006; Barry, Pickard, & Ansel, 2009). A cocky, tough, and sure-minded language profile seems to integrate many of these findings. Moreover, narcissistic people present as sexually explicit and prone to using profanity (Adams, Fiorell, Burton, & Hart, 2014); conversely, they tend not to use words reflecting uncertainty, anxiety, and fear. Collectively, this set of linguistic manifestations points to an attention-getting personality that is (at least on the surface) confident and certain.

## **Conclusions**

In sum, we have shown that narcissism has associations with select LIWC word categories, several of which were significant at stringent levels. This paper should be valuable for future studies as it suggests the LIWC correlates of narcissism (and other broad personality attributes) are modest in magnitude, so large sample sizes are needed when planning studies. Our lack of success at revealing pronounced linguistic patterns using the dictionary-based LIWC approach calls for future research using emerging text analytic methods (e.g., topic models) that are not based on pre-determined—and thereby from a measurement perspective likely suboptimal—word lists.

## **Acknowledgements**

We thank two anonymous reviewers, David Funder, and Howie Giles (the *JLSP* editor-in-chief), for helping to make this paper better.

## **Declaration of Conflicting Interests**

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

## **Funding**

None

**References**

- Aalsma, M. C., Lapsley, D. K., & Flannery, D. J. (2006). Personal fables, narcissism, and adolescent adjustment. *Psychology in the Schools, 43*, 481-491. doi: 10.1002/pits.20162
- Adams, J. M., Fiorell, D., Burton, K. A., & Hart, W. (2014). Why do narcissists disregard social-etiquette norms? A test of two explanations for why narcissism relates to offensive-language use. *Personality and Individual Differences, 58*, 26-30. doi: 10.1016/j.paid.2013.09.027
- Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality, 40*, 440-450. doi: 10.1016/j.jrp.2005.03.002
- Barry, C. T., Pickard, J. D., & Ansel, L. L. (2009). The associations of adolescent invulnerability and narcissism with problem behaviors. *Personality and Individual Differences, 47*, 577-582. doi: 10.1016/j.paid.2009.05.022
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2*, 6-10. doi: 10.1038/s41562-017-0189-z
- Bills, C. B., & Li, G. (2005). Correlating homicide and suicide. *International Journal of Epidemiology, 34*, 837-845. doi: 10.1093/ije/dyi079
- Bogolyubova, O., Panicheva, P., Tikhonov, R., Ivanov, V., & Ledovaya, Y. (2018). Dark personalities on Facebook: Harmful online behaviors and language. *Computers in Human Behavior, 78*, 151-159. doi: 10.1016/j.chb.2017.09.032
- Buss, D. M., & Shackelford, T. K. (1997). Susceptibility to infidelity in the first year of marriage. *Journal of Research in Personality, 31*, 193-221.

- Campbell, W. K., Bosson, J. K., Goheen, T. W., Lakey, C. E., & Kernis, M. H. (2007). Do narcissists dislike themselves “deep down inside”? *Psychological Science, 18*, 227-229.
- Carey, A. L., Brucks, M. S., Kufner, A. C. P., Holtzman, N. S., Deters, F. G., Back, M. D., Donnellan, M. B., Pennebaker, J. W., & Mehl, M. R. (2015). Narcissism and the use of personal pronouns revisited. *Journal of Personality and Social Psychology, 109*, E1-E15. doi: 10.1037/pspp0000029
- Cheng, J. T., Tracy, J. L., & Henrich, J. (2010). Pride, personality, and the evolutionary foundations of human social status. *Evolution and Human Behavior, 31*, 334-347. doi: 10.1016/j.evolhumbehav.2010.02.004
- Cheng, J. T., Tracy, J. L., & Miller, G. E. (2013). Are narcissists hardy or vulnerable? The role of narcissism in the production of stress-related biomarkers in response to emotional distress. *Emotion, 13*, 1004-1011. doi: 10.1037/a0034410
- Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology, 94*, 334-346. doi: 10.1037/0022-3514.94.2.334
- Feldman, G., Lian, H., Kosinski, M., & Stillwell, D. (2017). Frankly, we do give a damn: The relationship between profanity and honesty. *Social Psychological and Personality Science, 8*, 816-826. doi: 10.1177/1948550616681055
- García, L. V. (2004). Escaping the Bonferroni iron claw in ecological studies. *Oikos, 105*, 657-663. doi: 10.1111/j.0030-1299.2004.13046.x
- Holtzman, N. S., Vazire, S., & Mehl, M. R. (2010). Sounds like a narcissist: Behavioral manifestations of narcissism in everyday life. *Journal of Research in Personality, 44*, 478-484. doi: 10.1016/j.jrp.2010.06.001

- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, 2, 696-701. doi: 10.1371/journal.pmed.0020124
- Ireland, M. E., & Mehl, M. R. (2014). Natural language use as a marker of personality. In T. Holtgraves (Ed.), *The Oxford handbook of language and social psychology* (pp. 201-237). New York, NY: Oxford University Press. doi: 10.1093/oxfordhb/9780199838639.013.034
- Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, 22, 420-432. doi: 10.1037/a0019265
- Kern, M. L., Eichstaedt, J. C., Schwartz, H. A., Dziurzynski, L., Ungar, L. H., Stillwell, D. J., Kosinski, M., Ramones, S. M., & Seligman, M. E. P. (2014). The online social self An open vocabulary approach to personality. *Assessment*, 21, 158-169. doi: 10.1177/1073191113514104
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2, 196-217.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, 90, 862-877. doi: 10.1037/0022-3514.90.5.862
- Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, 74, 1197-1208. doi: 10.1037/0022-3514.74.5.1197
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001 [Computer software]. Mahway, NJ: Erlbaum. Retrieved From: <http://www.liwc.net/>

Preotiuc-Pietro, D., Carpenter, J., Giorgi, S., & Ungar, L. (2016). *Studying the Dark Triad of personality through Twitter behavior*. Paper presented at the Proceedings of the 25th ACM International Conference on Information and Knowledge Management. doi: <http://dx.doi.org/10.1145/2983323.2983822>

Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology*, *54*, 890-902. doi: 10.1037/0022-3514.54.5.890

Schütz, A., Marcus, B., & Sellin, I. (2004). Measuring narcissism as a personality construct: Psychometric properties of a long and a short version of the German Narcissistic Personality Inventory. *Diagnostica*, *50*, 202-218. doi: 10.1026/0012-1924.50.4.202

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., & Seligman, M. E. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLOS ONE*, *8*, e73791. doi: 10.1371/journal.pone.0073791

Sherman, R. A., & Funder, D. C. (2009). Evaluating correlations in studies of personality and behavior: Beyond the number of significant findings to be expected by chance. *Journal of Research in Personality*, *43*, 1053-1063. doi: 10.1016/j.jrp.2009.05.010

Steinberg, K., Smith, S., Stroup, D., Olkin, I., Lee, N., Williamson, G., & Thacker, S. (1997). Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. *American Journal of Epidemiology*, *145*, 917-925.

- Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). *Predicting dark triad personality traits from Twitter usage and a linguistic analysis of tweets*. Paper presented at the International Conference on Machine Learning and Applications.
- 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. doi: 10.1177/0261927x09351676
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1-48.
- Wilde, O. (1890/2014). *The picture of Dorian Gray*. Brooklyn, NY: Millennium Publications.
- Wolf, M., Horn, A. B., Mehl, M. R., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computer-aided quantitative textanalysis: Equivalence and reliability of the German adaptation of the Linguistic Inquiry and Word Count. *Diagnostica, 54*, 85-98. doi: 10.1026/0012-1924.54.2.85
- Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44*, 363-373. doi: 10.1016/j.jrp.2010.04.001

## **Author Biographies**

**Nicholas S. Holtzman, Ph.D.** is the lead author of this paper. He is an associate professor of psychology at Georgia Southern University in Statesboro, GA, U.S.A., and his research interests revolve around socially aversive personality traits—where they come from, how they manifest, and what consequences they have.

**Matthias R. Mehl, Ph.D.** is the senior author of this paper. He is a professor of psychology at the University of Arizona, Tucson, AZ, U.S.A., where he also holds an affiliate appointment in the department of communications. His research aims at understanding the psychological implications of people's daily social lives and everyday language use.

**SUPPLEMENTARY FILE FOR:****Linguistic Markers of Grandiose Narcissism:  
An Exploratory LIWC Analysis of 15 Samples****Are there Gender Differences in the LIWC Correlates of Narcissism?**

The second research question we intend to answer is whether there are consistencies across men and women in the profiles of correlations. Prior research in this area has shown that there are gender differences in the language correlates of narcissism (Fast & Funder, 2010); in particular, it has been demonstrated that for men, some facets of narcissism are positively correlated with first-person singular pronoun use, whereas for women this is not the case. Although follow-up research generally did not support this pattern with regard to first-person singular pronouns (Carey et al., 2015), it is important to consider the possibility that it could be relevant to a broader study of LIWC correlates of narcissism. Thus, we explore gender differences in the LIWC correlates of narcissism in our study.

**Correlations by Gender and a Test for Between-Correlation Gender Differences**

Among females, narcissism was most positively correlated with the “humans” category ( $r = .046$ ,  $p = .01$ , 95% CI [.011; .080]), and it was most negatively correlated with “anxiety or fear” words ( $r = -.061$ ,  $p < .001$ , 95% CI [-.097; -.026]). Among males, narcissism was most positively correlated with certainty ( $r = .050$ ,  $p = .048$ , 95% CI [.0004; .0992]), and it was most negatively correlated with tentative words ( $r = -.051$ ,  $p = .04$ , 95% CI [-.100; -.002]).

Next, we examined gender differences for each specific LIWC category; the full results are available in the Supplemental Material – Gender Differences (AKA Supple~1.XLS at

<https://osf.io/caf8n/files/>). Using the equation 2.8.5 from Cohen and Cohen (1983), which compares the magnitudes of correlations from independent samples, we found two gender differences that were significant ( $p < .05$ ). The two were as follows: The correlation between narcissism and word count was greater for women than men ( $Z = 2.204, p = .028$ ); and the correlation between narcissism and percent of words captured as LIWC dictionary words was greater for men than women ( $Z = 2.030, p = .042$ ). Overall, the number of gender differences we found is fewer than what can be expected by chance alone, given that we ran 72 statistical tests (at chance, 3.6 out of 72 would be significant at  $p < .05$  [under the null hypothesis]), and thus these two highlighted differences probably do not reflect real differences.

To determine whether the profile for females was similar to the profile for males, we ran a column-vector correlation (Funder & Sneed, 1993). This approach takes the column of correlations for females, the column of correlations for males, and then correlates them (after using the appropriate  $r$ -to- $z$  transformation); it indexes the degree of similarity between the profiles. In this case, the correlation was small for a profile correlation ( $r = .321$ ), indicating only modest similarity between the females and the males with regard to LIWC correlates of narcissism. It is worth mentioning that because the 72 correlations hovered around zero, this pattern may have stifled our ability to detect a large profile correlation, and thus it is possible that the profile correlation reported here is underestimated.

**References (for Supplementary File)**

Carey, A. L., Brucks, M. S., Kufner, A. C. P., Holtzman, N. S., Deters, F. G., Back, M. D.,

Donnellan, M. B., Pennebaker, J. W., & Mehl, M. R. (2015). Narcissism and the use of personal pronouns revisited. *Journal of Personality and Social Psychology, 109*, E1-E15.

doi: 10.1037/pspp0000029

Cohen, J., & Cohen, P. (1983). *Multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Earlbaum.

Fast, L. A., & Funder, D. C. (2010). Gender differences in the correlates of self-referent word use: Authority, entitlement, and depressive symptoms. *Journal of Personality, 78*, 313-338. doi: 10.1111/j.1467-6494.2009.00617.x

Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64*, 479-490. doi: 10.1037/0022-3514.64.3.479