



## Personality and language use in self-narratives

Jacob B. Hirsh\*, Jordan B. Peterson

Department of Psychology, University of Toronto, Sidney Smith Hall, 100 St. George St., Toronto, Ontario, Canada M5S 3G3

### ARTICLE INFO

#### Article history:

Available online 19 January 2009

#### Keywords:

Personality  
Language  
Big Five  
LIWC  
Self-narratives  
Word use

### ABSTRACT

Social and personality psychologists have recently begun examining patterns of natural language use in relation to psychological phenomena. One domain of interest has been the relationships between individual differences in personality and the types of words that people use. The current study extends this research by examining the association between personality traits and language use in the production of self-narratives. Ninety-four undergraduate students were led through an automated writing program that facilitated the telling of the past and the planning of the future. Word usage was categorized using James Pennebaker's Linguistic Inquiry and Word Count (LIWC) text-analysis software. Individual differences in the frequency of word use within these categories were correlated with measures of the Big Five personality traits. Every one of the Big Five was strongly and significantly associated with word use patterns theoretically appropriate to the trait, indicating strong connections between language use and personality.

© 2009 Elsevier Inc. All rights reserved.

### 1. Introduction

While the idea that language and personality are related is not new (Sanford, 1942), researchers have only recently begun to systematically examine how word use relates to individual differences in personality traits (Fast & Funder, 2008; Mehl, Gosling, & Pennebaker, 2006). Recent experimental discoveries have demonstrated that patterns of natural language use reveal a great deal about an individual's psychological characteristics (Pennebaker, Mehl, & Niederhoffer, 2003). Word usage appears characterized by considerable within-person stability, making it an appropriate variable for individual differences research (Mehl & Pennebaker, 2003; Pennebaker & King, 1999). Linguistic differences have thus become an important variable in social and personality psychology.

Initial investigations into the relation between word use and the five-factor model of personality demonstrated significant associations between the two domains, but the effects tended to be rather small (Pennebaker & King, 1999). The most obvious explanation for these results is that the relationships between personality traits and language use are genuinely minor. Before this conclusion is accepted, however, some alternative explanations might be considered. First, the initial research included a large number of relatively unstructured stream-of-consciousness writing exercises. Different participants were likely to approach these writing assignments in a variety of idiosyncratic manners, which potentially may have increased measurement error. Second, the length of the original writing assignments may not have been large enough for stron-

ger personality effects to be reliably observed. Finally, it is possible that the nature of the writing task could influence the magnitude of the observed relationships. In particular, stream-of-consciousness writing might be less powerfully related to personality traits than forms of writing that are more clearly linked to self-expression.

One form of writing that might be more closely linked to personality is the domain of self-narratives. A large body of research has now examined how personality is expressed through this broader linguistic domain (McAdams, 2001). This research is predicated on the idea that the self is structured in terms of personal and cultural narratives, which are largely mediated by language (Bruner, 1991). Narratives about an individual's life trajectory, including subjective descriptions of the past and the future, appear to be central to selfhood and identity. It is reasonable to assume that patterns of word usage within these self-narratives might reveal stronger relationships with personality traits than has been found using stream-of-consciousness or essay-writing exercises. Because personal narratives are extremely self-relevant, their content and style should be more likely to reflect individual differences in personality characteristics (McLean, Pasupathi, & Pals, 2007).

It is thus possible that the effect sizes obtained from the initial research relating personality traits and word usage may not generalize across all types of linguistic production. Indeed, subsequent research in which participants completed a semi-structured 1-h life history interview demonstrated stronger relationships between language use and personality trait scores (Fast & Funder, 2008), in addition to showing that word use predicted ratings of behavior. The current study therefore attempts to combine the

\* Corresponding author. Fax: +1 416 978 4811.

E-mail address: [jacob.hirsh@utoronto.ca](mailto:jacob.hirsh@utoronto.ca) (J.B. Hirsh).

word usage and narrative approaches to personality, by examining how self-reported personality traits relate to word use during the production of self-narratives. A group of undergraduate students were guided through two automated self-authoring processes, one of which facilitates the telling of the past and one the planning of the future. The advantages of this methodology are that it allowed for (a) a large sample of writing, (b) a standardized process, and (c) written expression in a highly self-relevant domain. It was expected that personality traits would be significant predictors of word use in the self-authoring narratives.

## 2. Methods

### 2.1. Participants

Participants included 94 undergraduate students from the University of Toronto (72 female), with an age range of 19–47 years ( $M = 22.0$ ,  $SD = 3.3$ ). Students were recruited from a third-year undergraduate psychology class. At the end of testing, participants received extra course credit for their time. The sample consisted mostly of students from European–Canadian (58.5%) and East-Asian (28.7%) backgrounds.

### 2.2. Materials

#### 2.2.1. Writing assignment

The writing assignment involved two components. In the first, “past-authoring” component, participants were asked to write about their past experiences. This process was facilitated by a training program that guides the participants through the storytelling process. Participants were initially asked to identify seven epochs in their lives. Within each of these epochs, participants then described up to six significant experiences. Participants were asked to discuss how each of these experiences influenced their lives, how it came about, and how it changed their self-views. In the second, “future-authoring” component of the writing assignment, participants were asked to write about their future goals. This involved writing about the ideal future, the specific goals that need to be established to realize this future, and the plans for overcoming any potential obstacles. Both parts of the assignment were completed gradually throughout the course of the semester.

#### 2.2.2. Linguistic Inquiry and Word Count (LIWC; Pennebaker, Francis, & Booth, 2007)

The LIWC software was used to analyze word frequencies in the writing assignments. This software derives frequency values for a large number of words that are pre-sorted into psychological and linguistic categories. Analyses in the current study were limited to the default LIWC2007 dictionary’s personal concerns and psychological word categories, which appeared the most psychologically informative. The LIWC categorization procedure is highly correlated with that of trained judges, indicating good external validity. The LIWC output presents the number of words relating to a given category as a percentage of all words in the writing sample.

#### 2.2.3. Big Five Aspect Scales (BFAS, DeYoung, Quilty, & Peterson, 2007)

The BFAS is an empirically validated instrument for measuring the broad Big Five dimensions of personality, as well as the lower-level aspects. The questionnaire features 100 descriptions with which respondents must rate their agreement on a 5-point Likert scale (e.g., “Sympathize with others’ feelings” and “Like to solve complex problems”). All items were empirically derived and validated as markers for the 10 aspect-level traits. As a measure of the broad domains, the BFAS has been validated against standard

Big Five instruments such as the BFI and the NEO PI-R with an average uncorrected correlation of  $r = .76$ . The scale also demonstrates internal (mean  $r = .83$ ) and test-rest (mean  $r = .81$ ) reliability.

At the aspect level, Extraversion separates into Assertiveness ( $\alpha = .87$ ;  $M = 3.4$ ;  $SD = .63$ ) and Enthusiasm ( $\alpha = .86$ ;  $M = 3.6$ ;  $SD = .65$ ), Agreeableness into Compassion ( $\alpha = .86$ ;  $M = 4.2$ ;  $SD = .49$ ) and Politeness ( $\alpha = .79$ ;  $M = 3.7$ ;  $SD = .59$ ), Conscientiousness into Industriousness ( $\alpha = .85$ ;  $M = 3.1$ ;  $SD = .65$ ) and Orderliness ( $\alpha = .82$ ;  $M = 3.4$ ;  $SD = .62$ ), Neuroticism into Volatility ( $\alpha = .93$ ;  $M = 2.9$ ;  $SD = .83$ ) and Withdrawal ( $\alpha = .87$ ;  $M = 3.1$ ;  $SD = .72$ ), and Openness/Intellect into Openness ( $\alpha = .73$ ;  $M = 4.1$ ;  $SD = .47$ ) and Intellect ( $\alpha = .84$ ;  $M = 3.8$ ;  $SD = .61$ ). Although the aspects from each domain are correlated with each other, they also show good divergent validity. The BFAS thus provides a good assessment of the broad Big Five domains, and provides the additional advantage of assessing an empirically derived aspect level of personality.

### 2.3. Procedure

Participants were given instructions for accessing the study materials online, and were free to complete the components at their own convenience. At the beginning of the semester, participants completed the BFAS and a demographics questionnaire. Participants were then asked to complete the writing assignments over the course of the semester, starting with the past-authoring and followed by the future-authoring. At the end of the semester, students were debriefed about the nature of the study and were given course credit for their time.

## 3. Results

The average word count was 10,693 ( $SD = 6379$ ) for the past-authoring and 5755 ( $SD = 2335$ ) for the future-authoring component. The vast majority of the words were recognized by the LIWC dictionary ( $M = 91.5\%$ ,  $SD = 2.0\%$ ). Prior to analyzing the word frequency data, a square-root transformation was applied to each LIWC category, producing more normally distributed variables. Because our sample contained substantial variability in familiarity with the English language, we statistically controlled for the number of years that each participant had been speaking English ( $M = 18.8$ ,  $SD = 5.6$ ). Partial correlations were conducted between each personality trait and the personal concern and psychological categories from the LIWC2007 dictionary. Table 1 presents the significant correlates of each personality trait, along with sample phrases to illustrate the writing content. We were also interested in examining the correlates of the 10 lower-order personality traits assessed by the BFAS. The same procedure was conducted with these lower-order traits. Results are presented in Table 2. Controlling for gender and word count did not change the obtained pattern of results.

Word choice in the writing assignment was found to be significantly associated with personality at both the broad five-factor level and the lower-order aspect level. An average correlation of  $r = .23$  was observed among the significant correlations between the Big Five traits and word usage. Previous analyses of the psychometric properties of the LIWC categories reported a mean Cronbach alpha coefficient of .59 (Pennebaker & King, 1999). Correcting for this reduced reliability (using the average reliability coefficient of .84 for the BFAS domains), increases the observed mean correlation to  $r = .33$ . These effects thus appear to be moderate in size when left uncorrected, but emerge as strong effects when correcting for reduced reliability (Hemphill, 2003). Overall, the results suggest that individual differences in personality are manifested in students’ word choice during the construction of self-narratives.

**Table 1**  
Linguistic correlates of each Big Five trait, with example phrases.

	<i>r</i>	Example sentences
<i>Extraversion</i>		
Humans	.25	"I feel that it facilitated my trust in <b>people</b> "
Social Processes	.22	"This experience contributed to my current love for public <b>speaking</b> "
Family	.21	"This goal will become increasingly important when I begin a <b>family</b> "
<i>Agreeableness</i>		
Certainty	.22	"I felt <b>total</b> security"
Inclusive	.22	"Now, I suddenly felt <b>included</b> "
Family	.21	"It was a hard decision that I had to make for the sake of my <b>family</b> "
Body	-.20	"It has caused me to have a relatively frail <b>body</b> nowadays"
Anger	-.26	"I <b>hated</b> my teacher"
<i>Conscientiousness</i>		
Achievement	.22	"We are high <b>achievers</b> and encourage one another to do our best"
Work	.21	"All the hard <b>work</b> was completely worthwhile"
Body	-.20	"I have very weak upper <b>body</b> strength"
Death	-.21	"I do not want to confront poverty, sickness and <b>death</b> "
Anger	-.23	"It made me <b>angry</b> and helpless"
Exclusive	-.24	"In my ideal future I would stop <b>just</b> reacting and start acting"
<i>Neuroticism</i>		
Sad	.29	"I walked around with a monstrous <b>sadness</b> "
Negative Emotion	.26	"It requires breaking a vicious circle of <b>guilt</b> "
Body	.22	"I felt this awkwardness in my <b>body</b> "
Anger	.20	"I will also be less <b>angry</b> with myself"
Home	.19	"I will stay <b>home</b> Saturday nights, when I will feel like it"
Anxiety	.19	"I was just chronically <b>scared</b> of the unknown"
Work	-.25	"I've been so busy with <b>work</b> "
<i>Openness</i>		
Perceptual Processes	.28	"I will start trying to <b>listen</b> "
Hear	.27	"I want to be able to talk to them and <b>hear</b> their voices"
Exclusive	.20	"I do not want to live <b>without</b> music in my future"

For all correlations,  $p < .05$ .

**Table 2**  
Linguistic correlates of each Big Five aspect.

Enthusiasm	<i>r</i>	Withdrawal	<i>r</i>
Humans	.26	Sadness	.30
Money	-.20	Negative Emotion	.25
		Anxiety	.21
Assertiveness		Body	.20
Anxiety	-.21	Work	-.24
		Volatility	–
Compassion		Sadness	.24
Certainty	.22	Negative Emotion	.23
Home	.21	Anger	.23
Humans	.20	Body	.19
		Work	-.22
Politeness		Openness	–
Inclusive	.20	Hearing	.40
Family	.19	Perception	.38
Anger	-.24	Religion	.36
Body	-.27	Sexuality	.25
Industriousness		Sadness	.21
Work	.25	Seeing	.21
Exclusive	-.24	Health	-.20
Body	-.26	Work	-.23
Death	-.28	Friends	-.24
Anger	-.28		
Orderliness		Intellect	–
Inclusive	.20	Tentativeness	.20
Achievement	.19	Exclusive	.20
Discrepancy	-.21	Body	-.21

For all correlations,  $p < .05$ .

#### 4. Discussion

Word usage during the production of self-narratives was significantly associated with the Big Five personality traits across a variety of psychological categories. While previous research demonstrated smaller effect sizes (Pennebaker & King, 1999), the current results suggest that personality-specific patterns of language use may be seen most clearly during the production of self-narratives (as opposed to stream-of-consciousness or essay-writing exercises). Additionally, the observed correlations demonstrate meaningful relationships between word use patterns and the Big Five traits themselves. Interestingly, many of the observed correlations involve the same linguistic categories identified by Fast and Funder (2008) as most likely related to personality. The strength of the observed effect sizes, which were broadly similar across traits, suggests that language use is indeed an important reflection of human personality.

Extraversion was associated with words related to humans, social processes, and family. These findings are consistent with the fact that extraverted individuals are active social explorers. Given that extraverts are more engaged with the social world, it would be expected that their descriptions of the past and goals for the future revolve around social processes. At the aspect level, the social element of extraverted writing appeared most strongly related to Enthusiasm. This is the aspect of Extraversion most closely related to sociability. Conversely, the Assertiveness aspect, which measures the dominance-related components of Extraversion, was negatively correlated with Anxiety words. More assertive individuals tend to be approach-oriented and less distressed by potential obstacles, and this appears to be reflected in their writing.

Agreeableness, like Extraversion, was related to family, as well as to inclusiveness, consistent with this trait's association with empathy and interpersonal concern. The negative relationship with anger is also supported by previous research looking at agreeableness and the inhibition of interpersonal aggression (Meier, Robinson, & Wilkowski, 2006). Highly agreeable people were also less likely to use body-related words. This may be because the improved health behavior associated with Agreeableness, leads to less somatic concerns (Booth-Kewley & Vickers Jr., 1994). Conversely, such individuals may be more concerned with interpersonal than self-related physical goals. The relation between Agreeableness and certainty-related words was both interesting and unexpected, and therefore deserves further investigation. It appears that agreeable individuals have a greater sense of certainty in their lives, and consequently think in more concrete terms. Perhaps the tendency of agreeable people to produce tight interpersonal bonds gives them a heightened sense of security.

Conscientiousness was associated positively with achievement and work-related words. Both of these findings are consistent with the strong work ethic and achievement orientation of highly conscientious individuals (Barrick & Mount, 1991). Negative correlations were observed for death and body-related words. Conscientious individuals, like their agreeable counterparts, are also characterized by higher levels of positive-health-related behaviors, along with lower mortality rates (Booth-Kewley & Vickers Jr., 1994). Perhaps health concerns are therefore less likely to afflict these individuals, leading to chronically reduced accessibility of body and death related thoughts. Finally, Conscientiousness has also been implicated in the control of anger (Jensen-Campbell, Knack, Waldrip, & Campbell, 2007) and this again appears reflected in the students' writing.

The trait of Neuroticism was clearly reflected in the students' writing samples, as it was correlated with negative emotion, anger, anxiety, and sad words. Additionally, neurotic individuals were more likely to discuss body-related topics. This may indicate the increased prevalence of physical problems in neurotic individuals

(Brown & Moskowitz, 1997), or the fact that body dysmorphia is closely associated with this trait (Phillips & McElroy, 2000). The negative relationship with work highlights the fact that Neuroticism can be a significant detriment to workplace performance (Barrick & Mount, 1991). The negative emotions associated with Neuroticism make it difficult to focus on work, especially during times of stress. Neurotic individuals also tend to have troubled home lives. It thus makes sense that those high in trait Neuroticism would place a greater emphasis on the home environment when discussing the past and desired future.

Openness, finally, was most strongly related to a greater prevalence of perceptual processes, including words related to hearing and seeing. Openness has previously been related to higher levels of creative functioning and esthetic sensitivity, as well as decreased gating of sensory input (DeYoung, Peterson, & Higgins, 2005). The current results suggest that this esthetic sensitivity also translates into the writing of highly open individuals, with a greater emphasis on perceptual processes.

## 5. Conclusion

Personality traits appear significantly and strongly related to patterns of word use during the telling of the past and the planning of the future. The fact that a relatively strong relationship emerged may be due to the nature of the self-authoring exercises employed, which combined free and structured writing, narrative in format, but remained standardized across participants. The writing exercises were also highly self-relevant, as the personal narratives that were produced constitute important aspects of an individual's identity and personality (McAdams, 2001; McLean et al., 2007). Furthermore, participants devoted large amounts of time and effort to these writing tasks, which were very detailed, personal and much lengthier than is easily obtainable in a lab study. While the typical double-spaced page characteristic of a student essay is under 300 words, the writing samples analyzed in this study averaged ~16,500 words – approximately 55 pages per student. It should be noted, however, that because we did not directly compare writing samples obtained using different instructions, we cannot be sure whether it was the increased self-focus of the writing assignment or the lengthier writing samples that contributed most to the larger effect sizes. Both of these factors are likely to play a role, but their relative contributions will have to be teased apart in future research.

The primary limitations of the study reflect the limitations of the LIWC itself. Although LIWC analysis has proved very useful for examining a number of psychological phenomena (Pennebaker et al., 2003), and is unique in its analytic approach, it still examines word usage only, rather than providing information about the context in which these words are embedded. Because written discourse operates at multiple levels of analysis simultaneously (e.g., word, sentence, paragraph, page), much if not most of the meaning of a writing sample will be lost during a simple word-count. Analytic techniques that take into account higher-order semantic structures would be of obvious use, if and when they

are developed. A second and related limitation is that each of the LIWC categories contains a variety of words that, although statistically related to each other, may encompass a number of different meanings. The categories should thus not be understood as perfect descriptions for their content, but rather as reasonably reliable clusterings of related words.

Despite these limitations, it appears that word usage patterns in self-narratives are intelligibly related to personality. Many of the specific results are consistent with previous research findings in personality psychology, supporting the validity of this analytic technique. Future research should continue examining how personality relates to the manner in which an individual's life experiences are construed and described. More generally, the relationship between language and personality continues to be a topic of interest and deserves further analysis.

## References

- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1–26.
- Booth-Kewley, S., & Vickers, R. R. Jr., (1994). Associations between major domains of personality and health behavior. *Journal of Personality*, *62*, 281–298.
- Brown, K. W., & Moskowitz, D. S. (1997). Does unhappiness make you sick? The role of affect and neuroticism in the experience of common physical symptoms. *Journal of Personality and Social Psychology*, *72*, 907–917.
- Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*, 1–21.
- DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect: Cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, *73*, 825–858.
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: Ten aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*, 880–896.
- 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.
- Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79.
- Jensen-Campbell, L. A., Knack, J. M., Waldrup, A. M., & Campbell, S. D. (2007). Do Big Five personality traits associated with self-control influence the regulation of anger and aggression? *Journal of Research in Personality*, *41*, 403–424.
- McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology*, *5*, 100–122.
- McLean, K. C., Pasupathi, M., & Pals, J. L. (2007). Selves creating stories creating selves: A process model of self-development. *Personality and Social Psychology Review*, *11*, 262.
- 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.
- Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*, *84*, 857–870.
- Meier, B. P., Robinson, M. D., & Wilkowski, B. M. (2006). Turning the other cheek: Agreeableness and the regulation of aggression-related primes. *Psychological Science*, *17*, 136–142.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2007). *Linguistic Inquiry and Word Count (LIWC): LIWC2007*. Mahwah: Lawrence Erlbaum Associates.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*, 1296–1312.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Reviews in Psychology*, *54*, 547–577.
- Phillips, K. A., & McElroy, S. L. (2000). Personality disorders and traits in patients with body dysmorphic disorder. *Comprehensive Psychiatry*, *41*, 229–236.
- Sanford, F. H. (1942). Speech and personality. *Psychological Bulletin*, *39*, 811–845.