Assessing the Big Five Personality Traits with Latent Semantic Analysis

Peter J Kwantes[1,2]

Natalia Derbentseva[1]

Quan Lam[1]

Oshin Vartanian[1,3]

Harvey H. C. Marmurek[4]

1. Defence Research & Development Canada
2. University of Queensland
3. University of Toronto at Scarborough
4. University of Guelph

Corresponding Author:

Peter Kwantes

DRDC Toronto Research Centre

1133 Sheppard Ave W.

Toronto, ON

M3K 2C9

Phone: +1 (416) 635-2028

Email: peter.kwantes@drdc-rddc.gc.ca

DRDC-RDDC-2014-P28

Abstract

We tested whether the characteristics of a person`s personality can be assessed by an automated analysis of the semantic content of a person's written text. Participants completed a questionnaire measuring the, so called, Big Five personality traits. They also composed five short essays in which they were asked to describe what they would do and how they would feel in each of five scenarios designed to invoke the creation of narrative relevant to the Big Five personality traits. Participants' essays were processed for content by Latent Semantic Analysis (LSA; T. Landauer & S. Dumais, 1997), a model of lexical semantics. We found that LSA could assess individuals on three of the Big Five traits, and we discuss ways to improve such techniques in future work.

Keywords: Personality, Big Five Traits, Latent Semantic Analysis

Assessing the Big Five Personality Traits with Latent Semantic Analysis

1. Introduction

In this article we build on classic work (e.g., Allport & Odbert, 1936; Cattell, 1943) that explores the role that words play in the description of personality. Specifically, we tested whether a collection of words that describes a trait can be used in an automated tool to assess a person's personality from information contained in his or her written text.

Current techniques of personality analysis from text samples are dominated by algorithms that tally and track word usage and map the patterns of usage across word categories onto personality traits. The basic idea of the text analytic approach is that personality influences word choice in one's speech or writing behaviour. To the extent that we can characterize the word usage patterns common to the different personality types, we should be able to assess personality based on an examination of language samples generated by a speaker/author. Among available software-driven text analytic techniques, the most widely cited is Pennebaker's Linguistic Inquiry and Word Count (LIWC; Pennebaker & King, 1999). LIWC is essentially a word frequency counter that tallies an author's use of words that are yoked to linguistic (e.g., prepositions, articles, numbers), psychological (e.g., optimism, anger, insight), or physical (e.g., work, sleep, sexuality) categories. The LIWC yields a profile of the speaker's words across the categories, and the pattern with which they are distributed across categories is believed to be driven in part by the author or speaker's psychological traits or states at the time.

Most of the language-based work focuses on personality as expressed by the, so called, Big Five personality traits (McCrae & John, 1992). The Big Five comprise: *Extraversion* (described as being friendly, assertive and sociable), *Conscientiousness* (described as being organized, dependable, and motivated), *Agreeableness* (described as being cooperative, trusting, and helpful), *Openness to New Experience* (described as being emotional, curious, creative, imaginative, and hereafter referred to as, *Openness*), and *Neuroticism* (associated with easily being made to feel upset, angry, anxious, or depressed).

Pennebaker and King (1999) and Yarkoni (2010), for example, reported that authors who scored high on Extraversion tended to use fewer negative emotion words and more social words (e.g., restaurant, meet) than introverts. By contrast, Yarkoni also reported that people high on the Neuroticism trait tended to use more negative emotion words than people low on neuroticism. In sum, the language we use to express ourselves seems to provide a window into how we feel, how we think, and how we are built psychologically. Current techniques to examine personality from language focus on the classification of word usage. The purpose of the current study was to augment such analyses of language by analysing the semantic content of a speaker or author's text.

Analysing text for word usage represents a categorical characterization of an author's text. Another way to characterize text is by the semantic content it carries. In this article we examine whether a formal analysis of word usage, like that provided by the LIWC, can be complemented with a meaning-based analysis of the text generated by an author. Over the past two decades computational models have been developed to create semantic representations for words encountered in text. One such model is Latent Semantic Analysis (LSA; Landauer & Dumais, 1997; Landauer, Foltz & Laham, 1998). LSA is a computational model that works on the notion that words with similar meanings tend to appear in similar contexts. It creates semantic representations for words by analysing the pattern with which words occur together in documents across thousands of text samples provided to it in a training corpus. Then, from an analysis of the words that do and do not co-occur in the corpus, the model estimates what words should occur in similar documents (i.e., contexts) and are, therefore, close to each other in semantic space.

Using LSA to explore aspects of an author's psychological make-up is not entirely new. For example, Campbell and Pennebaker (2003) reported results from participants who were asked, over several sessions, to write about an emotional time in their life or about some emotional experiences. They used LSA to characterise changes in writing content and style over time, and measured the extent to which changes were related to improved wellness. They found no evidence that the semantic content of patients' written text over the time spent in treatment was in any way correlated with changes in well-being. They did, however, find that changes in writing style over time, especially the use of pronouns, were related to changes in patients' well-being.

Little work has been done to test LSA as a tool in the evaluation of authors' or speakers' personalities. To fill that gap in research, we examined how strongly the semantic content of authors' text is driven by personality and ask whether we can use LSA to measure aspects of it. To do so, participants were presented with five scenarios, and for each, asked to describe how they would feel and what they would do. The five scenarios were designed to invoke the production of narratives relevant to each of the five dominant, or "Big Five", personality traits (McCrae & John, 1992). We postulate that when participants write about themselves in the scenarios, they will use terms that express their status on each of the Big Five traits. We hypothesize further that the more strongly a participant identifies with a trait, the more his or her narrative will contain text relevant to the trait, and that such differences can be detected using models like LSA.

2. Method

2.1 Participants.

One hundred and fifteen first-year (19 male) undergraduate students in an introductory psychology course at the University of Guelph participated in the study for course credit. Average age of participants was 19 years old (range 18-23 years), All but 12 of the participants reported English as their first language.

2.2 Materials.

2.2.1. Testing Materials.

Five scenarios were developed to describe situations in which participants were to imagine themselves. For each scenario participants were prompted to ponder how they would feel and what they would do. Each of the five scenarios, reproduced in Appendix A, was designed to be relevant to one of the Big Five personality traits. The scenarios were devised by a focus group of three researchers at Defence Research & Development Canada (DRDC) Toronto. Validation and fine-tuning of the scenarios was then done using feedback from a separate sample of three researchers at DRDC Toronto.

The Big Five Inventory (BFI; John, Donahue & Kentle, 1991; John, Naumann & Soto, 2008) was used in the study. The BFI is a 44 item test wherein respondents indicate their agreement with statements about themselves on a five-point scale. John and Srivastava (1999) reported alpha reliabilities for the five scales of between .75 and .90 and test-retest reliability between .80 and .90. They also report strong agreement (M= .87) between the BFI and other tests like the Costa and McRae's (1985) NEO Five Factor Inventory, and Trait Descriptive Adjectives (Goldberg, 1992) which also assess the Big Five personality traits.

2.2.2. LSA.

As mentioned above, LSA is an algorithm that generates a *semantic space* from a statistical analysis of frequencies with which words co-occur in a large collection of documents (i.e., contexts) during an initial training phase[1]. After training, a semantic space comprises a set of vectors containing the semantic features for each word encountered during training. We refer to the vectors in the semantic space as, *semantic vectors*. Generally speaking, the more documents contained in a training corpus, the more contextual information the system has to semantically differentiate or align words. We used different corpus sizes to ensure that, if LSA failed to detect differences in authors' personalities, the results might suggest whether it was because of an insufficient number of documents during the training phase.

The other important aspect to consider when building LSA's semantic space is the choice of training corpus. LSA builds its semantic knowledge by exploiting the associations among words within the thousands of training documents. As a consequence, how the training corpus uses language and how words are associated in the training corpus will drive the nature of the system's interpretation of a word. For example, if LSA were trained on a document collection dominated by sports-related articles, its semantic representation for the word *play* would have different close associates than if the collection were dominated by, say theatre-related articles. We trained LSA on two types of corpora. For one version, we used a random collection of articles from Wikipedia. For the other, we trained LSA on Wikipedia articles relevant to the Big Five personality traits. Done this way, LSA's understanding of the words in the collection was in the context of materials related to the personality traits, and might therefore amplify the extent to which terms a person uses are considered to be related to the five personality traits.

We trained LSA separately on seven training corpora. The corpora were constructed by creating collections of varying sizes using different criteria for selecting documents. For the first three corpora, we trained the system on randomly selected articles taken from Wikipedia[2]. The three corpora differed in size. One corpus contained 12000 articles, another 30000, and the final one 50000 articles. For the next three corpora, we used the Lucene (lucene.apache.org) indexer to search the Wikipedia corpus for terms relevant to the Big Five personality traits. Articles were selected by forming a query from terms (mainly adjectives) that describe one extreme on the trait's continuum. The terms were taken from the traits' definitions as reported in Wikipedia and those contained in the BFI. Adjectives from reverse-keyed items on the BFI were changed to their antonyms. The terms in the query are listed in Table 1. Again, we selected three different corpus sizes with collections of the 5000, 10000 and 15000 most relevant documents to train the system. The seventh corpus was also one which only contained documents relevant to the Big Five, but was constructed slightly differently. Instead of extracting documents from Wikipedia using a query that included search terms relevant to all five personality traits, we extracted five 1000-document collections, each relevant to a single trait, and combined the results to create a 5000-document corpus containing document about all five traits. The rationale for creating the final corpus the way we did was to ensure that the proportion of documents relevant to each of the five traits was equal across them.

Once a semantic space is created from a training corpus, the semantic vectors it creates for words can be used to create semantic vectors for new documents. Creating a document vector is straightforward and involves summing the semantic vectors of a document's content words. Once created, pairs of documents can be compared by calculating the cosine between their vectors. The cosine behaves much like a correlation in that a cosine of 0 indicates that two vectors are orthogonal and a cosine of 1 indicates that they have identical projections in high-dimensional space. Words and documents with high cosines project in similar directions in semantic space and are therefore considered by

LSA to be semantically related. Likewise, the lower the cosine between two vectors, the less related they are considered to be.

2.3 Procedure

2.3.1. Data Collection.

Participants were tested individually and completed the task at their own convenience using their own personal computer with an internet connection. The task was completed using an online survey administration program installed at the University of Guelph.

The participants completed two tasks. First, participants were given scenarios in which to imagine themselves and gave open-ended responses detailing what they would do or how they would feel in each situation (See Appendix A for the instructions and the list of scenarios). After completing the five essays, the participants completed the BFI (John, Donahue, & Kentle, 1991).

2.3.2. LSA Processing of the Essays.

For each of the seven semantic spaces, we created a single semantic vector for the text generated by each participant—we refer to it as the *essay vector*. To do so, words specific to the scenarios (e.g., the words “paint” and “colour” from the essays about the *Openness* trait) were removed because they would be common across all essays and irrelevant to the traits under consideration. Then, the vectors for the remaining content words across all five essays were summed to create a single essay vector for the participant. A *trait vector* for each trait was formed by summing the LSA term vectors from the trait-appropriate terms listed in Table 1.

Each participant’s *essay vector* was assessed by calculating its similarity to each of the five *trait vectors*.

3. Results

The data reported here will be limited to the 87 participants who completed all five essays and the BFI. For each participant we calculated the BFI subscale score values corresponding to each of the Big Five traits (shown in Table 2). We also calculated the cosines between each participant’s essay vector and each of the trait vectors for each of the seven semantic spaces we built. Average cosines across traits and semantic spaces are shown in Table 3. As is clear in the table, the text generated by participants had strong associations to all five trait vectors. Table 4 shows the correlations between BFI subscale scores and the cosine between essays to trait vectors. That is, the extent to which the semantic similarities of participants' text to each of the five traits vectors predicts the authors' scores on each of the five traits, as measured by the BFI. We found no indication that Agreeableness and Conscientiousness could be predicted from the semantic content of participants' text. By contrast, as is clear in Table 4, there were statistically significant correlations relating semantic content and Extraversion, Neuroticism, and Openness. In other words, LSA provided reliable predictors of an author's status on three of the five traits.

There was a notable difference between results generated from an LSA trained on domain-relevant documents and one trained on a random selection of articles from Wikipedia. For the three traits that LSA could be used to predict, LSA models trained on domain-relevant documents had significantly higher correlations (M = .24, SD = .04) across all three traits than models trained on a random selection of articles (M = .20, SD = .05), t (19) = 1.89, p = .04.

Finally, it is interesting to note that the size of the training corpus made little difference on the strength of relationship between test scores and the text's alignment with trait vectors. Typically, LSA is trained using tens of thousands of documents. The results here suggest that modestly sized corpora can be adequate for training, especially when trait-relevant documents are used.

4. Discussion

We found that for three of the Big Five personality traits, there was a reliable relationship between a person's subscale scores on the BFI and how closely matched his/her essay's semantic content was to the related trait vector. Generally speaking, for the traits where the technique worked, it did so because the more strongly a person possessed a trait, the more their text tended to contain narrative terms related to it.

The work presented here complements such other linguistic-based techniques for personality assessment (e.g., LIWC; Pennebaker & King, 1999) by placing less constraint on the tokens that one chooses during speech or writing. Our technique adds flexibility to the assessment by measuring the semantic association between the language sample and trait vector containing the blend of semantics from terms describing the trait. Done this way, the language sample is not assessed for what terms it contains (indeed, the sample may contain none of the terms that went into creating the trait vector), but rather, it is assessed for the extent to which it is semantically aligned with the trait vector. We are not claiming that our technique is superior to others like the LIWC; however, we would argue that it has promise as a supplemental technique.

With respect to Agreeableness and Conscientiousness, the traits for which our analysis did not work, there are two possible reasons. First, it may be that differences in Agreeableness and Conscientiousness cannot be easily gauged by an analysis of the semantics of an author or speaker's narrative. The other possibility is that the problem stems from our choices of scenario and the classes of words they invoked authors to use. To explore the notions, we appealed to Chung and Pennebaker's (2008) report on the relationship that word usage from semantic categories had to the Big Five.

Chung and Pennebaker (2008) asked participants to write an essay about themselves. Instructions included a directive to describe, who they are, who they want to be, and how they are seen by others. Like ours, their participants also completed John and Srivastava's (1999) Big Five questionnaire. The words used in the essays were analysed using principal components analysis to determine what semantic categories they fell into. Then, trait scores were correlated with the authors' usage of words from the various word categories. They found, for example, that authors who score high on Neuroticism had an increased use of words related to Evaluation (*smart, lazy*) and Negativity (*horrible, scared*) relative to those who score low on the trait.

Interestingly, and consistent with our analysis, Chung and Pennebaker (2008) found little relationship between Agreeableness scores and the frequency of word usage in any of the word categories they uncovered. The failure of both studies to detect individual differences on Agreeableness might have occurred because authors tend to focus on their own feelings and habits when asked to write about themselves, and not on how "forgiving" "trusting" or "helpful" they are in relation to others. Perhaps we would have detected individual differences in Agreeableness if the scenario we devised was one in which authors were given a more provocative situation to evoke discussions around the critical dimensions of Agreeableness like, for example, interpersonal behaviour. For example, instead of the painting scenario, consider a car accident scenario where the other driver, who is at-fault and slightly injured, needs help. When asked how they would feel and what they would do, perhaps authors who score highly on Agreeableness would be more likely to provide essays containing discussions relating to cooperativeness, forgiveness, and help for the other driver. In sum, our failure to detect differences in Agreeableness might reflect the scenario's failure to prime the right kind of language use.

Likewise, we believe that our scenario for Conscientiousness did not prime the relevant semantic knowledge. Chung and Pennebaker (2008) reported a significant tendency for Conscientiousness to be associated to words related to Evaluation (e.g., *smart, lazy*) and Maturity (e.g., *capable, responsible*). We take Chung and Pennebaker's pattern to suggest that if authors had described their thoughts and actions in a scenario where they had greater responsibility and where failure had serious consequences, we might have been able to detect individual differences in Conscientiousness.

A clear limitation of the work reported here is that we did not undertake preliminary explorations of multiple scenarios for their ability to invoke trait-relevant text from participants. Hence, it is unclear whether our inability to detect individual differences on Agreeableness and Conscientiousness occurred because of the scenario or the trait itself. In order to develop robust remote assessment techniques based on language use, future work will need to develop several candidate scenarios and validate them with a large participant sample.

Future work should also examine the extent to which LSA can be used to assess aspects of personality from stream-of-consciousness text. In many reports published by Pennebaker and colleagues, the LIWC is applied to text samples written by participants who are not given strict instructions about what to write about, but rather are assigned the more general task of writing freely about themselves. It remains to be seen whether the LSA-based technique we describe in this article is sensitive enough to identify personality traits from more unstructured text.

Our results show promise for the development of remote personality assessment techniques when the subject of assessment is inaccessible or must be assessed covertly. The technique we propose needs refinement, but shows promise to complement or augment currently available text-based assessment techniques.

5. References

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. Albany, NY: Psychological Review Company.

Campbell, R.S. & Pennebaker, J. W. (2003). The Secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*, 60-65.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology 38* (4): 476–506

Chung and Pennebaker's (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.

Costa, P. T., Jr., & McCrae, R. R. (1985). The NEO personality inventory manual. Odessa, FL: Psychological Assessment Resources.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26-42.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semanctic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review , 104* , 211-140

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25*, 259-284.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York, NY: Guilford Press

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory--Versions 4a and 54. Berkeley, CA: University of California,Berkeley, Institute of Personality and Social Research.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford Press.

McCrae, R.R., & John, O.P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.

Pennebaker, J. W. & King, L.A., (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Individual Differences*, 77, 1296-1312.

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(3), 363–373.

Table 1

Terms Used to Create Trait Vectors

Trait	Terms
Openness	Original, novelty, curious, different, ingenious, active, imaginative, inventive, artistic, aesthetic, reflective, sophisticated, artistic, musical, literate, unpredictable, fearless, open, creative, adventurous, explore, brave, openness
Conscientiousness	Conscientious, thorough, accurate, reliable, organize, organized, diligent, persevere, persevering, efficient, plan, planning, persist, persistent, focus, focussed, careful, work, painstaking, meticulous, scrupulous, particular, selfless, caring, empathetic
Extraversion	Talkative, outgoing, energetic, enthusiastic, boisterous, assertive, eager, friendly sociable, lively, social, open, chatty, meet, interaction, energized, public
Agreeableness	Agreeable, helpful, help, unselfish, altruistic, agree, agreement, forgive, forgiving, trust, trusting, warm, friendly, friend, considerate, kind, polite, cooperate cooperative, easygoing, accommodating
Neuroticism	Neurotic, depressed, blue, agitated, stressed, tense, worried, worry, emotionally, emotional, unstable, upset, moody, restless, tense, nervous, unstable, anxiety, compulsive, obsessed, indecisive, maladjusted, anxious, uneasy, irritable

Table 2

Mean Trait Scores with Standard Deviations in Parentheses

		Traits			
	AGREE	CONC	EXTRA	NEURO	OPEN
Trait Scores	3.79 (.63)	3.59 (.61)	3.21 (.82)	3.28 (.61)	3.52 (.59)

AGREE = Agreeableness, CONC = Conscientiousness, EXTRA = Extraversion, NEURO = Neuroticism, OPEN = Openness.

Table 3

Average Cosines between Participants' Essays and the Trait Vectors. Standard Deviations are in Parentheses.

	Traits				
	AGREE	CONC	EXTRA	NEURO	OPEN
Randomly Generated Corpora					
12,000 Documents	.77 (.02)	.69 (.03)	.62 (.02)	.58 (.02)	.65 (.02)
30,000 Documents	.85 (.01)	.79 (.02)	.82 (.02)	.70 (.02)	.78 (.02)
50,000 Documents	.81 (.02)	.71 (.03)	.64 (.02)	.64 (.02)	.69 (.03)
Big Five-Related Corpora					
5,000 Documents	.71 (.02)	.72 (.03)	.69 (.02)	.54 (.03)	.65 (.03)
10,000 Documents	.74 (.02)	.73 (.03)	.71 (.02)	.56 (.03)	.66 (.03)
15,000 Documents	.75 (.02)	.72 (.03)	.73 (.02)	.57 (.02)	.67 (.03)
5 traits x 1,000Documents	.64 (.02)	.67 (.03)	.56 (.02)	.53 (.03)	.64 (.03)

AGREE = Agreeableness, CONC = Conscientiousness, EXTRA = Extraversion, NEURO = Neuroticism, OPEN = Openness.

Table 4

Correlation Coefficients Indicating the Extent to Which Trait Scores on the BFI can be Predicted by the Cosine Between Participants' Essays and Trait Vectors.

	Traits				
	AGREE	CONC	EXTRA	NEURO	OPEN
Randomly Generated Corpora					
12,000 Documents	.03	.09	.20*	.21*	.17+
30,000 Documents	-.06	.05	.19*	.15	.23*
50,000 Documents	-.02	.06	.14	.26**	.25**
Big Five-Related Corpora					
5,000 Documents	-.03	.08	.20*	.21*	.30**
10,000 Documents	-.01	.08	.20*	.19*	.32**
15,000 Documents	-.02	.08	.19*	.22*	.31**
5 traits x 1,000 Documents	-.03	.11	.17+	.28**	.29**

AGREE = Agreeableness, CONC = Conscientiousness, EXTRA = Extraversion, NEURO = Neuroticism, OPEN = Openness. + = $p < .06$, * = $p < .05$; ** = $p < .01$

Appendix

Instructions for the Essay Component of the study

Please imagine that you are in the following scenario. Knowing yourself as you do, describe **how you would feel and what you *would* do** (not what you *should* do) in each situation. Do not spend too much time thinking about your answer--instead, start writing, and express what comes to mind. Do not labour over your spelling and grammar. Your writing will not be graded. Instead, it is important that you express your thoughts while you have them. Please try to write between 200 and 300 words (approximately one half of a page, single-spaced)

1. Question relevant to the Quality of Conscientiousness

You're working alone late at the office and you notice a strange smell and a hazy mist hanging in the air of the corridor. You suspect it's some gas or vapor leak from some equipment or machinery in the building. You have no idea whether the leaked vapor is hazardous**. As honestly as possible, describe what you would do in this situation.**

2. Question relevant to the Quality of Extraversion

Your friend wants you to attend an important party to which he/she has been invited. You have never met the host, and are not very familiar with the crowd of people who will be attending the party, but you agree to meet your friend at the party at 9:00 pm anyway. When you arrive there, you realize that your friend is late. **How would you feel, and what would you do while you waited for your friend**?

3. Question relevant to the Quality of Openness

You have won an Air Canada paid vacation package for one person to any destination in the world. Your package includes round trip plane tickets, accommodations for any type of lodging, and $5,000 spending money. **Assuming that you were available to go, where would you choose to go and why**?

4. Question relevant to the Quality of Agreeableness

Your housemate decides to paint her bedroom a new colour. One night, when you come home from class, you discover that she also painted your room in the same colour because she had paint left over and didn't want it to go to waste. **As realistically as possible, describe how you would feel and how you would you handle the situation**.

5. Question relevant to the Quality of Neuroticism

You have developed an email friendship with someone. In your latest email, you ask your friend a more personal question. Your friend usually replies quite promptly, but has taken unusually long to reply to your latest questions. **Discuss how you would interpret this long period of silence, how you would react and what you would do about it**?

Endnotes

1. The version of LSA we used here was developed for DRDC under contract, and can be downloaded at:

mall.psy.ohio-state.edu/wiki/index.php/Semantic_Models_Package_(SEMMOD)

2. Wikipedia documents can be quite long. In response, only the first 200 words of every selected document used during training.

Acknowlegements

This work was supported by funds granted to PJK by DRDC's Applied Research Project Program