SAGE researchmethods cases

Developing a New Method for Psychological Investigation Using Text as Data

Contributors: Eyal Sagi
Pub. Date: 2018
Access Date: February 20, 2018
Academic Level: Postgraduate
Publishing Company: SAGE Publications Ltd
City: London
Online ISBN: 9781526442604
DOI: http://dx.doi.org/10.4135/9781526442604

Abstract

Over the past few decades, the study of psychology has been undergoing a methodological transformation. The increasing availability and quantity of real-world big data have prompted researchers to look for new ways to use these data to gain psychological insights. In this case study, I trace the process of developing a new method for testing psychological theories using corpora. Instead of bringing participants to the lab, I analyze statistical patterns of co-occurrence in naturally occurring texts obtained from a variety of sources, including political speeches, literature, and the Internet. The basic assumption I make is that these patterns reflect the representations and cognitive processes of their author. I present three different applications of the method and use them to describe how such data can be analyzed and used to answer a range of questions in psychology and social science. The first application examines the linguistic question of the relationship between word form and meaning. The second application identifies cognitive frames as they are found in text. The final application uses texts to measure the style of moral reasoning individuals apply in particular contexts. This case study provides insight into the process of developing new methodologies for hypothesis testing, as well as demonstrating how to formulate hypotheses that can be tested using corpora. In addition, several key pitfalls in the process of adapting statistical methods to new uses are identified and discussed.

Learning Outcomes

By the end of this case, students should be able to

- Understand how to use traditional hypothesis testing methods and designs on textual data
- Be familiar with the use of texts to supplement traditional psychological lab-based studies
- Be able to adjust hypotheses so that they can be tested using big data
- Recognize the importance of interdisciplinary knowledge and collaboration in research

---

Project Overview and Context

Psychology focuses on understanding individuals, mostly humans. Therefore, it is not surprising that most psychological studies involve observations of human (or animal) behavior in various circumstances. Through such observations, psychology has advanced our understanding of what makes people tick and how we can guide ourselves to become better. However, human behavior can be observed indirectly, as well as directly. For example, archeologists learn about extinct cultures by studying the ruins and artifacts they have left behind. Historical linguists learn about the principles that guide the evolution of language by examining commonalities among languages and by comparing old textual records. What if we could augment the study of psychology by studying indirect evidence similarly? That is the promise of several disciplines in psychology, such as sociobiology and evolutionary psychology.

In this case study, I will describe a different approach that I have been developing over the past decade. It focuses on the hypothesis that we can learn about humans by examining how they use language. One of the early methods that were based on this idea is James Pennebaker's Linguistic Inquiry and Word Count (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007). More recently, I, and others, have proposed using more elaborate methods that we believe are better suited for measuring psychological variables and constructs via language (Sagi & Dehghani, 2014a; Sagi, Diermeier, & Kaufmann, 2013).

Language as a Reflection of the Person

Psycholinguists have long observed that our cognitive makeup and personality affect our use of language. In some sense, it seems almost trivial to argue that our thoughts and experiences affect what we say. After all, we believe that the primary role of language is to allow us to communicate needs, desires, and ideas and therefore it is required that our use of language reflect these.

However, the psychological claim runs deeper—empirical evidence suggests that the specific words and phrases we choose to use, as well as our choice of grammar, reflect beliefs and representations that are often implicit and not readily available to the conscious mind. Essentially, language allows us to communicate the same ideas in a variety of ways. Our personality, beliefs, and subconscious tendencies guide us to choose the particular manner of expression that communicates our conscious intent (cf. Tausczik & Pennebaker, 2010).

Moreover, the context in which we operate also affects these choices and evidence suggests that we adapt our use of language to better fit the language use of individuals we communicate with. One influential account, developed by Martin Pickering and Simon Garrod (2004), argues that, in the course of a dialogue, individuals *align* their use of language to each other. That is, their choice of grammar, phrasing, and representation becomes more similar over time. Pickering and Garrod identify several distinct levels of language use over which this phenomenon has been observed, from the way words are pronounced (phonetics) to the mental representation that the interlocutors use, which is often referred to as a *situation model* in this context.

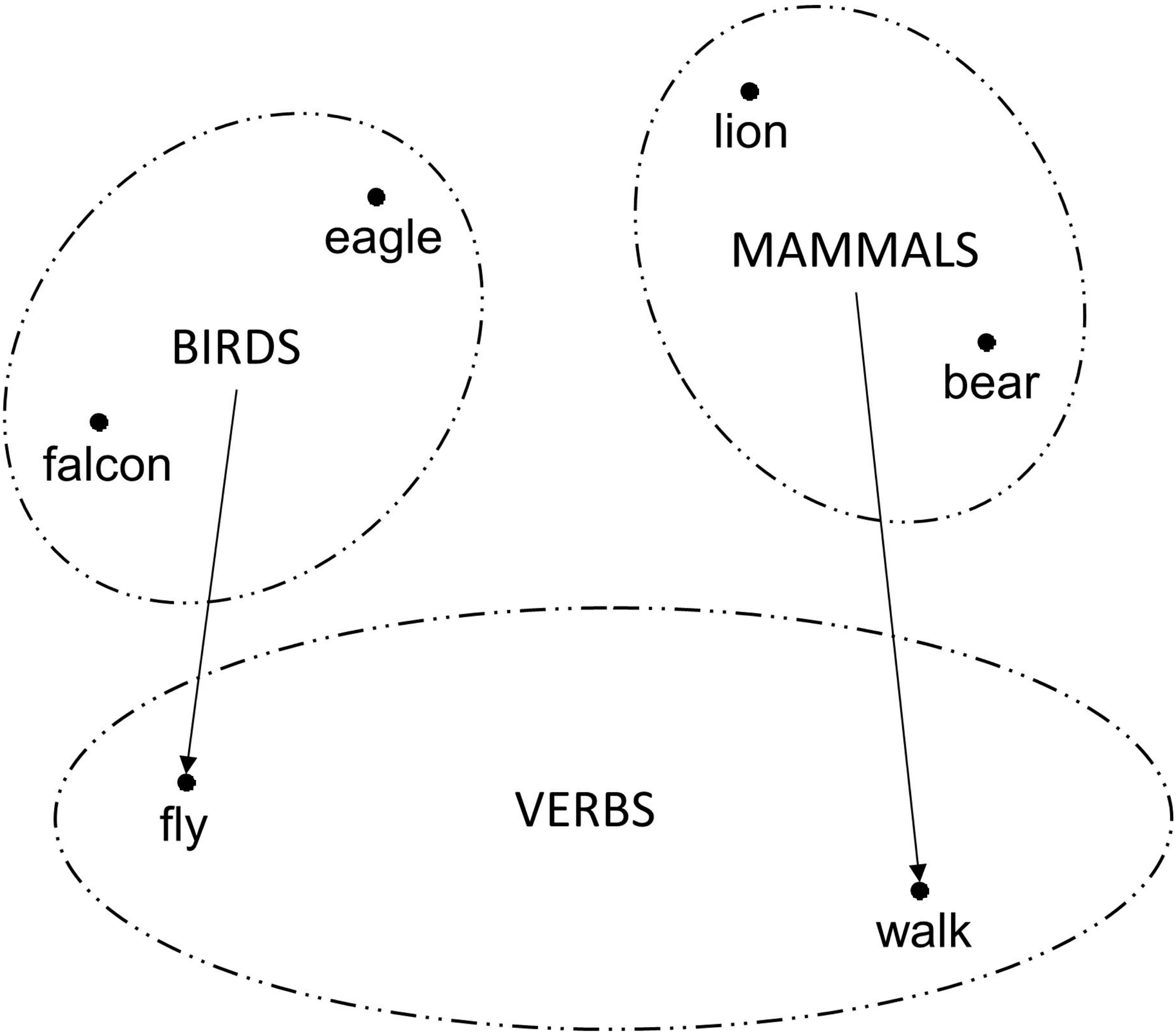
Taken together, these studies make a strong case that the language we use to communicate can also provide a window into our thoughts, feelings, and beliefs. Language is therefore more than a mere means of communication—it reflects the person producing it. Consequently, it should be possible to use the language produced by individuals to learn about them and gain insight into their psyche. All that is required is a principled method of identifying patterns in language use and connect them with their psychological correlates.

Using Language as an Instrument of Measurement

The idea that the meaning of a word not entirely contained with it is not new. It has been postulated by several influential thinkers, including Firth (1957), who wrote, “You shall know a word by the company it keeps.” Once computers arrived on the scene, it was inevitable that they will be recruited to help researchers wade through texts to uncover the hidden patterns within. As I started graduate school, some of these methods were already well established. Moreover, Thomas Landauer’s approach, Latent Semantic Analysis (LSA; Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998), has already come under criticism for their ambitious claims regarding the ability of such statistical methods to model cognitive processes and linguistic structures.

At their core, these models attempt to distill the essence of meaning through the statistical analysis of patterns of word co-occurrence. This is a direct application of Firth’s postulate, as the underlying hypothesis is that words that tend to occur in close textual proximity are likely to refer to the same, or similar, entities because they share the same underlying context. The result of such an analysis is frequently represented as a high-dimensional space in which words that are frequently found together are closer than those that rarely appear in the same context. Figure 1 depicts a reduced dimensional space for some birds and mammals, as well as relevant verbs of motion. As can easily be seen in the figure, similar terms tend to cluster together, whereas terms with clearly distinct meanings are further apart. The positions of the words were generated using multidimensional scaling (MDS) to reduce a 100-dimensional space based on co-occurrence patterns in the British National Corpus (2007). Terms cluster by their category (bird, mammal, verb) and are also related by semantic properties (e.g., “fly” is closer to birds, while “walk” is closer to mammals).

Figure 1. A two-dimensional representation of the relative positions of the words "eagle," "falcon," "lion," "bear," "fly," and "walk."



In addition to the attempts to apply these methods as models of knowledge representation and language processing, similar methods were successfully used by computational linguists. A wide range of applications were proposed, such as answering questions based on texts (Mohler & Mihalcea, 2009), automatic summarization of texts (Yeh, Ke, Yang, & Meng, 2005), machine grading (Foltz, Laham, & Landauer, 1999; Graesser et al., 2000), and cross-linguistics translation (Tam, Lane, & Schultz, 2007). Many of the current technologies employ similar approaches for their semantic processing. Moreover, the underlying characteristics of the space have been demonstrated to correspond to human performance in various tasks, with semantic priming and similarity judgments being especially prominent examples (Günther, Dudschig, & Kaup, 2016; Landauer & Dumais, 1997).

The catalyst which made me look at these methods as possible tools for testing psychological theories, as opposed to modeling processes, was A. Kimball Romney's cultural consensus model (CCM; Romney, Weller, & Batchelder, 1986). CCM is based on the observation that not all cultural informants are created equal. Some are more knowledgeable about the culture than others and will therefore provide "better" answers. CCM uses statistics to identify *consensus* answers—answers that most informants agree on and then uses these answers to gauge the correctness of each informant. Informants are then weighted by the observed quality of their answers, allowing the researcher to more reliably study aspects of the culture that are less well known and widely agreed upon.

Like LSA, the CCM examines a set of data for common patterns. In contrast to LSA, CCM does not aim to provide a *model* of culture or cultural knowledge, but rather simply to identify relevant information hidden in a heap of statistical noise. Nevertheless, the two models have very similar mathematical underpinnings, and languages are often associated with cultures. So why not use methods like LSA not as models of cognition, but as tools for comparing groups, much like we design traditional studies in psychology?

The Traditional Logic of Hypothesis Testing

To properly use a new source of information for statistical inference, we must first be cognizant of the tools we propose to use for its analysis. As mentioned earlier, I intended to use text-based data as stand in for individuals—that is, as if it was a measurement of the person who wrote the text. Essentially, I wanted to follow the standard logic of hypothesis testing, but with a different source of information.

In hypothesis testing, as normally practiced by psychologists, the researcher is interested in whether two (or more) distinct groups of measurements differ from each other. Most frequently, these differences involve some type of experimental treatment that is related to a prediction made by the researcher based on a theory. For example, if a social scientist would like to test the hypothesis that conservatives are more likely to be concerned with the rights of a fetus to life than liberals, they can measure individuals' political leanings and opinions regarding the rights of a fetus and compare the opinions of liberals and conservatives using one of several statistical tests (e.g., *t*-test, ANOVA).

It is important to note that these tests make several assumptions about the underlying measures and how they were collected. Two common assumptions are that the mean of the data follows the normal distribution, and that each measurement is independent of other measures. The latter can often be corrected with appropriate tests (e.g., repeated-measures

ANOVA), but such dependencies need to be explicitly stated in the statistical model.

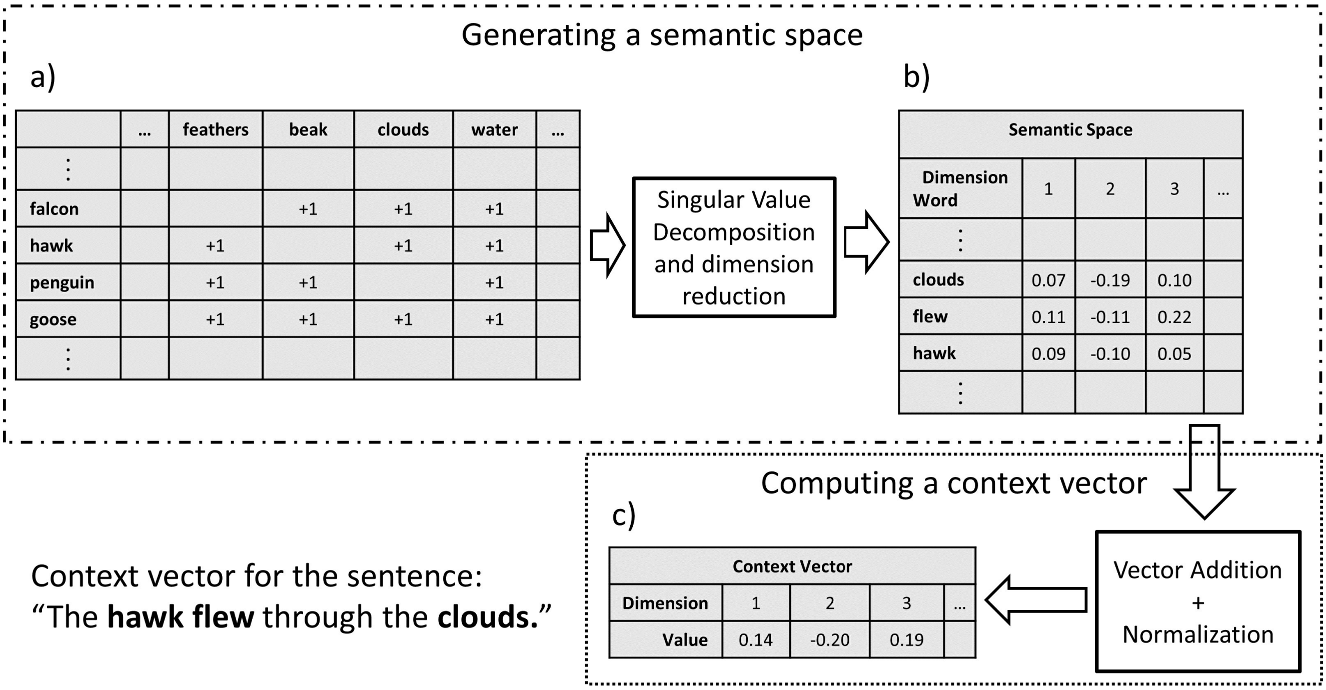
The first assumption can generally be expected to hold for the mean any set of data that is based on a random sampling of independent data points. This follows from the *Central limit theorem*, which states that as the sample size tends to infinity, the distribution of the sample mean will be normal. However, the second assumption, that the data points are independent of each other, is more difficult to ascertain. Nevertheless, in most cases of interest, the source of the measurement (i.e., the origin of the text in our case) needs to be known to a certain degree; otherwise, it would be largely impossible to assign it the appropriate study group. Consequently, it should be possible to identify which measurements originate from the same individual. As that is the level of dependency normally considered as relevant in psychological studies, it stands to reason that it should also be the standard we apply here.

Research Practicalities

It is now time to flesh out the method of measurement. As mentioned earlier, the basic idea is to measure the similarity of words and terms by examining the contexts in which they appear. If two words, for example, “hawk” and “falcon,” appear with similar words in their proximity, such as “feathers” and “clouds,” we will judge them more similar to each other than to a word such as “penguin” which is more likely to appear in proximity to “water” and not “feathers” or “clouds.”

Several mathematical models have been proposed for achieving this goal, and the specifics of these models are less important for this present case than their general goal. Therefore, I will only illustrate one of these approaches, Wordspace, which was developed by Hinrich Schütze (1997, 1998) briefly (see Figure 2). References to more detailed explanations of the methods used in LSA and Topics Modeling are provided in the “Further Reading” section.

Figure 2. An overview of the process of generating and using a semantic space.



	…	feathers	beak	clouds	water	…
⋮						
falcon			+1	+1	+1	
hawk		+1		+1	+1	
penguin		+1	+1		+1	
goose		+1	+1	+1	+1	
⋮						

Semantic Space				
Dimension / Word	1	2	3	…
⋮				
clouds	0.07	-0.19	0.10	
flew	0.11	-0.11	0.22	
hawk	0.09	-0.10	0.05	
⋮				

Context Vector				
Dimension	1	2	3	…
Value	0.14	-0.20	0.19	

Note: (a) A section of a table of co-occurrences generated from a fictional corpus. (b) The resulting semantic space after computing SVD and reducing the number of dimensions. (c) An aggregated vector generated from the sample fragment “the hawk flew through the clouds.” The bolded words indicate the content words used in the computation of the context vector.

The first step in Wordspace is to create a table collating the occurrences of the words in a corpus with other words. Each row in the table represents a word, and the columns represent words which appear within a particular *context window* of it (by default, within 15 words). Because most words never occur near each other (e.g., “falcon” and “atom” are unlikely to co-occur), this table will be populated mostly by zeroes. However, factoring approaches such as *Singular Value Decomposition* (SVD) can be used to identify the latent factors which predict such co-occurrences. When this table is then reconstructed based on only the most important such factors, it will have very few zeroes.

Essentially, these tables represent a multidimensional semantic space in which each column is an axis. Each word therefore corresponds to a position in this space and the distance between words in the space relates to how often they co-occur with similar words (e.g., if both occur with “clouds” with equal frequency, they would be closer together than if they did not, all else being constant). These positions are identified in the table and the extent from the origin point of the space to that point is the *vector* associated with this word in the table. Several studies have demonstrated that the distances measured in such a space correspond with other measures of word similarity (e.g., Maki, McKinley, & Thompson, 2004).

It is also possible to locate a phrase in this space by “averaging out” its component words. This is accomplished by adding the vectors together and then reestablishing the length of the vector to be 1 (which is the normalized length of the vectors representing each word). Individual measurements can therefore be computed from a text using this approach, and each measurement is a *vector* within this space. Importantly, once such a semantic space is constructed, it can be used to compute context vectors for any text—not just texts that were part of the initial corpus. Nevertheless, as language use varies across domains and genres, a space generated from a similar content type to the measured texts is likely to provide better results.

Although standard measurements of human performance are generally described using a single number (a *scalar*), textual measurements as described here involve a vector—an ordered set of numbers. As most standardized statistical approaches compare the means of numbers rather than vectors, it would be useful to compute such numbers from the vectors. Indeed, a number can be derived from each *pair* of textual fragments by computing the distance between them. Because this involves a pair of measurements rather than a single measure, this will affect the type of hypotheses we can readily test and our distributional assumptions.

Method in Action

Are Some Word Forms Associated With Meaning?

I initially applied these methods to questions that were fundamentally linked to language as such a connection made the role the texts played more straightforward. Also, because the statistics were based on the distance between two measurements rather than individual measurements, it was easier to draw hypothesis about the *variance* in a sample rather than its mean. For example, asking whether words consisting of a group were more similar among themselves than to other words translates to a hypothesis that the semantic space distances between words belonging to this group are smaller than those between words in general. Essentially, this hypothesis suggests that there is less variability in the meaning of words within the group compared with a random sample of other words.

Such a hypothesis exists in the case of phonesthemes (“Phonestheme,” 2017)—units of sound that appear to be related to meaning while not having independent semantic content. Prominent examples in English are the prefixes *gl*- (often referring to visual stimuli and light; e.g., glimmer, glimpse, glisten) and *sn*- (frequently associated with the mouth and/or nose; e.g., snout, snicker, snarl). Importantly, many such phonesthemes were suggested in the literature, but there was not much empirical support for them. In particular, although there were examples of words in support of the proposed meaning of each phonestheme, there were also counter examples (e.g., glycerin, snow). It was therefore difficult to determine which phonesthemes, if any, were actually related to an aspect of meaning.

The hypothesis that a particular phonestheme (say, *gl*-) exists is essentially a hypothesis suggesting that words following the phonestheme's form (e.g., all words in English that start with *gl*-) are more likely to share some aspect of their meaning than words in general. That is, they will be more similar to each other than would be expected by chance alone. That is the hypothesis Katya Otis and I (Otis & Sagi, 2008) set out to test in our study.

Because we were interested in testing general uses of words, we used a semantic space generated from a collection of literary works. For this purpose, we put together a corpus of approximately 4,000 literary works from Project Gutenberg (see http://www.gutenberg.org/), which provides electronic versions of classical works that are out of copyright. Although this provided us with a substantial sample of English language (about 290 million words), it was also an older sample, consisting largely of 19th century works.

This means that we could only identify phonesthemes that existed in that period. However, as language change is generally a slow and gradual process, we felt that the corpus was appropriate. Moreover, many of the phonesthemes we were interested in were proposed in the first half of the 20th century and using a sample of language that predated the original proposals by a few decades was arguably a better choice for testing them than a later sample. We have since replicated this analysis using a semantic space based on the British National Corpus (BNC Consortium, 2007), which contains a selection of about 100 million words of British English from texts of the late 20th century. Such replication strengthens the hypothesis that these phonesthemes not only exist, but that they are relatively stable over time.

To test whether words exhibiting a particular phonestheme were more similar to each other than might be expected by chance, we first computed vectors representing all frequently occurring words in the corpus. Next, we identified all words that matched each of 47 phonesthemes that were suggested in the literature. This provided us with our measurements—word vectors representing the range of meaning exhibited by each proposed phonestheme.

To test our hypothesis, we needed to compare the variance of the words exhibiting each phonestheme with the overall variance of words in the corpus. To reduce our dependence on the distributional properties of our samples, we patterned our statistical analysis after the Monte Carlo method. We repeatedly sampled from the pair of all possible words and the pair of words following a particular phonestheme and compared the pairwise distances of the two groups.

Using this approach allowed us to estimate the variability of our sample of pairwise distances as well as its mean, allowing for more rigorous statistical testing. We found that although some phonesthemes were statistically supported, others were not. To further validate the method we were using, we also tested a couple of phonetic units that we were not expecting to be phonesthemes (*z*- and *br*-). As predicted, these two did not turn out as statistically supported phonesthemes.

Framing Effects in Text

Now that we have seen that this method can be applied to questions that relate to the processing of language, I turn to a case that is more closely tied to reasoning and representation—*framing effects*.

As described by Tversky and Kahneman (1981), framing effects are a type of cognitive bias where presenting information in a particular manner affects how people process and react to it. This idea has been extended and used throughout the social sciences and is particularly prevalent in political science where researchers are interested in how the framing of an idea affects its reception by the target audience.

For example, in the United States, the difference between the liberal and conservative position on the topic of abortion can be described as being based on different representation of the individuals involved in the procedure—Liberals focus on the rights of the pregnant individual (i.e., their right to choose), whereas conservatives focus on the right of the developing fetus (i.e., their right to life). Can we apply the idea that texts represent individuals and identify such differences in framing and representation by examining texts produced by people with different views?

I attempted to address this question, together with Daniel Diermeier and Stefan Kaufmann (Sagi et al., 2013). The texts we examined in this study were speeches given by U.S. Senators on the Senate floor, and our semantic space was therefore based on a corpus of such speeches from 1989 to 2006. As we were interested in each senator’s representation of abortion, we identified each occasion where they used the word and examined the context in which it was used. We constructed vectors representing each of these contexts and compared the distances from these vectors with similar vectors for the uses of the words “choice” and “life,” which are key terms representing the expected frames. As predicted, we found that democrats preferred a frame of “choice” (the contexts of “abortion” and “choice” were more similar than those of “abortion” and “life”), whereas republicans preferred a frame of “life.” Following this different focus on the individual whose rights are more important, we also compared the distances of “abortion” from the neutral term “woman” and the more fetusoriented term “mother,” with similar results.

The Reflection of Morality in Language

Following this cognitive-based study, I, together with Morteza Dehghani (Sagi & Dehghani, 2014a, 2014b), applied these ideas to a theoretically driven set of frames about morality, as proposed by Jonathan Haidt and his colleagues (Moral Foundation Theory; Haidt & Joseph, 2004; Koleva, Graham, Iyer, Ditto, & Haidt, 2012). Moral Foundation Theory identifies five distinct moral intuitions or concerns (e.g., caring for others, being loyal, and acknowledging authority). Instead of measuring the reliance of individuals on these intuitions, we examined how they manifest in textual contexts. For example, we identified that in the context of abortion, democratic senators are more likely to be concerned with *fairness*, whereas republican senators are more likely to bring forth ideas based on *purity*. For this purpose, we again used a semantic space based on political speeches given in the senate.

We also examined how these ideas influence the structure of social networks and demonstrated that the concern of *purity* is an important factor for determining who an individual user will choose to follow on twitter (Dehghani et al., 2016). Once we identified this connection, we wanted to substantiate it and examine it outside the confines of text and social networks. We therefore took this result that we identified in a naturally occurring corpus of tweets and constructed a psychological study to examine it further. This study demonstrated that individual preferences regarding purity are not unique to social networks and the Internet, but are prevalent in other contexts as well.

Practical Lessons Learned

Over the course of applying this method to a variety of hypotheses, I gradually realized that each study needs specific modifications and considerations. For instance, when analyzing the use of the word *abortion* by senators, what is the appropriate unit of measurement? Should each use be treated as an independent measure? Should they be aggregated by speaker? By political affiliation?

I decided to be conservative and averaged the semantic distances over all speakers from each party for every year. However, analyzing them independently or by speaker yielded similar results. This made me feel more confident in the robustness of the result, as well as the method. As a general rule, I believe that if the specific result varies greatly depending on how it is aggregated (when the aggregation is of little statistical significant), it should be considered suspect and additional verification be sought.

However, as the number of data points tends to be large compared with traditional psychological studies, that standard of statistical significance should also be modified accordingly. By using the year as a unit of measurement, I ended up with a number of data points that is in line with psychological studies. In contrast, if I would have analyzed the approximately 11,000 contexts of abortion individually, the statistical tests would have likely been highly significant even for a negligible effect size. A possible way to address this, which I employed in my study of phonesthemes, is to report effect sizes as well as significance levels.

Conclusion

Grounding in more than one discipline often highlights some deficiencies in either field, allowing the researcher to devise a useful and non-trivial research program. The method I describe here is a novel application of methods from natural language processing to the study of psychology. In developing it, I drew on domains that I was already familiar with—psychology and linguistics. In addition, developing new methodologies for scientific research requires a careful consideration of their statistical underpinnings. It is especially important to carefully consider and make explicit the assumptions your approach relies on, such as the independence of measurements.

Using a new method to conduct research requires some creativity and flexibility, especially when it comes to selecting appropriate hypotheses. Whereas there is ample evidence for hypotheses appropriate for established methods, a new approach often requires the hypothesis to be adapted. For example, it is easier to use the text-based method I present here to compare the variability of samples than to directly assess differences in means. Nevertheless, given a theoretically grounded set of concepts, such as those provided by Moral Foundation Theory (Haidt & Joseph, 2004), it is possible to adapt the method for measuring scalar values which are more appropriate for standard hypothesis testing statistics (e.g., *t*-test, ANOVA).

Exercises and Discussion Questions

1. How would you use texts to measure personality variables? What would you look for? (e.g., introversion/extroversion)
2. One common application for this type of methods in Natural Language Processing is called *sentiment analysis* and is concerned with identifying whether a text is positive or negative. What psychological hypotheses might be testable using sentiment analysis?
3. What types of hypotheses would be easy to test using language-based data? What types would be difficult?
4. What are the possible consequences of using a non-optimal semantic space for an

analysis? (e.g., using a literary corpus to test a hypothesis about political speeches)

5. What is the relationship between studies conducted in the lab and studies conducted on texts? What are their respective advantages and disadvantages?
6. How can methods based on the analysis of text be used as part of a lab-based experiment? Is such a combination useful?

Further Reading

Iliev, R., **Dehghani, M.**, & **Sagi, E.** (2015). Automated text analysis in psychology: Methods, applications, and future developments. *Language and Cognition*, 7, 265–290.

Griffiths, T. L., **Steyvers, M.**, & **Tenenbaum, J. B.** (2007). Topics in semantic representation. *Psychological Review*, 114, 211–244.

Landauer, T. K., **Foltz, P. W.**, & **Laham, D.** (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284.

Sagi, E., & **Dehghani, M.** (2014a). Measuring moral rhetoric in text. *Social Science Computer Review*, 32, 132–144.

Sagi, E., **Diermeier, D.**, & **Kaufmann, S.** (2013). Identifying issue frames in text. *PLoS ONE*, 8(7), e69185.

Web Resources

LSA @ CU Boulder: http://lsa.colorado.edu

Center of Data on the Mind: http://www.dataonthemind.org/

Project Gutenberg: http://www.gutenberg.org/

References

BNC Consortium. (2007). *British National Corpus version 3* (BNC XML ed., Distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium). Available from http://www.natcorp.ox.ac.uk/

Dehghani, M., **Johnson, K.**, **Hoover, J.**, **Sagi, E.**, **Garten, J.**, **Parmar, N. J.**, … **Graham, J.** (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, 145, 366–375. doihttp://dx.doi.org/:10.1037/xge0000139

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. In **C. E. Bazell** (Ed.), *Studies in linguistic analysis: Special volume of the philological society* (pp. 1–31). Oxford, UK: Blackwell.

Foltz, P. W., **Laham, D.**, & **Landauer, T. K.** (1999). The intelligent essay assessor: Applications to educational technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, 1, 939–944.

Graesser, A. C., **Wiemer-Hastings, P.**, **Wiemer-Hastings, K.**, **Harter, D.**, Tutoring Research Group, & **Person, N.** (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments*, 8, 129–147.

Günther, F., **Dudschig, C.**, & **Kaup, B.** (2016). Latent semantic analysis cosines as a cognitive similarity measure: Evidence from priming studies. *The Quarterly Journal of Experimental Psychology*, 69, 626–653.

Haidt, J., & **Joseph, C.** (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.

Koleva, S. P., **Graham, J.**, **Iyer, R.**, **Ditto, P. H.**, & **Haidt, J.** (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, 46, 184–194.

Landauer, T. K., & **Dumais, S. T.** (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240. doi:http://dx.doi.org/10.1037/0033-295X.104.2.211

Landauer, T. K., **Foltz, P. W.**, & **Laham, D.** (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284.

Maki, W. S., **McKinley, L. N.**, & **Thompson, A. G.** (2004). Semantic distance norms computed from an electronic dictionary (WordNet). *Behavior Research Methods, Instruments, & Computers*, 36, 421–431. doi:http://dx.doi.org/10.3758/BF03195590

Mohler, M., & **Mihalcea, R.** (2009). Text-to-text semantic similarity for automatic short answer grading. In **Gardent, C.**, & **Nivre, J.** (Eds.), *Proceedings of the 12th conference of the European chapter of the association for computational linguistics* (pp. 567–575). Stroudsburg, PA: Association for Computational Linguistics.

Otis, K., & **Sagi, E.** (2008). Phonaesthemes: A corpus-based analysis. In **B. C. Love**, **K. McRae**, & **V. M. Sloutsky** (Eds.), *Proceedings of the 30th annual meeting of the cognitive science society* (pp. 65–70). Austin, TX: Cognitive Science Society.

Pennebaker, J., **Chung, C.**, **Ireland, M.**, **Gonzales, A.**, & **Booth, R.** (2007). *The development and psychometric properties of LIWC2007*. Austin, TX: LIWC. Available from www.Liwc.Net

Phonestheme. (2017, January 3). *In Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Phonestheme&oldid=758135051

Pickering, M. J., & **Garrod, S.** (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–189.

Romney, A. K., **Weller, S. C.**, & **Batchelder, W. H.** (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, 88, 313–338.

Sagi, E., & **Dehghani, M.** (2014a). Measuring moral rhetoric in text. *Social Science Computer Review*, 32, 132–144. doi:http://dx.doi.org/10.1177/0894439313506837

Sagi, E., & **Dehghani, M.** (2014b). Moral rhetoric in Twitter: A case study of the US federal shutdown of 2013. In **P. Bello**, **M. Guarini**, **M. McShane**, & **B. Scassellati** (Eds.), *Proceedings of the 36th annual conference of the cognitive science society* (pp. 1347–1352). Austin, TX: Cognitive Science Society.

Sagi, E., **Diermeier, D.**, & **Kaufmann, S.** (2013). Identifying issue frames in text. *PLoS ONE*, 8(7), e69185.

Schütze, H. (1997). *Ambiguity resolution in language learning: Computational and cognitive models*. Stanford, CA: CSLI Publications.

Schütze, H. (1998). Automatic word sense discrimination. *Computational Linguistics*, 24, 97–123.

Tam, Y.-C., **Lane, I.**, & **Schultz, T.** (2007). Bilingual LSA-based adaptation for statistical machine translation. *Machine Translation*, 21, 187–207.

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.

Tversky, A., & **Kahneman, D.** (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458. doi:http://dx.doi.org/10.1126/science.7455683

Yeh, J.-Y., **Ke, H.-R.**, **Yang, W.-P.**, & **Meng, I.-H.** (2005). Text summarization using a trainable summarizer and latent semantic analysis. *Information Processing & Management*, 41, 75–95.