

Love, Hope, Perspective, and Leadership in the Ngram Database: Solace for Modern Times

William H. Zywiak^{1,2}, Gao Niu¹

¹Mathematics Department, Bryant University, Smithfield, USA

²Center for Alcohol and Addiction Studies, Brown University, Providence, USA

Email: wzywiak@bryant.edu

How to cite this paper: Zywiak, W. H., & Niu, G. (2021). Love, Hope, Perspective, and Leadership in the Ngram Database: Solace for Modern Times. *Open Journal of Social Sciences*, 9, 159-166.

<https://doi.org/10.4236/jss.2021.911013>

Received: October 20, 2021

Accepted: November 21, 2021

Published: November 24, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The Ngram database allows searches that identify words that are top of mind in a culture at a specific time. We used this database to discover the character strengths that have been loudest in the conversation of the American English publication record during the last 100 years, and in particular in 2019. Love, perspective, and leadership have been increasing since 1920 and hope has been increasing since 1987. Results suggest that these character strengths could be increased even more with greater emphasis upon them. The Ngram database was used to filter the 24 constructs identified by Peterson and Seligman (2004) perhaps identifying four of the most important constructs now and ever.

Keywords

Love, Hope, Perspective, Leadership, Ngram

1. Introduction

Depression rates are increasing since 2005 (Weinberger et al., 2017) and the media and politicians in the United States often engage in divisive rhetoric. To counter these effects, we sought to identify character strengths that are prominent in the United States literature and also to predict changes in these character strengths in the future. More specifically, we sought to see what character strengths were prominent, since they could be amplified further to counter the negative influences of pessimism and despondency. Peterson and Seligman (2004) oversaw a comprehensive examination of six core values (“wisdom and knowledge”, courage, humanity, justice, temperance, and transcendence) and 24 character strengths (e.g., perspective) nested under these core values. For data we used Google’s Ngram Viewer database (Aiden & Michel, 2013; Michel et al.,

2011a; Zywiak, Bobroff, & Niu, 2021). Aiden and Michel (2013) introduced the world to Google's Ngram Viewer through a best-selling book and a Science article (Michel et al., 2011a).

The Ngram database includes words and phrases, up to 5 words, from books created between 1500 and 2019. These words and phrases are anchored to the year of publication. Sources include over 5 million books. The Ngram software plots the frequency of a given word across a selected time span. The authors have been having undergraduate students in statistics courses use this database to conduct and interpret regression analyses and lagged correlation analyses. The sheer volume of the dataset allows students to conduct analyses on terms that are of interest to them. This data is open access and is available if you google Ngram Viewer.

Aiden and Michel (2013) previously focused on fame (see 4.4. in the discussion), tragedies (Lusitania, Pearl Harbor), war (the U.S. Civil War, the Great War, World War II), and diseases (e.g., cancer, TB, diabetes, heart attack, Michel et al., 2011b). Aiden and Michel's focus on negative concepts and our focus on positive concepts, parallels psychology's original focus on psychopathology (e.g., the work of Freud) and the balance of a focus on positive aspects of the mind by third force psychologists (Maslow, 1943; Rogers, 1961).

2. Method

In this paper we used a single corpus which is the typical approach in Ngram studies. We used the corpus of American English, based on familiarity. The raw data were the frequencies of the 24 character strengths. Peterson & Seligman (2004) and colleagues conducted an exhaustive and comprehensive review of character strengths explicated in historical literature (e.g., religious texts, philosophy, and more recently psychology). In a few cases we used synonyms if the term used by Peterson and Seligman contained more than one word ("appreciation of beauty and excellence" was replaced with "aesthetics") or the term had prominent different meanings ("citizenship" was replaced with "civics"). Data were retrieved from Ngram Viewer from 1920 through 2019 every 5 years, including 2019 (i.e., 21 data points). We chose a period of 99 years, since the most recent easily accessed data includes 2019, 100 years is a very common time frame for retrospectives in America (Moore & Pitlor, 2015; Peter Pauper Press, 2009; Winkowski & Sullivan, 1998) and Aiden and Michel (2013) used 100 years for many of their ngram graphs. We set smoothing to zero to be able to view the pattern in the raw data. Six digits after the last leading zero were recorded. We first determined the four most prevalent character strengths as of 2019 and these were: love, hope, perspective, and leadership. These fell under the core values of humanity, transcendence, "wisdom and knowledge", and justice, respectively. Data were analyzed using regression analyses in SPSS 26.

3. Results

The year was used to predict the frequency of each term. We multiplied the

Ngram percentages by 10,000 (and dropped the %) to avoid rounding errors on the beta weights in the resulting regression models, and so we could report frequency per million words following the convention of Aiden and Michel (2013), please see Figure 1. A linear transformation: adding or multiplying the dependent variable by a constant does not affect the p -value or R^2 . Significant results were found for three of the four terms (all but hope) and these are reported in Table 1. The frequency of hope ebbed in the 70s and 80s. We conducted an extra regression analysis for hope for the period from 1975 through 2019 (with the minimum required 10 data points for a regression). This second analysis was statistically significant: $F(1,8) = 14.82$, $p = .005$, $m = .864$, $b = -1647.127$.

4. Discussion

4.1. Love

Deconstructing love into Greek terms, in order of 2019 prevalence we have: agape, eros, philia, pragma, ludus, and philautia (Kzarnic, 2011, please see Figure 2). Agape or self-less love represents 49.5% of the overall prevalence of these six terms in 2019. This term overlaps with the idealized version of motherly love as well as Rogers' (1961) unconditional positive regard. Eros [lust (as a thought) and sex (as a behavior)] represents 30.6%, and philia, or friendship, represents

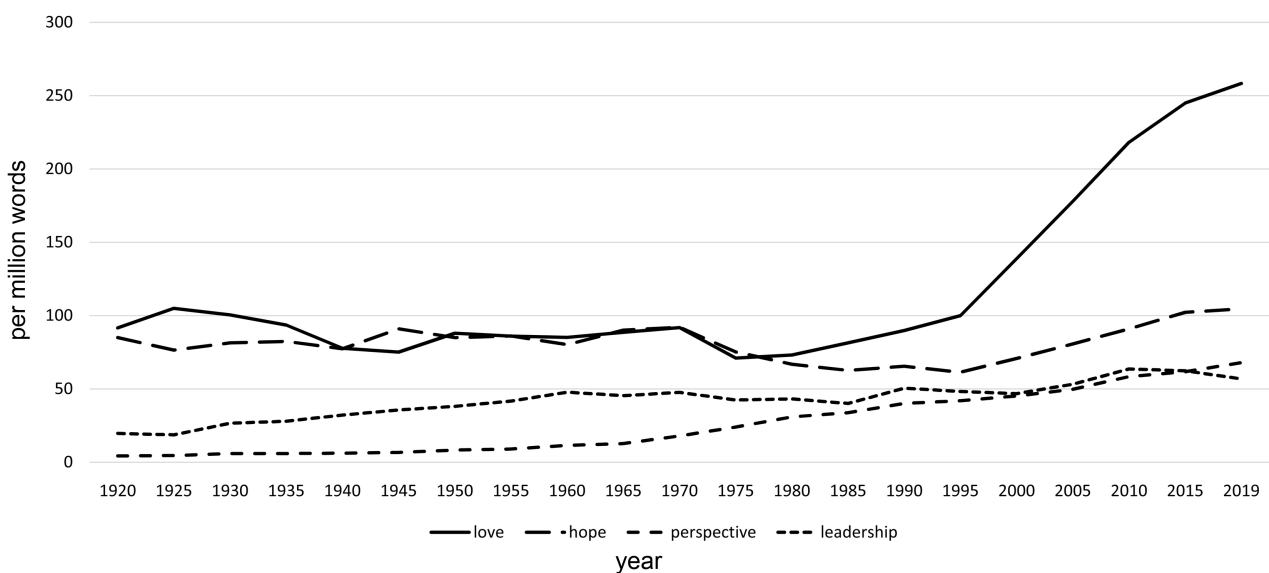


Figure 1. Love, hope, perspective, leadership: frequency per million words by year.

Table 1. Regression results.

	2019 value	R^2	p	$m (\beta_1)$	$b (\beta_0)$
love	258.262	.44	.001	1.272	-2390.449
hope	104.511	.00	.75	NA	NA
perspective	67.8564	.91	<.001	.661	-1275.896
leadership	56.7179	.85	<.001	.369	-685.196

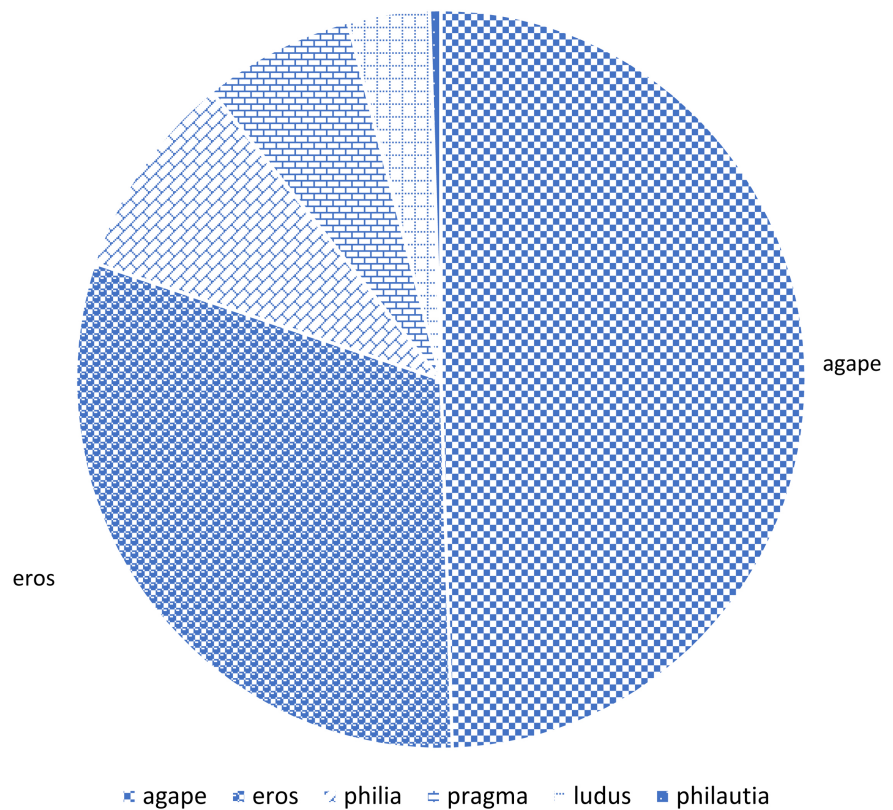


Figure 2. Six Types of Love, % per word presented in a clockwise order.

9.1%. Pragma, ludus, and philautia, represent 6.7%, 3.6%, and 0.5%, respectively. Pragma is evident in marriages that last decades, or mature love. Ludus refers to playful love, or flirtation. Philautia refers to being able to love yourself (so that you can love others, Kzarnic, 2011). All of these contribute to the sustainability of society, but agape and eros are getting the greatest emphasis in the American English corpus.

4.2. Hope

The data from **Figure 1** suggest that “hope” enters a tailspin in the early seventies. We take a more fine-grained look at this pattern by plotting annual data from 1970 to 2010 in **Figure 3**. The figure shows that “hope” decreases in early 1970s, hits a low in 1987, and rebounds to the 1970 level by 2009. Historical events that may have contributed to this deterioration in hope include the assassinations of MLK Jr. and RFK in 1968, the National Guard shooting war protestors at Kent State in 1970, Vice President Spiro Agnew’s resignation in 1973 associated with accusations of political corruption, and President Richard Nixon’s resignation in 1974 following the Watergate scandal. During this period there was also widespread coverage of the Vietnam War on the evening news, as well as the jarring photo of Phan Thi Kim Phuc, a young Vietnamese girl burned by napalm in 1972. In 1975 Saigon, fell to the North. American soldiers were humiliated by the public during their return. The hopeful music of the flower power

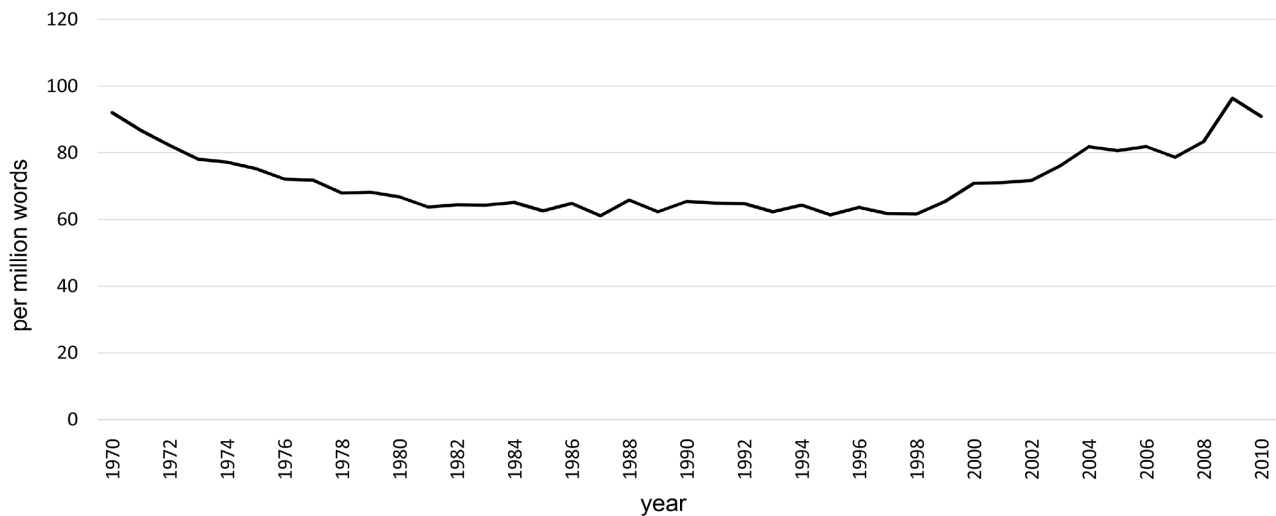


Figure 3. Hope from 1970 to 2010.

period was replaced by the despondency and malaise of Hotel California. Admittedly there were some signs of hope during this time; the first Earth Day was celebrated in the US in April 1970. In 1971, Secretary of State Henry Kissinger visited China. In 1975, there was the orbiting handshake between the U.S. Apollo space module and the USSR Soyuz space module.

4.3. Perspective

There are various dimensions that can be used to typify perspective. The humanistic psychologist Maslow (1943) describes five sets of goals 1) physiological (e.g., food and water) 2) safety 3) love 4) esteem (achievement, adequacy, confidence, reputation, and prestige) and 5) self-actualization (to be the best at what one is best at, to give to one's community). When each successive basic need is met, the person can pursue higher needs to a greater extent. Eckhart Tolle's (2004) *"The Power of Now"* is all about maintaining perspective. Neurologically, one can see the distinction between the reptile brain and the cerebral cortex. In this context addictive behavior can be seen as a short cut to activating the reward system (e.g., the limbic system) in an anti-evolutionary way. The limbic system developed to help animals thrive, and in the case of humans to thrive within the society and to help the society thrive. Triggering the reward system through addictive behavior sabotages the reason for the limbic system. Another dimension that can be considered in the context of perspective is geographical and is defined by these areas: individual, couple, household, community, state, nation, and globe. For instance, many industries may be wonderful for a nation's economy, but terrible for the climate of earth. In Ngram Viewer, the term perspective eclipses leadership during the last few years, suggesting the impact of globalization, and possibly the evolution of human consciousness.

4.4. Leadership

The term leadership increases consistently during the last 100 years. It should be

noted that this may not be applauding good leadership, but may be critiquing current leaders. The characteristics necessary for a good leader will be determined to some extent by the organization s/he is leading. A good president, will be different than a good pope, who will be different than a good general. In some ways, a good leader must emanate the epitome of characteristics that the specific organization holds valuable. Other great leadership characteristics may generalize across different types of organizations.

Aiden and Michel (2013) review leaders of each birth year as defined by the frequency with which their names appear in the Ngram corpus. This identifies such “leaders” (and some self-actualized people) as (in chronological order) Winston Churchill, Thomas Mann, Pierre Teilhard de Chardin, Ezra Pound, Ho Chi Minh, Mao Zedong, Margaret Mead, George Orwell, Laurence Olivier, Mother Teresa, Henry Kissinger, MLK Jr., Lee Harvey Oswald, Barbara Streisand, and Clarence Thomas. The appearance of Lee Harvey Oswald in this list exemplifies that the appearance of the term leadership in our top four character strengths, may be criticizing some leaders, i.e., that the term may have a negative valence.

Leadership is a focus of many bestsellers and these include Sun Tzu’s (2005) *The Art of War*, Valerian Albanov’s (2000) *In the Land of White Death*, Dale Carnegie’s (1936) *How to Win Friends and Influence People*, Stephen Covey’s (1989) *The 7 Habits of Highly Effective People*, and Jim Collins’s (2001) *Good to Great*.

4.5. Conclusion

The good news is that the four most prominent character strengths are increasing in frequency in the American English corpus, and possibly the American culture at large. Statistically, perspective and leadership most closely follow a linear function, as reflected in the R^2 s approaching one. Results suggest that efforts to boost love, hope, perspective, and leadership among individuals may be fruitful. Further research could use this same approach to study the British English corpus, and native speakers of other languages could search the various other corpora available through Ngram Viewer using relevant terms. More specifically, it could be hypothesized that the concept of leadership might be more prevalent in cultures where collectivism (Triandis, Bontempo, Villareal, Asia, & Lucca, 1988) receives greater emphasis than individualism, i.e., where organizations are more highly valued. The two authors have students at Bryant University use the Ngram database to conduct a statistics project on constructs that are of interest to them. We have used the Ngram database to filter the 24 constructs identified by Peterson and Seligman (2004) and perhaps have identified four of the most important constructs now and ever. Bryant University strives to increase character strengths in its community by having all Freshmen take GFCL 100—Global Foundations of Character and Leadership. Different faculties tailor the course based on their expertise and interests. It is notable that one such fa-

culty requires the reading and discussion of Stephen Covey's (1989) best-seller cited in this article. Other universities offer World Religions as an elective.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Aiden, E., & Michel, J.-B. (2013). *Uncharted: Big Data as a Lens on Human Culture*. Penguin. (Riverhead)
- Albanov, V. (2000). *In the Land of White Death: An Epic Story of Survival on the Siberian Arctic*. Random House.
- Carnegie, D. (1936). *How to Win Friends and Influence People*. Simon and Schuster.
- Collins, J. (2001). *Good to Great: Why Some Companies Make the Leap ... and Others Don't*. HarperCollins.
- Covey, S. (1989). *The 7 Habits of Highly Effective People: Restoring the Character Ethic*. Free Press.
- Kzarnic, R. (2011). *How Should We Live? Great Ideas from the Past for Everyday Life*. Blue Bridge.
- Maslow, A. (1943). A Theory of Human Motivation. *Psychological Review*, 50, 370-396. <https://doi.org/10.1037/h0054346>
- Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011a). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, 331, 176-182.
- Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011b). *Supporting Online Material for Quantitative Analysis of Culture Using Millions of Digitized Books*.
- Moore, L., & Pitlor, J. (2015). *100 Years of the Best American Short Stories*. Houghton Mifflin Harcourt Publishers.
- Peter Pauper Press (2009). *A Century of Christmas Memories: 1900-1999*. White Plains, Peter Pauper Press.
- Peterson, C., & Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook of Classification*. Oxford University Press.
- Rogers, C. R. (1961). *On Becoming a Person*. Houghton Mifflin.
- Tolle, E. (2004). *The Power of Now: A Guide to Spiritual Enlightenment*. Namaste Publishing.
- Triandis, H., Bontempo, R., Villareal, M., Asi, M., & Lucca, N. (1988). Individualism and Collectivism: Cross-Cultural Perspectives on Self-Group Relationships. *Journal of Personality and Social Psychology*, 54, 323-338. <https://doi.org/10.1037/0022-3514.54.2.323>
- Tzu, S. (2005). *The Art of War*. Shambala.
- Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., & Goodwin, R. D. (2017). Trends in Depression Prevalence in the USA from 2005 to 2015: Widening Disparities in Vulnerable Groups. *Psychological Medicine*, 48, 1308-1315.

<https://doi.org/10.1017/S0033291717002781>

Winkowski, F., & Sullivan, F. D. (1998). *100 Planes 100 Years*. Smithmark.

Zywiak, W. H., Bobroff, R. P., & Niu, G. (2021). Black Swan Years in American English, French, German, Hebrew, and Russian: Years That Reverberate in Ngram Viewer. *Advances in Historical Studies*, 10, 208-214. <https://doi.org/10.4236/ahs.2021.103013>