Optimizing Online Suicide Prevention: A Search Engine-Based Tailored Approach

Florian Arendt and Sebastian Scherr

Department of Communication Science and Media Research, University of Munich (LMU)

ABSTRACT

Search engines are increasingly used to seek suicide-related information online, which can serve both harmful and helpful purposes. Google acknowledges this fact and presents a suicide-prevention result for particular search terms. Unfortunately, the result is only presented to a limited number of visitors. Hence, Google is missing the opportunity to provide help to vulnerable people. We propose a two-step approach to a tailored optimization: First, research will identify the risk factors. Second, search engines will reweight algorithms according to the risk factors. In this study, we show that the query share of the search term "poisoning" on Google shows substantial peaks corresponding to peaks in actual suicidal behavior. Accordingly, thresholds for showing the suicide-prevention result should be set to the lowest levels during the spring, on Sundays and Mondays, on New Year's Day, and on Saturdays following Thanksgiving. Search engines can help to save lives globally by utilizing a more tailored approach to suicide prevention.

We ask you, the reader, to imagine the following scenario: Consider you are suffering from a deep crisis. There are several causes, including—but not limited to—the death of a loved one, and financial problems due to job loss. You are facing feelings of loneliness, helplessness, and hopelessness. In your eyes, the whole world is against you. You perceive everything other people say as being against you, and you believe that nobody cares about you, as if you were wearing glasses that bathed your environment in a negative, dark light. You might think about self-harming through excessive alcohol consumption, or by consuming illegal drugs or other chemicals—maybe as a "cry for help"—or you might even seriously think about taking your own life. You are "on the edge" and ruminate: Should I or should I not?

Common wisdom, as well as previous research (Joiner & Rudd, 2002), suggests that input from the environment (e.g., family, friends, coworkers, or the media) is of greatest importance during this phase of life. Information on where to find help, such as telephone counseling services or crisis chat rooms, can save the lives of thousands of people around the globe each year (Lester & Rogers, 2012). Importantly, suicides are preventable (Wasserman, 2016). If you read this paragraph and decide to make use of professional help, you have a good chance of being able to overcome your current crisis.

Search engines are increasingly used to seek suicide-related information online, which can serve both harmful and helpful purposes. Google acknowledges this fact and presents a suicide-prevention result for particular search terms. Unfortunately, the result is only presented to a limited number of visitors. Hence, Google is missing the opportunity to provide help to vulnerable people. We propose a two-step approach to a tailored optimization of search engines' algorithmic decision making. The goal is to increase the frequency of the suicide-prevention result showing up, especially for those in high-risk situations. The optimized detection of vulnerable individuals is of the greatest importance because search engines are perhaps the only stakeholder that can adequately react and provide immediate help at the exact moment when a vulnerable person is seeking suicide-related information online, at any time, 24/7.

Online Suicide Prevention

The media are considered as a key factor in the dissemination of such helpful information (Mann et al., 2005; Niederkrotenthaler et al., 2010). An increasing amount of literature acknowledges the importance of the Internet for suicide prevention (Mehlum, 2000). Especially, search engines such as Google have become powerful gatekeepers (Shoemaker & Vos, 2009) for suicide-related information (Gunn & Lester, 2013). Importantly, suicide-related information seeking via search engines can serve both harmful and helpful purposes (Biddle, Donovan, Hawton, Kapur, & Gunnell, 2008): Online users can search for harmful information, for example, detailed descriptions of suicide methods. At no time in human history has it been so easy to find detailed information about how to die by suicide in a fast, uncomplicated, and anonymous way. However, online users can also seek out helpful information when feeling suicidal. One primary strategy, which has been recommended in international guidelines that were developed based on previous scholarly work, concerns referrals to crisis-intervention professionals (Draper, Murphy, Vega, Covington, & McKeon, 2015).

CONTACT Florian Arendt florian.arendt@ifkw.lmu.de, florian.arendt@univie.ac.at Department of Communication Science and Media Research, University of Munich (LMU), Oettingenstr. 67, Munich 80538, Germany.

Google—currently the most popular search engine (Pew Research Center, 2012)—has already acknowledged the importance of online suicide prevention and presents information seekers with a “suicide-prevention result” (Zeiger, 2010). For certain search queries, Google prominently presents a suicide-prevention result above all others that depicts important online and offline resources for acute cases of increased suicidality, such as country-specific helpline telephone numbers and relevant websites (Cohen, 2010). Unfortunately, this suicide-prevention result is only presented to a limited number of vulnerable individuals. One previous study found, for example, that the suicide-prevention result was shown in only 11% of all cases involving helpful search terms (e.g., “help when having suicide ideas”) and in only 25% of all search queries where harmful search terms were used (e.g., “best method for suicide”) (Haim, Arendt, & Scherr, in press). Even though it is undoubtedly laudable that Google is supporting and investing in suicide prevention, there is clearly room for considerable improvement.

A Tailored Optimization

We outline a two-step approach for a tailored optimization of search engines’ algorithmic decision making. The ultimate anifold, but should be targeted at variables that can enable the adjustment of the search engines’ algorithms. For example, the knowledge of peaks in suicidal behavior on specific days could be employed as an adjustment factor during machine-based decision making. As online search behaviors are reliably indicative of behavior in the real world, while also being more up-to-the-minute than official statistics are (Ginsberg et al., 2009; Preis, Moat, Stanley, & Bishop, 2012), peaks in search queries regarding certain harmful search terms might be indicative of high-risk times. Second, search engines should reweight algorithmic decision making according to the ascertained factors. Although algorithms are complex (Lazer, 2015), they are all man-made, and can thus be adjusted. In fact, algorithm-based decision-making by search engines is constantly tested and improved upon (Lazer, Kennedy, King, & Vespignani, 2014).

It is important to note that we are not arguing that search engines should stop presenting the suicide prevention result on days not identified as high-risk times. Conversely, our aim is that search engines generally increase the frequency of presentation of the suicide prevention result, but especially on high-risk days. Although search engines should ideally present a suicide prevention result whenever a risk term such as “suicide” has been entered, search engines are run by market-driven companies with economic interests. It comes to no surprise that they are very skeptical when it comes to the change of their algorithmic decision making. The aim of the tailored approach outlined in this article is to put emphasis on identified risk factors first, which should lead to a general improvement with beneficial consequences for public health second.

We illustrate the proposed approach in the context of suicidal behavior by poisoning. Poisoning is one of the most widely used suicide methods: For example, in total, of 1,061,277 attempted suicides and 3,790 completed suicides by poisoning were registered in the United States between 2006 and 2010 (Beauchamp, Ho, & Yin, 2014). In order to identify poisoning-related high-risk times, we investigated temporal variations in poisoning-related online information seeking.

Research Question

Epidemiologically well-known peaks in actual suicidal behavior by poisoning guided the identification process (i.e., spring, Sundays and Mondays, New Year’s Day, and days following family-oriented holidays; see Christodoulou et al., 2012; Beauchamp et al., 2014). It could be interpreted that spring, Sundays and Mondays, and New Year’s Day all have a similar symbolic meaning as they all represent a new beginning. Some individuals may hold out the hope of things getting better in their lives—especially during the holidays with increased family contact—but may have become disappointed when their situation remains unchanged as they approach the new period of time (Gabennesch, 1988).

RQ1: Does query share of the search term “poisoning” on Google show substantial peaks corresponding to peaks in actual suicidal behavior?

Method

Online information seeking was conceptualized as query share: Data are based on the search term “poisoning” users entered into Google. We used the search term “poisoning” due to its direct conceptual correspondence to suicidal behavior by poisoning. Data between December 2009 and January 2016 were used for the region of the United States (see Appendix). All analyses used averaged scores across this time period to partial out idiosyncrasies of single years. We downloaded the raw data for “poisoning” from Google Trends on February 13, 2016.

Google Trends data reflect the frequency of particular search terms entered into the search engine relative to the total search volume. The website provides normalized scores, with 100 being defined as the peak search volume for the requested time periods, search terms, and regions. For example, if a 2-month period is chosen for the search term “poisoning” (e.g., December 2015–January 2016), the day with the highest relative search query volume got the value 100 within the requested region (e.g., the United States). Query volume of the other days is divided by the query volume of the day with the highest query volume and multiplied by 100 (Fond, Gaman, Brunel, Haffen, & Llorca, 2015). For example, a query share value of 50 represents 50% of the highest observed search proportion during the observation period. Query share indirectly corrects for Internet access and population size.

Results

Research question 1 asked whether query share of the search term "poisoning" on Google shows substantial peaks corresponding to peaks in actual suicidal behavior. To answer this research question, we investigated variations in query share as a function of these peak days.

Season elicited a highly significant main effect, $F(3, 309) = 30.51$, $p < .001$. Spring ($M = 76.80$, $SD = 7.74$) showed higher values than winter ($M = 73.61$, $SD = 6.62$), $t(154) = 2.76$, $p = .006$, summer ($M = 70.01$, $SD = 7.80$), $t(159) = 5.54$, $p < .001$, and fall ($M = 66.73$, $SD = 5.15$), $t(152) = 9.45$, $p < .001$. This finding is visualized in Figure 1. As already noted, this spring peak has also been revealed for method-specific suicidal behavior by poisoning (Beauchamp et al., 2014).

Figure 2 shows that the query share was highest on Sundays and Mondays. Sunday ($M = 79.41$, $SD = 9.60$), $t(624) = 10.73$, $p < .001$, and Monday ($M = 76.74$, $SD = 9.05$), $t(624) = 7.50$, $p < .001$, showed increased levels of online information seeking compared to Tuesday ($M = 71.05$, $SD = 9.90$). Sunday showed the highest query share from all days, including in a direct comparison with Monday, $t(624) = 3.58$, $p < .001$. Comparisons for the other days can be easily made when interpreting the confidence intervals in Figure 2.

The figure also shows—visualized as a scatter plot on the top right—that there was a strong correspondence between actual suicidal behavior by poisoning data obtained from previous research (Beauchamp et al., 2014) and the "poisoning"-related online information-seeking measure obtained in the present study. Interestingly, actual suicidal behavior by poisoning was also strongest on Sundays and Mondays according to Beauchamp and colleagues (2014).

Previous research revealed that there was an epidemiologically well-known peak in actual suicidal behavior by poisoning on New Year's Day (Beauchamp et al., 2014). Our analysis replicates this pattern using Google Trends data: Figure 3 shows online information seeking for the period of December to January, averaged for the seven turns of the year during the observation period. New Year's Day clearly showed the highest query share value. New Year's Day ($M = 99.57$, $SD = 0.79$) showed a higher query share compared to December 31 ($M = 72.43$, $SD = 3.51$), $t(6) = 20.61$, $p < .001$, January 2 ($M = 84.71$, $SD = 11.21$), $t(6) = 3.34$, $p = .016$, and January 3 ($M = 81.29$, $SD = 10.70$), $t(6) = 4.34$, $p = .005$.

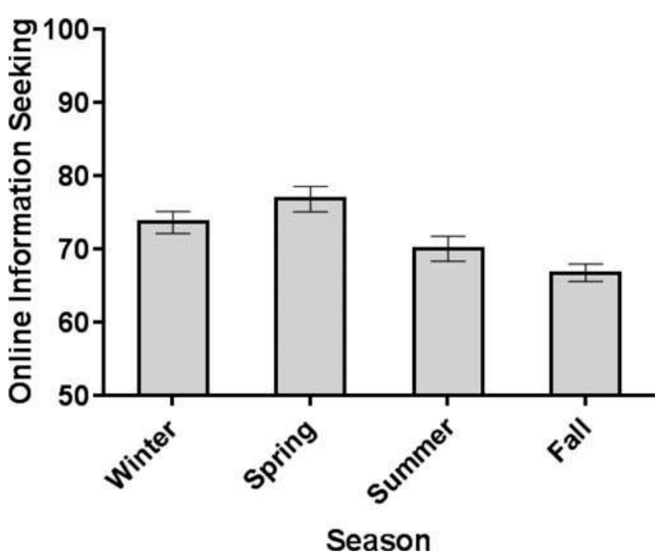

Figure 1. Seasonal variations in online information seeking. Error bars indicate the 95% confidence interval.

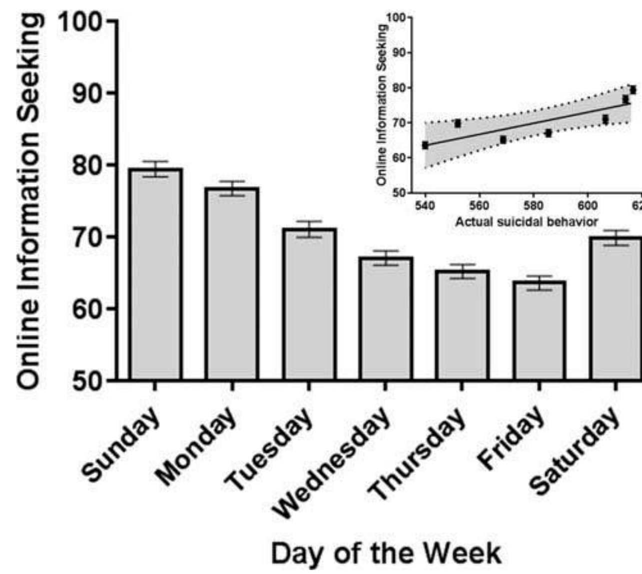

Figure 2. Daily variations in online information seeking. Each day represents the average within the observation period of a total of $N = 2191$ days. Error bars indicate the 95% confidence interval. The top right chart illustrates a scatter plot of actual suicidal behavior by poisoning (Beauchamp et al., 2014) against online information seeking.

We tested query share differences around a further family-oriented holiday: Thanksgiving. As can be seen in Figure 4, there was an increase in online information seeking on the Saturday following Thanksgiving, which is always celebrated on a Thursday. The Saturday ($M = 90.00$, $SD = 2.45$) following Thanksgiving showed a significantly higher value than the Saturday 1 week before ($M = 77.33$, $SD = 5.28$), $t(5) = 8.33$, $p < .001$, the Saturday 2 weeks before ($M = 73.33$, $SD = 6.15$), $t(5) = 7.79$, $p = .001$, and the Saturday 3 weeks before Thanksgiving ($M = 74.00$, $SD = 2.83$), $t(5) = 12.92$, $p < .001$. Importantly, the increase was limited to the Saturday immediately after Thanksgiving.

Discussion

Search engines are increasingly used to seek suicide-related information online, which can serve both harmful and helpful purposes. We propose a two-step approach to a tailored optimization of online suicide prevention: First, research will identify the risk factors. Second, search engines will reweight algorithms according to the risk factors. We recommend that search engines should generally increase the frequency of presentation of the suicide prevention result—especially at identified high-risk times. In this pilot study, we show that the query share of the search term "poisoning" on Google shows substantial peaks in spring, on Sundays and Mondays, on New Year's Day, and on Saturdays following Thanksgiving.

Our findings emphasize the importance of online suicide prevention: Vulnerable individuals who experience suicidal ideation on a Saturday (e.g., following Thanksgiving) and Sunday may have limited access to psychiatric health care professionals, and could

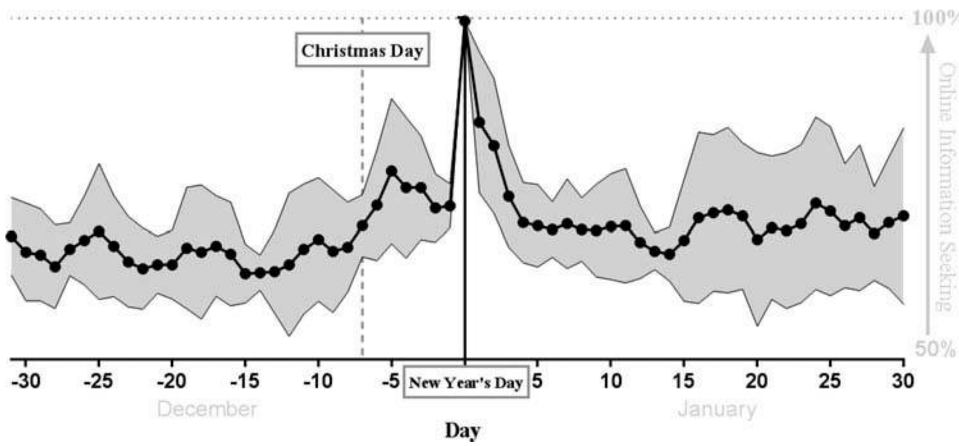

Figure 3. Daily fluctuation of "poisoning" queries around New Year's Day. Each day represents the average from seven turns of the year. The area indicates the 95% error band. The dashed gray vertical line represents Christmas Day (December 25).

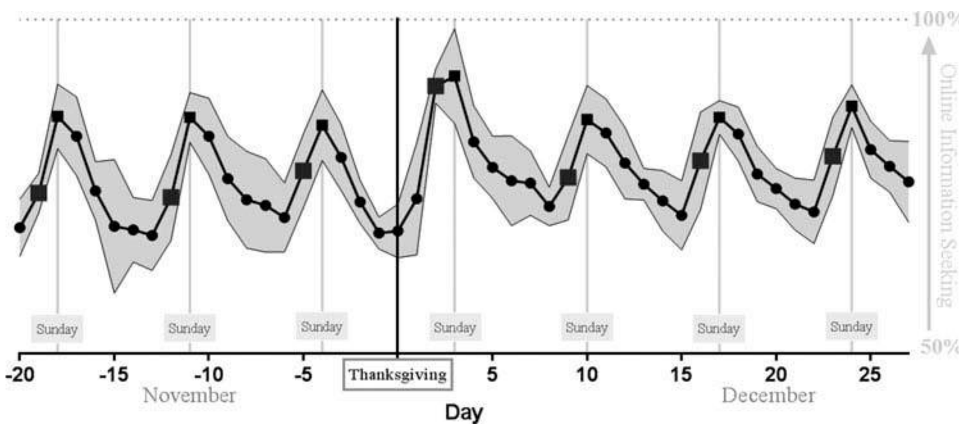

Figure 4. Daily fluctuation of "poisoning" queries around Thanksgiving. Each day represents the average from the 6-year observation period. Thanksgiving is always celebrated on a Thursday. Therefore, the figure shows the daily pattern clearly (see Figure 2). The area indicates the 95% error band. Big squares indicate Saturdays.

therefore benefit from helpful online content that is available 24/7. Especially at weekends, information provided by search engines may have an increased potential to tip the scales between life and death in favor of choosing life.

The fact that peak days indicate higher risk times is the result of the first step of the proposed two-step approach: We recommend setting search engines' algorithmic thresholds to lower levels especially in spring, on Sundays and Mondays, on New Year's Day, and on Saturdays following Thanksgiving (region: United States). The second step—the reweighting of algorithmic decision making according to the identified factors—rests in the hands of information-age hegemons such as Google. Importantly, search engines can help to save lives globally by utilizing a more tailored approach to suicide prevention.

Limitations

As with all research projects, this study has its limitations. First, we investigated the query share of one search term (i.e., "poisoning") only. Thus, our pilot study is limited to the poisoning context. We illustrated the proposed approach in the context of suicidal behavior by poisoning, because poisoning is one of the most widely used suicide methods (Beauchamp et al., 2014). However, future studies should identify high-risk times in other contexts as well.

Second, we relied on data provided by Google Trends. Google provides relative query share data. Although Google Trends data are not created for scientific research (see Lazer et al., 2014), research has shown that query share data are reliable predictors of online information seeking (Nuti et al., 2014). Moreover, the findings of the present study are stable (see the 95% error bands in the figures). However, data were used for the region of the United States. Thus, we cannot generalize findings to other regions (e.g., analyses for smaller countries).

Conclusion

In 1776, Isaac Newton used the phrase "standing on the shoulder of Giants," referring to the fact that scientific knowledge is cumulative and is generated by building on previous discoveries (Historical Society of Pennsylvania, 2016). Importantly, for some of its scholarly services, Google uses this quote as its motto. In accordance with this motto, the search engine prominently presents a suicide-prevention result. Google deserves great credit for this. Unfortunately, helpful information is only presented to a limited number of vulnerable individuals due to the current algorithm's lack of precision in identifying them. Although the meaning of numerous search terms such as "poisoning" (e.g., "unintentional food poisoning") or "hanging" (e.g., "hanging out in a pub") can be ambiguous, Google's alarm bells should ring louder, especially at high-risk times. The proposed two-step approach can contribute to the optimization of the detection of vulnerable individuals. Finally, one should bear in mind, however, that the cost of a false alarm is substantially less than the cost of a miss, which may ultimately be the death of a vulnerable member of our global society.

References

Beauchamp, G., Ho, M., & Yin, S. (2014). Variation in suicide occurrence by day and during major American holidays. *The Journal of Emergency Medicine*, *46*, 776–781. doi:10.1016/j.jemermed.2013.09.023

Biddle, L., Donovan, J., Hawton, K., Kapur, N., & Gunnell, D. (2008). Suicide and the Internet. *British Medical Journal*, *336*, 800–802. doi:10.1136/bmj.39525.442674.AD

Christodoulou, C., Douzenis, A., Papadopoulos, F. C., Papadopoulou, A., Bouras, G., Gournellis, R., & Lykouras, L. (2012). Suicide and seasonality. *Acta Psychiatrica Scandinavica*, *125*(2), 127–146. doi:10.1111/acps.2011.125.issue-2

Cohen, N. (2010). *'Suicide' query prompts Google to offer hotline*. Retrieved from http://nyti.ms/1RkKOYf

Draper, J., Murphy, G., Vega, E., Covington, D., & McKeon, R. (2015). Helping callers to the National Suicide Prevention Lifeline who are at imminent risk of suicide: The importance of active engagement, active rescue, and collaboration between crisis and emergency services. *Suicide and Life-Threatening Behavior*, *45*, 261–270. doi:10.1111/sltb.2015.45.issue-3

Fond, G., Gaman, A., Brunel, L., Haffen, E., & Llorca, P. (2015). Google trends: Ready for real-time suicide prevention or just a Zeta-Jones effect? An exploratory study. *Psychiatry Research*, *228*, 913–917. doi:10.1016/j.psychres.2015.04.022

Gabennesch, H. (1988). When promises fail: A theory of temporal fluctuations in suicide. *Social Forces*, *67*, 129–145. doi:10.1093/sf/67.1.129

Ginsberg, J., Mohebbi, M., Patel, R., Brammer, L., Smolinski, M., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*, 1012–1014. doi:10.1038/nature07634

Gunn, J., & Lester, D. (2013). Using Google searches on the Internet to monitor suicidal behavior. *Journal of Affective Disorders*, *148*, 411–412. doi:10.1016/j.jad.2012.11.004

Haim, M., Arendt, F., & Scherr, S. (2016). Abyss or shelter? On the relevance of Web search engines' search results when people Google for suicide. *Health Communication*. Advance online publication. doi:10.1080/10410236.2015.1113484

Historical Society of Pennsylvania (2016). *Letter from Sir Isaac Newton to Robert Hooke*. Retrieved from http://digitallibrary.hsp.org/index.php/Detail/Object/Show/object_id/9285.

Joiner, T., & Rudd, M. (2002). *Suicide science: Expanding the boundaries*. Boston, MA: Kluwer.

Lazer, D. (2015). The rise of the social algorithm. *Science*, *348*, 1090–1091. doi:10.1126/science.aab1422

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science*, *343*, 1203–1205. doi:10.1126/science.1248506

Lester, D., & Rogers, J. (2012). *Crisis intervention and counseling by telephone and the Internet*. Springfield, IL: Charles C Thomas.

Mann, J., Apter, A., Bertolote, J., Beautrais, A., Currier, D., Haas, A., … Handin, H. (2005). Suicide prevention strategies: A systematic review. *Journal of the American Medical Association*, *294*, 2064–2074. doi:10.1001/jama.294.16.2064

Mehlum, L. (2000). The Internet, suicide, and suicide prevention. *Crisis*, *21*, 186–188. doi:10.1027//0227-5910.21.4.186

Niederkrotenthaler, T., Voracek, M., Herberth, A., Till, B., Strauss, M., Etzersdorfer, E., … Sonneck, G. (2010). Role of media reports in completed and prevented suicide: Werther v. Papageno effects. *British Journal of Psychiatry*, *197*, 234–243. doi:10.1192/bjp.bp.109.074633

Nuti, S., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R., Chen, S., & Murugiah, K. (2014). The use of Google Trends in health care research: A systematic review. *PLoS ONE*, *9*, e109583. doi:10.1371/journal.pone.0109583

Pew Research Center. (2012). *Internet & American Life Project: Search engine use 2012*. Retrieved from http://www.pewinternet.org/2012/03/09/search-engine-use-2012

Preis, T., Moat, H., Stanley, H., & Bishop, S. (2012). Quantifying the advantage of looking forward. *Scientific Reports*, *2*, 350. doi:10.1038/srep00350

Shoemaker, P., & Vos, T. (2009). *Gatekeeping theory*. New York, NY: Routledge.

Wasserman, D. (2016). *Suicide: An unnecessary death*. Oxford, England: University Press.

Zeiger, R. (2010). *Helping you find emergency information when you need it*. Retrieved from http://blog.google.org/2010/11/helping-you-find-emergency-information.html

Appendix

Data Collection

All analyses used averaged scores across the observation time period to partial out idiosyncrasies of single years. On February 13, 2016, we entered the search term and downloaded raw data for "poisoning" from Google Trends. Our search was limited to the United States from December 1, 2009, to January 31, 2016 (see the following).

Season

Raw data were collected and merged for the whole period of 6 years between January 2010 and December 2015. As Google Trends does not provide data for each day when specifying the whole year, we requested query shares per week. For each downloaded week, information about the season was manually entered: winter (December 21), spring (March 21), summer (June 21), and fall (September 23). We used the following rule to code weeks including the target dates:

2010-06-20, 2010-06-26, = > spring (because not all days represent summer)

2010-06-27, 2010-07-03, = > summer (because all days represent summer)

Day of the Week

Raw data were collected in 3-month periods (January–March, April–June, July–September, October–December). Again, this was done because Google Trends does not provide data for each day when specifying the whole year. We calculated descriptive statistics (mean, standard deviation, *N*) for each day and calculated the 95% confidence intervals based on this output.

New Year's Day

Raw data were collected in 2-month periods (December–January: 2009/2010, 2010/2011, 2011/2012, 2012/2013, 2013/2014, 2014/2015, and 2015/2016). Query share was assessed at each day (December 1–January 31). We calculated statistics for each day ranging from −31 (December 1) through 0 (January 1) to 30 (January 31).

Thanksgiving

Raw data were collected in 2-month periods (November–December: 2010, 2011, 2012, 2013, 2014, and 2015). Query share was assessed at each day (November 1–December 31). Thanksgiving is always celebrated at the third Thursday in November. Raw time series were used to calculate the mean for each day, ranging from −20 through 0 (Thanksgiving) to 27. The specific start and end date differed each year, because Thanksgiving was on different dates (2010: November 25; 2011: November 24; 2012: November 22; 2013: November 28; 2014: November 27; 2015: November 26). We decided to start on −20 because all years provide data for this day.